GBE

The Evolution of the GPCR Signaling System in Eukaryotes: Modularity, Conservation, and the Transition to Metazoan Multicellularity

Alex de Mendoza[1,2], Arnau Sebé-Pedrós[1,2], and Iñaki Ruiz-Trillo[1,2,3,*]

[1]Institut de Biologia Evolutiva (CSIC-Universitat Pompeu Fabra) Passeig Marítim de la Barceloneta, Barcelona, Spain

[2]Departament de Genètica, Universitat de Barcelona, Spain

[3]Institució Catalana de Recerca i Estudis Avançats (ICREA) Passeig Lluís Companys, Barcelona, Spain

*Corresponding author: E-mail: inaki.ruiz@multicellgenome.org, inaki.ruiz@ibe.upf-csic.es.

Accepted: February 19, 2014

Abstract

The G-protein-coupled receptor (GPCR) signaling system is one of the main signaling pathways in eukaryotes. Here, we analyze the evolutionary history of all its components, from receptors to regulators, to gain a broad picture of its system-level evolution. Using eukaryotic genomes covering most lineages sampled to date, we find that the various components of the GPCR signaling pathway evolved independently, highlighting the modular nature of this system. Our data show that some GPCR families, G proteins, and regulators of G proteins diversified through lineage-specific diversifications and recurrent domain shuffling. Moreover, most of the gene families involved in the GPCR signaling system were already present in the last common ancestor of eukaryotes. Furthermore, we show that the unicellular ancestor of Metazoa already had most of the cytoplasmic components of the GPCR signaling system, including, remarkably, all the G protein alpha subunits, which are typical of metazoans. Thus, we show how the transition to multicellularity involved conservation of the signaling transduction machinery, as well as a burst of receptor diversification to cope with the new multicellular necessities.

Key words: arrestin, phosducin, Ric8, GRK, heterotrimeric G protein complex.

Introduction

A molecular system to receive and transduce signals from the environment or from other cells is key to multicellular organisms (Gerhart 1999; Pires-daSilva and Sommer 2003), although molecular signaling pathways are not uniquely required within a multicellular context. Unicellular species face similar signaling needs as multicellular organisms, dealing with a changing environment and, in some cases, coordinating different cells (e.g., density sensing) (Crespi 2001; King 2004; Rokas 2008).

Both animals (metazoans) and plants have evolved complex signaling pathways to govern their embryonic development, and, according to current genomic data, some of these pathways appear to be specific to either metazoans or plants. This is the case of the metazoan-specific WNT and Hedgehog signaling pathways (Ingham et al. 2011; Niehrs 2012) and the land plant-specific auxin and cytokinin (Rensing et al. 2008). Other signaling pathways, such as the metazoan Notch pathway, have instead been assembled from various, more ancient components by domain shuffling (Gazave et al. 2009). Finally, other signaling pathways were already present in the unicellular ancestors and were subsequently co-opted for multicellular functions. A good example are the receptor tyrosine kinases, which emerged and expanded in unicellular holozoans (i.e., choanoflagellates and filastereans), and were later recruited for developmental control in metazoans (King et al. 2008; Manning et al. 2008; Suga et al. 2012). The reuse of previously assembled signaling systems is indeed an important mechanism of signaling pathway co-option in multicellular lineages (King et al. 2008).

One of the major eukaryotic signaling pathways is the G-protein-coupled receptors (GPCRs) and their associated signaling modules (Fritz-Laylin et al. 2010; Anantharaman et al. 2011; Krishnan et al. 2012), which are conserved from excavates to animals. GPCRs are involved in many processes apart

from developmental control, such as cell growth, migration, density sensing, or neurotransmission (Bockaert and Pin 1999; Pierce et al. 2002; Rosenbaum et al. 2009). GPCRs are able to sense a wide diversity of signals, including proteins, nucleotides, ions, and photons. Structurally, GPCRs have a seven transmembrane (TM) domain (they are also known as 7TM receptors), which forms a ligand-binding pocket in the extracellular region, and a cytoplasmic G-protein-interacting domain (Pierce et al. 2002; Lagerström and Schiöth 2008), which binds to G proteins to mediate intracellular signaling. G proteins form a heterotrimeric complex that is disassembled when activated by the GPCR, which acts as a guanidine exchange factor (GEF), and transduce the signal into downstream effectors (Oldham and Hamm 2008). The G protein heterotrimeric complex has three different subunits of distinct evolutionary origin, *alpha*, *beta,* and *gamma*. G protein heterotrimeric signaling is, in turn, regulated by various proteins families, including Regulators of G protein Signaling (RGS) and GoLoco-motif-containing proteins (Pierce et al. 2002; Siderovski and Willard 2005; Wilkie and Kinch 2005). The combination of GPCR, G proteins, and their regulators results in many diverse signaling outputs.

Besides the classic GPCR-G protein signaling system described earlier, there are alternative upstream and downstream molecules (fig. 1). For instance, seven TM receptors associated to RGS antagonize "self-activated" G *alpha* proteins in some lineages, acting as GTPase-accelerating proteins (GAP) receptors (Urano et al. 2012; Bradford et al. 2013). In plants, a single pass TM receptor has been recently characterized to interact with G *alpha* proteins (Bommert et al. 2013). Moreover, monomeric G protein *alpha* activation by Ric 8 (resistance to inhibitors of cholinesterase 8) is also GPCR independent (Wilkie and Kinch 2005; Hinrichs et al. 2012), and *beta/gamma* heterodimers are regulated via phosducins (Willardson and Howlett 2007). Complementarily, GPCRs can perform downstream signaling independently of G proteins by G protein-coupled receptor kinases (GRKs), Arrestins, and Arrestin domain-containing proteins (ARDCs) (Gurevich VV and Gurevich EV 2006; Reiter and Lefkowitz 2006; DeWire et al. 2007; Liggett 2011; Shenoy and Lefkowitz 2011).

Most of the proteins involved in the GPCR signaling pathway have previously been analyzed as single units in various phylogenetic contexts (Blaauw et al. 2003; Fredriksson and Schiöth 2005; Alvarez 2008; Oka et al. 2009; Anantharaman et al. 2011; Krishnan et al. 2012; Mushegian et al. 2012; Bradford et al. 2013). However, not much attention has been paid to the system-level evolution of the entire pathway, and given the modularity of the system, it is important to investigate its evolution from a global point of view.

In this article, we provide an update on the evolutionary histories of all components of the GPCR signaling system using a genomic survey that includes representatives of all eukaryote supergroups. We analyze the modular structure of the signaling pathway and show how different parts of the system coevolved in complementary or independent patterns. We also reconstruct the GPCR signaling system in the last common ancestor of eukaryotes (LECA) and track its evolution in various lineages. Finally, we analyze the evolution of the system in the transition from unicellular ancestors to metazoans. We observe strong conservation in the pathway components associated with cytoplasmic signaling transduction, whereas receptors radiated extensively in metazoans, becoming one of the largest gene families in metazoan genomes (Fredriksson and Schiöth 2005). The dissimilarity between the pattern of evolution in preadapted signaling transduction machinery and active diversification of receptors provides clues on how key innovations in metazoan complexity could have evolved from pre-existing machineries.

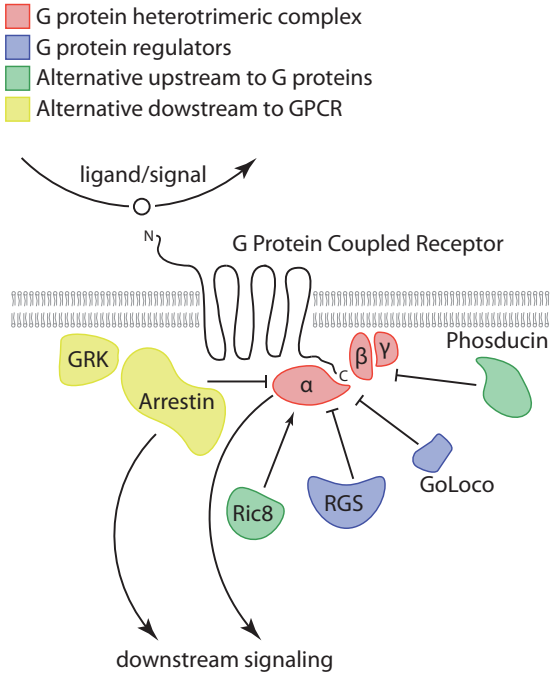

Fig. 1.—Schematic representation of the GPCR signaling pathway. Protein families belonging to similar functional categories are grouped as specified in the color legend.

Materials and Methods

Taxon Sampling and Data Gathering

The 75 publicly available genomes used in this study were downloaded from databases at National Center for Biotechnology Information, The Joint Genome Institute, and The Broad Institute. Data from some unicellular holozoan species come from RNAseq sequenced in-house (*Pirum gemmata*, *Abeoforma whisleri,* and *Corallochytrium limacisporum*) or from The Broad Institute "Origin of Multicellularity

Database" (*Ministeria vibrans* and *Amoebidium parasiticum*). The RNAseq transcripts were translated into six frames.

All the protein domains that are components of the GPCR signaling machinery were selected from the literature and the PFAM database (Punta et al. 2012). All proteomes were scanned using PfamScan with PFAM A version 26 as query and selecting the gathering threshold option. Gathering threshold is important in the case of GPCRs, because it helps to disambiguate between different GPCR families by selecting the most significant hit. Additionally, PfamScan gathering threshold avoids the spurious partial hits typical of TM proteins and is a conservative approach to minimize false positives that may arise with other more sensitive methods (Punta et al. 2012). General distribution patterns were obtained by counting proteins with at least one domain belonging to the GPCR signaling machinery present in the PfamScan proteomic outputs. The same files were used to obtain multidomain architectures, with the exception of the TM domains analyzed in RGS proteins, which were obtained using the TMHMM software (Krogh et al. 2001). In the case of G protein *gamma* subunits, additional TBlastN searches against reference genomes were performed to avoid false negatives using bikont and opisthokont sequences as query. Gene loss is very difficult to assess due to the different degrees of incompleteness of the available genomes. To overcome this problem we used, when possible, more than one taxa for each eukaryotic clade. Transcriptome data do not account for gene loss, as genes can be missed due to low expression, but in our data set most species with transcriptomic data have sister species with genome sequence available.

Heatmaps, Principal Component Analysis, and Parsimony Reconstruction

Heatmaps were built using R heatmap.2 function, from the gplots package. Principal component analysis (PCA) was carried out using the built-in R prcomp function, with scaling and a covariance matrix, and were plotted using the R bpca package. We assumed Dollo parsimony to infer ancestral gains and secondary loss reconstructions in figure 2 using Mesquite (Maddison WP and Maddison DR 2011).

Phylogenetic Analyses

Arrestins/ARDCs, Ric8, G *alpha* subunit, G *beta* subunit, Phosducin, Kinase, and RGS domains were used for phylogenetic analyses. The alignments were obtained using MAFFT with the L-INS-i option (Katoh and Standley 2013), and these alignments were manually trimmed to avoid ambiguous regions. Seed alignments are available in supplementary file 1, Supplementary Material online. The amino acid model of evolution used for phylogenetic inference was LG, with a discrete gamma distribution of among-site variation rates (four categories) and a proportion of invariable sites.

Maximum likelihood analyses were performed using RAxML version 7.2.6. (Stamatakis 2006). The best-tree topology depicted in the figures was obtained by selecting the best tree out of 100 replicates. Bootstrap support was obtained using 100 bootstrap replicates of the same alignment. Bayesian inference trees were inferred using PhyloBayes v3.3 (Lartillot et al. 2009). The resulting tree and posterior probabilities were obtained when two parallel runs converged (tracecomp standard values), after surpassing at least 500.000 generations. The runs were sampled every 100 generations, and the burn-in was established using a bpcomp maxdiff < 0.3.

Results

GPCR Families: Ancient Origins and Architecture Diversifications

A widely accepted classification of the metazoan GPCR complement is the GRAFS system, which is based on both phylogeny and structural similarity (Fredriksson et al. 2003; Fredriksson and Schiöth 2005; Lagerström and Schiöth 2008; but see Pierce et al. [2002] for an alternative classification). The GRAFS system divides GPCRs into five different families, Glutamate (also known as Class C), Rhodopsin (Class A), Adhesion (Class B), Secretin (class B), and Frizzled (Class F). This system can be extended to GPCR types described in nonmetazoans, including the cAMP (Class E), ITR-like and GPR-108-like families, as well as several lineage-specific receptor families such as insect odorant receptors, nematode chemoreceptors, or vertebrate vomeronasal receptors (Nordström et al. 2011). Fungi also have well-defined GPCR families such as Ste2 and Ste3 (both included in Class D), and Git3 and plant Abscisic acid receptors are also thought to be GPCRs (Plakidou-Dymock et al. 1998; Tuteja 2009; Krishnan et al. 2012). Most GPCR families are associated with a characteristic PFAM domain (Fredriksson et al. 2003; Fredriksson and Schiöth 2005; Lagerström and Schiöth 2008).

First, we assessed the presence and abundance of GPCR family domains in diverse eukaryotic genomes (see fig. 2 for a complete taxon sampling). Our data show that the distribution of GPCR families in eukaryotes follows two distinct evolutionary patterns. Some families are pan-eukaryotic, whereas others are biased toward amorpheans (unikonts). For instance, GRAFS are more abundant in amorpheans, especially in metazoans, although some (Glutamate, Adhesion/Secretin, and Rhodopsin) are also observed in some bikonts. Other families, such as cAMP receptors, Git3, ITR-like, GPR-108-like, and Abscisic acid receptors, are found in similar abundance among eukaryotes. Interestingly, non-GRAFS GPCR families are never expanded in any species (<10 members in all genomes). We also surveyed the taxonomically restricted metazoan families, and although we found chemosensory receptors (7tm_7) and the Serpentine type chemoreceptors Srw

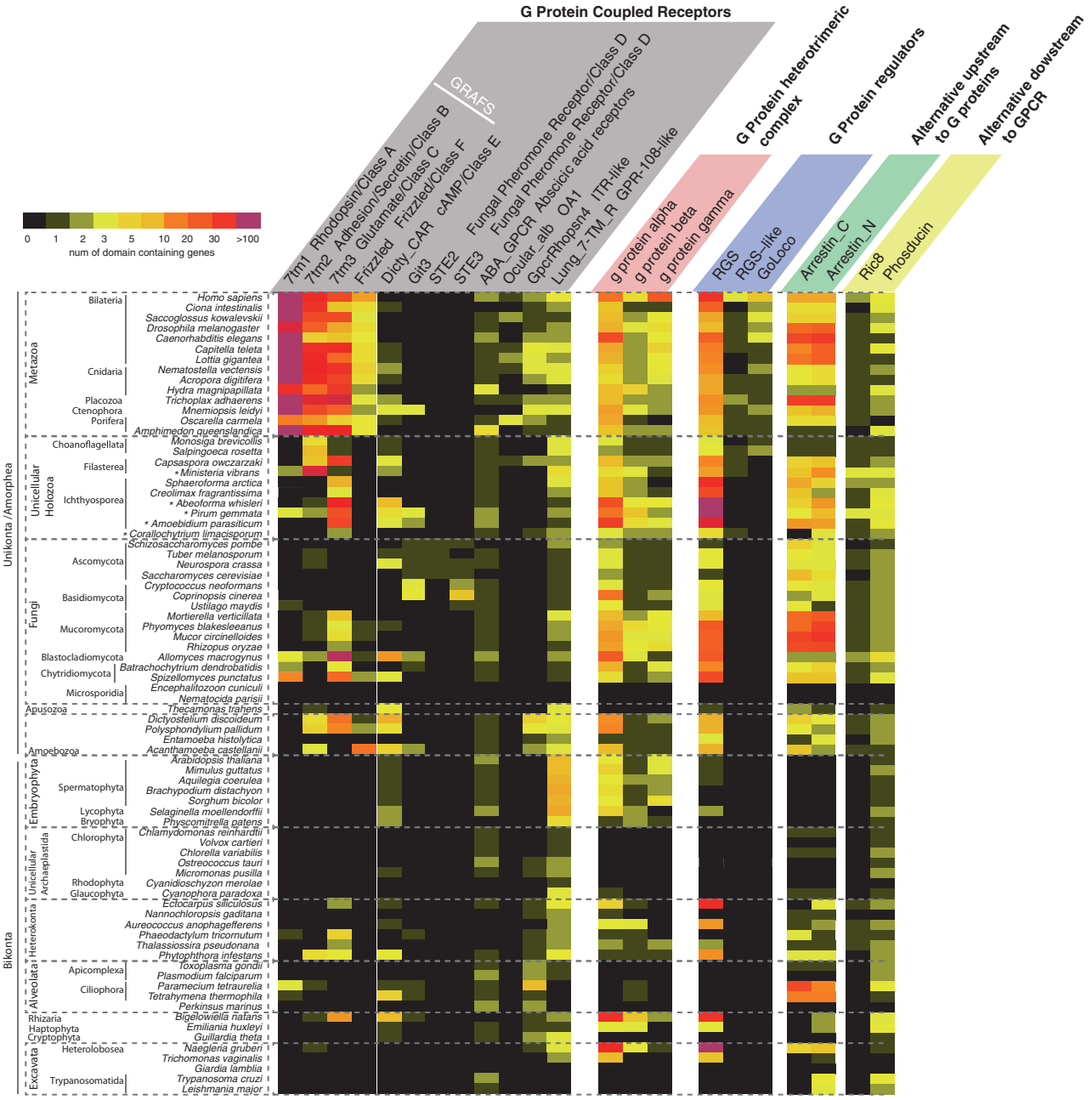

Fig. 2.—Distribution and abundance of GPCR signaling components in 78 eukaryotic genomes. Numbers and abundance of domain containing proteins are depicted according to the color legend in the upper left, being black absence of the given domain in a given species. Yellow color indicates smaller amounts, whereas the scale to purple indicates more abundance. The various domains are grouped into functional modules specified in figure 1, as shown in the schema at the bottom right. Species marked with an asterisk are only covered by RNA-seq data, therefore gene absence is not definitive. The original numbers of the heatmap are available at supplementary table S1, Supplementary Material online.

and Srx in some previously unreported metazoan genomes, none were observed in nonmetazoan eukaryotes (supplementary figs. S1 and S2, Supplementary Material online), with exception of OA1 (Ocular Albinism receptor), which is specific to metazoans and *Capsaspora owczarzaki*. These results indicate that most GPCR families have ancient origins in the last eukaryotic common ancestor.

Diversification of ancient GPCR families is usually accompanied by architectural diversification of the N-terminal protein domain (Lagerström and Schiöth 2008). Thus, we analyzed the architectural diversity of each GPCR family in each genome and observed two types of GPCRs in terms of N-terminal domain diversity (diversifying vs. nondiversifying in supplementary fig. S2, Supplementary Material online). Some,

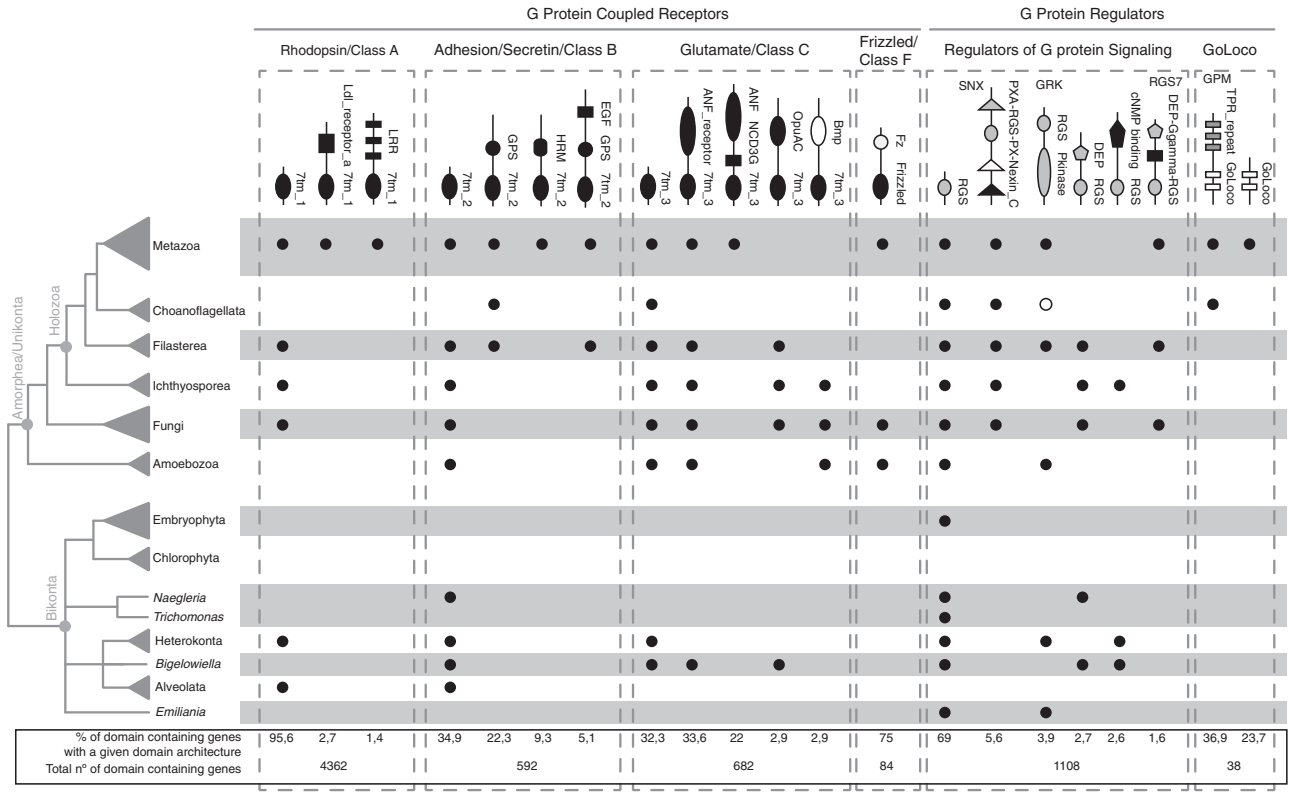

Fig. 3.—Conservation of the domain architecture of different GPCR signaling components across eukaryotic genomes. A black dot indicates the presence of a given domain architecture. A white dot refers to similar domain architecture, Tyrosine Kinase instead of Serine/Threonine kinase in the case of Choanoflagellate GRK-like genes. For simplicity, only the most common architectures are shown. The percentage of genes found with a given architecture within a family is indicated at the bottom part of the table, as well as the total number of genes within the family. GPM in the first column of GoLoco motif containing proteins stands for G Protein Modulator/Rapsynoid. The complete domain architectures of the GPCR signaling system components are found in supplementary figures S3, S4, and S6, Supplementary Material online.

such as Glutamate, Adhesion/secretin, and, to a lesser extent, Rhodopsin, are susceptible to the recruitment of new domains in the N-terminal region, especially in Metazoa, whereas others, such as cAMP, Git3, OA1, Abscisic acid receptors, GPR108-like, and ITR-like, have substantially lower diversity of protein domains at the N-terminal. This result suggests that some GPCR families have functional constraints, whereas others are prone to diversify through recruitment of concurrent domains.

To gain further insights into domain diversification, we searched for evolutionary conservation of specific protein domain architectures (fig. 3) and found that some architectures are highly conserved across lineages. For example, Glutamate receptors (7tm_3) have protein domain configurations that are conserved in distant eukaryotic lineages, including those with Venus Flytrap module (ANF_receptor), OpuAC, or Bmp domains (fig. 3). Additionally, several nonmetazoan species have diversified their own species-specific configurations of glutamate receptors (supplementary fig. S3, Supplementary Material online). The Adhesion family is also quite structurally diverse, especially in metazoans and, to a lesser extent, unicellular holozoans (supplementary fig. S3, Supplementary Material online). Similarly, the Rhodopsin family is architecturally diversified, mainly in metazoans. Finally, Fz-Frizzled, RpkA (cAMP-PIP5K domain architecture), and Git3-Git3_C protein domain architectures could be identified in several eukaryotic genomes (supplementary fig. S4, Supplementary Material online), expanding the previous distribution of those architectures at LECA or at the root of Amorphea/Unikonta. Remarkably, most of the GPCR complex architectures belong to GRAFS families and are mostly diversified and conserved within metazoans.

Heterotrimeric G Protein Complex

GPCRs typically signal through G proteins. In an inactive state, the three G protein subunits (*alpha*, *beta,* and *gamma*) form a heterotrimeric complex (Pierce et al. 2002; Oldham and Hamm 2008) (fig. 1). When a ligand activates a GPCR, it acts as a GEF, promoting GDP to GTP exchange in the G *alpha* subunit. This exchange alters G *alpha* subunit conformation and promotes the disaggregation of the heterotrimeric

complex. The active G *alpha* subunit and an active dimer of *beta* and *gamma* subunits mediate further downstream signaling through various effectors (Milligan and Kostenis 2006; Oldham and Hamm 2008). G *alpha* is a low-efficiency GTPase, whereas G *beta* has various WD-40 repeats (PF00400) and G *gamma* is a small protein containing a conserved domain (Milligan and Kostenis 2006; Anantharaman et al. 2011).

Using the signature domains of each G protein, we surveyed our data set to find their general distribution patterns and found that the abundance of each subunit varies markedly across eukaryotes and that some taxa have lost these three subunits entirely (Anantharaman et al. 2011). G protein *alpha* is the most susceptible to diversification, and, interestingly, *beta* and *gamma* subunits have multiple copies in G *alpha* rich species. Although combination of the three elements is important for signaling plasticity, G *alpha* is the most evolutionarily dynamic of the three G proteins.

To gain further insights into the evolution of G *alpha* proteins, we performed phylogenetic analyses using our eukaryotic data set (fig. 4), and the resulting tree shows that several groups have lineage-specific diversifications, such as those in *Naegleria gruberi*, *Bigelowiella natans*, and *Emiliania huxleyi*. The opisthokonts have a diverse but conserved repertoire of G *alpha* proteins. Fungi have four distinct paralogs (GPA-1 to 4) present in Ascomycota, Basidiomycota, Mucoromycotina, and Chytridiomycetes (families reviewed in Li et al. 2007) and therefore were most likely present in the fungal ancestor. Holozoa also have four ancient paralogs, Gαs, Gαq/12/13, Gαi/o, and Gαv (described for Metazoa in Oka et al. 2009). It is worth mentioning that all the metazoan G *alpha* families are conserved in the unicellular relatives of Metazoa, indicating that they originated prior to the diversification of metazoans from the rest of holozoans.

We also identified a new and divergent family of holozoan G *alpha* subunits that branches out from the Opisthokonta clade, comprising *Nematostella vectensis*, *Lottia gigantea,* and other holozoans (fig. 4). Additionally, we observed a cluster of conserved G *alpha* subunits in several distant eukaryotic lineages (what we call conserved-eukaryotic group I): Ichthyosporea, *Allomyces macrogynus,* and dictyostelids within the Amorphea, and *B. natans* and *Ectocarpus siliculosus* within the bikonts. It is likely that this particular family originated in the LECA and was lost many times during eukaryotic evolution.

We also performed a phylogenetic analysis of eukaryotic *beta*-subunits, to compare the evolutionary histories of *alpha* and *beta* (supplementary fig. S5, Supplementary Material online). Our tree shows that holozoans have a particular ancient duplication, Gβ1-4 and Gβ5, with the more derived Gβ5 known to interact with G *gamma*-like subunits, such as RGS7 (Sondek and Siderovski 2001; Anderson et al. 2009), a multidomain protein that contains a G *gamma* domain. We identified RGS7 in both chytrid fungi and holozoans (fig. 3 and supplementary fig. S6, Supplementary Material online), and therefore, the ancient duplication of G protein *beta* and its partner, RGS7, are ancient features of holozoans.

Regulatory Proteins: RGS and GoLoco

Regulation of G proteins is a key step in GPCR signaling that involves two main protein families, RGS and GoLoco motif-containing proteins (Siderovski and Willard 2005; Wilkie and Kinch 2005). RGS proteins act as GAP, turning GTP into GDP and thereby promoting the formation of the G protein heterotrimeric complex and completing G *alpha* signaling (Siderovski and Willard 2005). Nevertheless, not all RGS domains act as GAP proteins in G protein signaling, and some have lost their GAP activity and have developed scaffolding functions (Anantharaman et al. 2011). GoLoco-motif-containing proteins (also known as G protein regulators) act as guanine dissociation antagonists, inhibiting the dissociation of the heterotrimeric complex by binding to G *alpha*-GDP and blocking downstream signal transduction (Siderovski and Willard 2005).

We traced the distribution and abundance of RGS and GoLoco motif proteins in eukaryotes and found that RGS is present in many different eukaryotes, mainly coinciding with the presence of heterotrimeric subunits (fig. 2). The number of RGS varies from one single copy in some taxa to numerous copies in other lineages. For example, some eukaryotes such as *N. gruberi* (229), *B. natans* (39), *E. siliculosus* (47), or the ichthyosporeans (22–119) have more RGS proteins than *Homo sapiens* (34), whereas other multicellular lineages such as plants possess only one copy. In contrast, the GoLoco motif appears to be exclusive to metazoans and choanoflagellates (figs. 2 and 3), and although its copy number may vary, it is less abundant than RGS. Therefore, our data show that the eukaryotic RGS system underwent independent radiations in lineages including amoebozoans, ichthyosporeans, heteroloboseans, and rhizarians, whereas GoLoco is a later development that originated prior to the divergence of choanoflagellates and metazoans.

We then examined the architectural configurations of RGS proteins, as they are known to combine with many other protein domains (Siderovski and Willard 2005; Anantharaman et al. 2011). Our survey shows that species with distant phylogenetic relationships to each other evolved their own architectural repertoires and generally have unique configurations that are not found elsewhere (supplementary fig. S6, Supplementary Material online). Moreover, many configurations evolved independently, recruiting the same domain in different configurations. For example, DEP, cNMP binding, Kinases, Rho GTPase, Leucine-Rich Repeat, START, and Ankyrin repeats are all present in various combinations in RGS genes from divergent taxa (shown in red in supplementary fig. S6, Supplementary Material online). However, some complex multidomain architectures are evolutionary

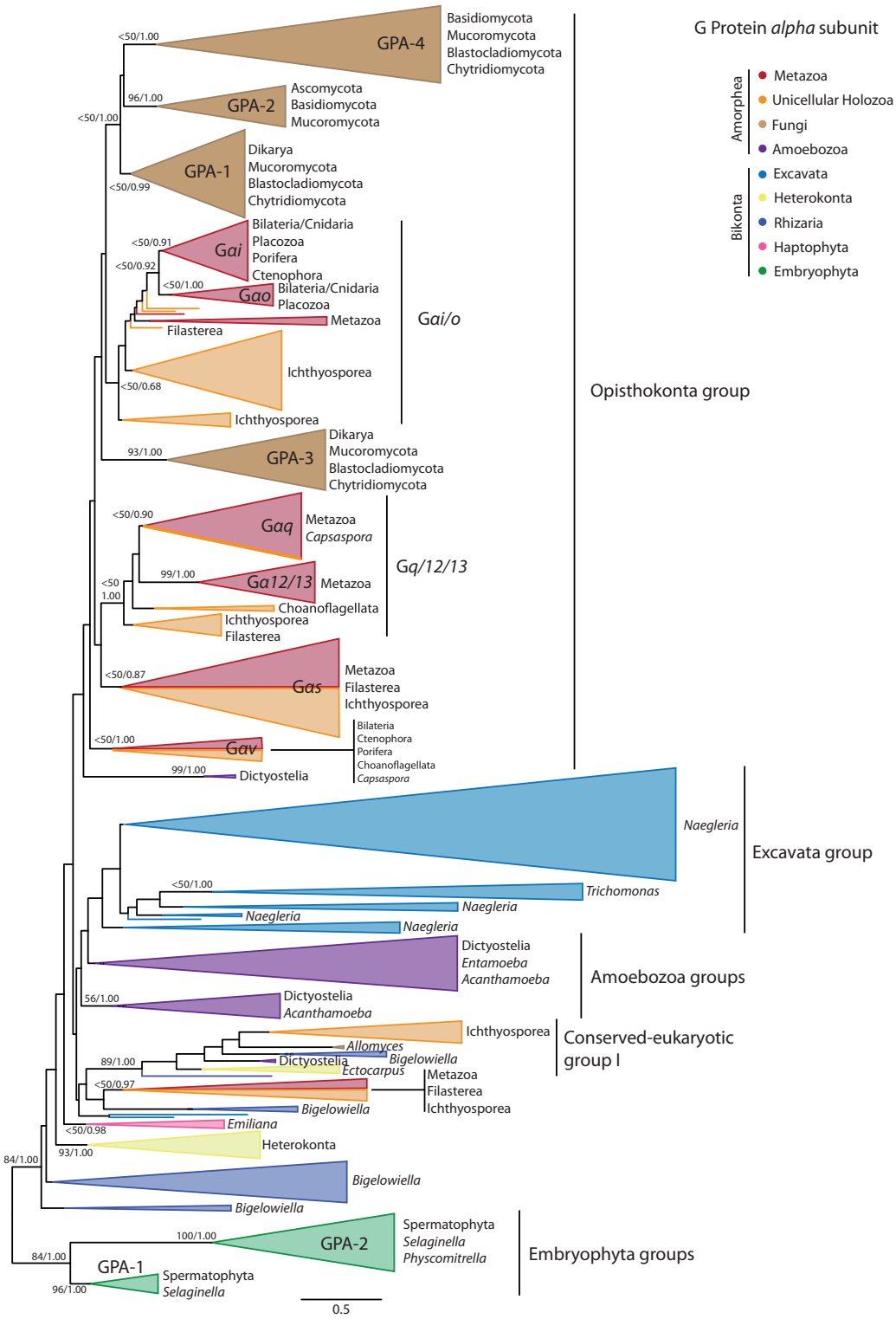

Fig. 4.—Maximum likelihood (ML) phylogenetic tree inferred by the G protein *alpha* subunit. Different eukaryotic lineages are represented by a color code depicted in the legend. Within the gene family clades, the specific taxonomic groups which comprise eukaryotic lineages represented in that clade (i.e., eumetazoans and placozoans) are shown on the right. Nodal supports indicate 100-replicate ML bootstrap support and Bayesian posterior probability (BPP). Supports are only shown for nodes recovered by both ML and Bayesian inference, with BPP > 0.9.

conserved (fig. 3 and supplementary fig. S6, Supplementary Material online). For example, opisthokonts share some common RGS architectures, namely Sorting Nexins (SNX13/14/25) and the previously mentioned RGS7. Additionally, the RGS-like domain, typical of PDZ-RhoGEF, is an innovation of Holozoa (fig. 2), whereas RGS12 and Axin are metazoan innovations (supplementary fig. S6, Supplementary Material online). Our results emphasize that metazoans and their unicellular relatives have conserved elements of RGS complement, which is quite susceptible to diversification through domain rearrangements.

Of specific interest are RGS proteins with TM domains (Anantharaman et al. 2011; Urano et al. 2012; Bradford et al. 2013), as they localize to the cell membrane next to heterotrimeric G proteins. We found that in most lineages, RGS is fused to at least one TM domain (supplementary fig. S7, Supplementary Material online) but in apusozoans, amoebozoans, and haptophytes. In plants and other eukaryotes, RGS domains have been observed together with 7TM organizations, somehow resembling a GPCR but with the opposite effect on G proteins (Urano et al. 2012; Bradford et al. 2013). Many bikonts possess 7TM-RGS architectures, but we found that chytrid fungi, filastereans, and ichthyosporeans also have this type of receptors, whereas metazoans do not, suggesting that metazoans dispensed with GAP TM signaling and restricted on typical GPCR signaling.

GoLoco motif-containing proteins are also part of multidomain proteins. Our results show that choanoflagellates have a unique configuration (SH2-GoLoco) and a shared architecture with metazoans, G-protein-signaling modulator/Rapsynoid (fig. 3 and supplementary fig. S6, Supplementary Material online). Metazoans have some additional conserved architectures, such as RGS12/RGS14 and Rap1GAP (supplementary fig. S6, Supplementary Material online).

Upstream Alternative Regulators: Ric8 and Phosducin

Ric8 is a long domain that acts as a GEF, activating G *alpha* subunits in the absence of GPCR signaling, or as a chaperone to stabilize G *alpha* (Hinrichs et al. 2012; Chan et al. 2013). Ric8-mediated activation of monomeric G *alpha* is involved in development and signaling in metazoans, fungi, and *Dictyostelium* (Hinrichs et al. 2012; Kataria et al. 2013). Although we found Ric8 in almost all amorpheans, suggesting it was secondarily lost in some species (Microsporidia, *Thecamonas trahens,* and *Entamoeba histolytica*) (fig. 2), it is rare in bikonts and found only in a small number of Heterokonta. The presence of Ric8 in only a few heterokonts could be explained by horizontal gene transfer, although our phylogenetic analysis does not support this hypothesis (supplementary fig. S8, Supplementary Material online), but suggests instead that Ric8 was present in the LECA and secondarily lost in many eukaryotic lineages.

Phosducins belong to a small and ancient gene family, Phosducin-like (Blaauw et al. 2003; Willardson and Howlett 2007), and act as cochaperones of the G *beta/gamma* dimers, allowing normal dimer configuration and transiently inhibiting their junction with G *alpha* (Willardson and Howlett 2007). We performed a phylogenetic analysis of Phosducin-like proteins, and the resulting tree shows three great clades: Phosducin I, Phosducin II/III, and orphan phosducin (supplementary fig. S9, Supplementary Material online). The only one known to interact with G protein *beta* subunits is the Phosducin-I or Phosducin/PhLP1 clade (Blaauw et al. 2003), and this is further reinforced by the fact that most species that have Phosducin I proteins also possess the heterotrimeric *beta* subunit. Conversely, the phosducin-II/III clade includes chlorophyte sequences, a group that lacks G protein signaling. This suggests that proteins belonging to the phosducin-II/III clade have substrates other than G proteins (Willardson and Howlett 2007).

Alternative Signaling Inputs: GRK, Arrestins, and ARDCs

GPCRs can also signal independently of G proteins, which is mainly achieved through interactions with GRKs and Arrestins, where Arrestins can either antagonize G protein signaling or connect GPCRs to other signaling modules (Gurevich VV and Gurevich EV 2006; Reiter and Lefkowitz 2006; DeWire et al. 2007; Shenoy and Lefkowitz 2011). GRKs have an active kinase domain and an inactive RGS domain, which allows it to scaffold with GPCRs. Similar to other kinases (e.g., PKC and PKA), GRKs phosphorylate active GPCR receptors in a process called desensitization, inhibiting the GPCR and allowing Arrestin binding. Arrestin binding promotes receptor internalization by endocytosis, which can result in ubiquitination or recycling of the GPCR (Pierce et al. 2002; Gurevich VV and Gurevich EV 2006; DeWire et al. 2007). Additionally, Arrestins can also act as adaptors for other signal transduction pathways such as MAPK or Akt (DeWire et al. 2007). Thus, understanding the evolutionary dynamics of Arrestin/GRK signaling is key to building a complete picture of GPCR signaling.

We found that GRK-like proteins are present in a reduced subset of eukaryotes, including Holozoa, Dictyostelida, Heterokonta, and Haptophyta (Mushegian et al. 2012) (supplementary figs. S10 and S11, Supplementary Material online). Our phylogenetic analysis supports the duplication of GRKa and GRKb paralog groups at the root of Holozoa, as some sequences belonging to filastereans and ichthyosporeans branch within the GRKa clade (supplementary fig. S10, Supplementary Material online). Nevertheless, using the kinase domain to unravel the evolutionary history of GRKs, some RGS-kinase architectures seem to be convergent, choanoflagellate and dictyostelid RGS are fused to a tyrosine kinase like, instead of being fused to an AGC kinase (supplementary fig. S11, Supplementary Material online). Although the absence of GRK in many GPCR rich genomes is not surprising,

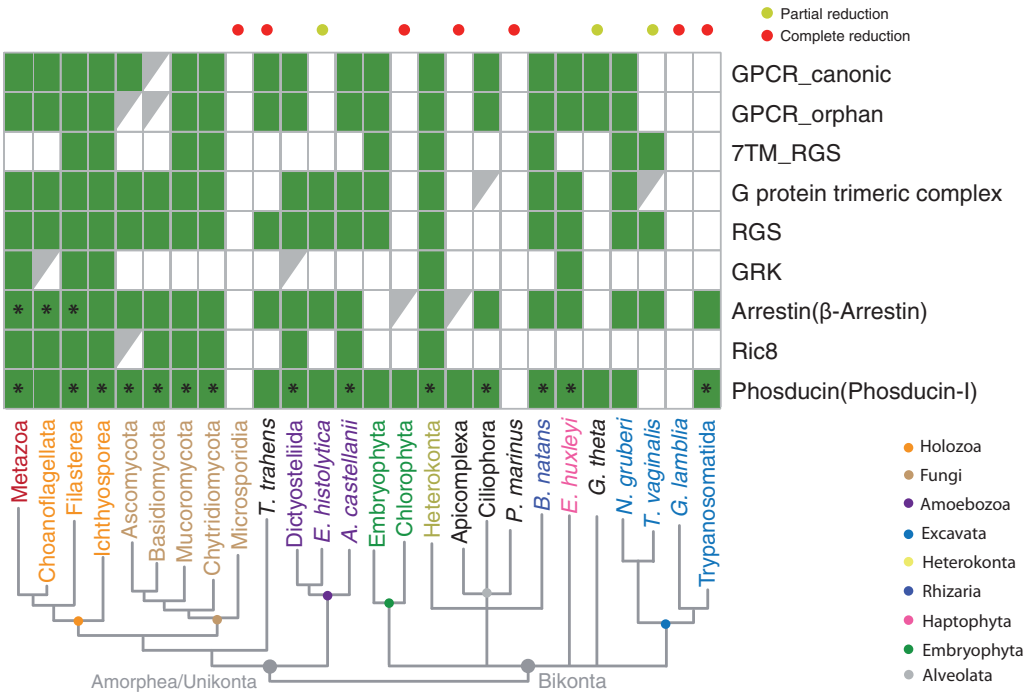

Fig. 5.—Schematic representation of the functional modules in eukaryotic lineages that were analyzed in the study. Green boxes indicate the presence, white the absence, and half-filled squares the presence with some simplification or uncertain affiliation. Asterisks in Arrestin and phosducin rows indicate the presence of orthologs of a subfamily (B-Arrestins and Phosducin-I clade), as discussed in the main text. In the upper part of the table, red dots indicate full reduction of GPCR signaling and green dots indicate severe simplifications but with some conserved functional modules.

because other kinases can replicate this function, holozoans retained two paralogs of this specialized kinase.

Although GRKs are rather scarce in eukaryotes, ARDCs are broadly distributed, and our survey shows that most eukaryotes have a variable number of ARDCs (fig. 3). To gain insights into the evolutionary history of Arrestins and ARDCs, we performed a phylogenetic analysis and identified three major clades, though with low nodal support (supplementary fig. S12, Supplementary Material online). One clade includes metazoan Arrestins, as well as several sequences from unicellular holozoans, making Arrestins a premetazoan invention. The tree also shows a large lineage-specific expansion of ARDCs in Ciliophorans, fungal clades dominated by Mucoromycotina sequences, and the metazoans *Caenorhabditis elegans*, *Drosophila melanogaster,* and *Trichoplax adhaerens*. Interestingly, both Arrestins and ARDCs are known to interact with GPCR (Alvarez 2008), and therefore, their presence and expansion suggest a complementary system to G protein signaling.

GPCR Signaling System

After addressing the evolutionary histories of the various components of GPCRs and their signaling modules, we analyzed them at system level by reducing the diversity of molecules into the main functional categories and analyzing their coevolution (fig. 5). Our data show that holozoans, fungi, amoebozoans, heterokonts/stramenopiles, haptophytes, rhizarians, and heteroloboseans have most of the components of the GPCR signaling system, whereas others, such as *Giardia lamblia* and the miscrosporidians, are completely reduced. Other lineages have retained only a subset of the components involved in GPCR signaling, which challenges general views on the basic mechanics of the system. First, Abscisic acid receptors (PF12430) and GPR-108-like (PF06814) are present in genomes where most of the GPCR signaling system has been lost (such as *Cyanidioschyzon merolae* and *Leishmania major*, see fig. 2), which implies that their role as GPCRs is doubtful, as previously suggested (Maeda et al. 2008; Anantharaman et al. 2011).

Furthermore, there are other taxa in which some GPCRs are present, even though the heterotrimeric complex is absent (or partially absent). For example, the apusozoan *T. trahens,* which lacks heterotrimeric subunits, has four cAMP receptors and one Adhesion receptor, all of which are canonical GPCRs. Similarly, ciliophorans, which only have the G protein subunit *beta*, have members of Rhodopsin, Adhesion, cAMP, and ITR-like receptors. Interestingly, both *T. trahens* and ciliophorans have ARDCs, in high numbers in the latter group, suggesting that ARDCs might provide an alternative link between GPCRs

and other signal transduction pathways in those lineages. This is not the case in *Guillardia theta*, however, which has cAMP and ITR-like GPCRs but neither G proteins nor ARDCs. All these data suggest that GPCRs might be connected to alternative signaling modules other than G proteins.

The modularity of the GPCR signaling system is further supported by the fact that various G protein subunits can be found independently of the other subunits. For example, the G *alpha* subunit, but not the G *beta* and *gamma* subunits, is present in *Trichomonas vaginalis* and *Cyanophora paradoxa*. The former has 7TM-RGS proteins, which, in the absence of GPCR and two of the components of the heterotrimeric complex, may be interacting with other signaling pathways (Bradford et al. 2013), but no RGS is detected in *C. paradoxa*. Ciliophorans only have the G *beta* subunit but have several Phosducin-like genes, which may also imply that ciliophorans have co-opted Phosducin and G protein *beta* into a distinct function. Additionally, *T. trahens* has an RGS protein with no obvious function due to the absence of G *alpha* subunits. Thus, the evolutionary conservation of some components in simplified genomes underpins the modular plasticity of the GPCR signaling system.

We also performed a PCA of our eukaryote data set with the aim of elucidating different evolutionary tendencies (supplementary fig. 13, Supplementary Material online). We observed at least three clusters among eukaryotes that illustrate different patterns of evolution: expansion, simplification, and conservation of the GPCR signaling system. Principal component 1 is principally loaded by the core functional categories of the GPCR signaling system, clustering the most simplified taxa together, including strict parasites such as microsporidians, *G. lamblia*, trypanosomatids, *Perkinsus marinus,* or apicomplexans. Interestingly, many autotrophic lineages, such as Archaeplastida and Cryptophyta, also have a considerably reduced complement of GPCRs. On the other hand, PC2 differentiates between the two kinds of diversification of the GPCR signaling system. In a cluster characterized by the loading of G *alpha* and *beta* subunits, RGS, and cAMP receptors, we find some ichthyosporeans (*A. whisleri*, *P. gemmata* and *Am. parasiticum*), *N. gruberi*, *B. natans,* and *Al. macrogynus*. Metazoans are differentiated in PC2 by the presence of 7tm1, 7tm2, GoLoco, and Frizzled. Therefore, our data indicate that the composition of the GPCR signaling system evolved repeatedly toward a more complex pathway in various eukaryotic lineages. In particular, metazoans developed a more complex system through the expansion of GPCR signaling components.

Reconstruction of GPCR Signaling Components in LECA

We reconstructed the evolutionary stories of the various modules throughout the eukaryotic branch of the tree of life (fig. 2) using the amorphea–bikont root for eukaryotes (Derelle and Lang 2012) and taking into account the topology from the most recent phylogenomic studies (Brown et al. 2012; Burki et al. 2012; Torruella et al. 2012). Our data show that most GPCR families are ancient and that some of the specific architectures of each family can be traced back to the eukaryotic ancestor. Therefore, the LECA already had a complex GPCR signaling system, as well as many other diversified gene families (Derelle et al. 2007; Fritz-Laylin et al. 2010; Wickstead et al. 2010; Grau-Bové et al. 2013). Most interestingly, some complex GPCR architectures are conserved in bikonts (being *B. natans* the major example), contradicting the hypothesis that claims that canonical GPCR signaling through G proteins evolved in amorpheans (Bradford et al. 2013).

Discussion

Our genomic survey and evolutionary reconstruction show that the LECA had a complex repertoire of GPCRs (fig. 6). Independent expansions of the GPCR signaling system occurred in some eukaryotic lineages, and, interestingly, most of the species that have these expansions are unicellular or colonial, such as *B. natans*, *N. gruberi,* and ichthyosporeans (supplementary fig. 13, Supplementary Material online). This supports the view that unicellular lifestyles also require complex signaling machineries (Crespi 2001). In fact, multicellular fungi such as the Basidiomycota *Coprinus cinereus* and the Ascomycota *Tuber melanosporum* have rather simpler complements of GPCRs than other unicellular, including chytrids and Mucoromycotina. Similarly, embryophytes possess a reduced GPCR signaling system. Of course, other signaling pathways are also present in eukaryotes, such as Histidine kinases, Serine/Threonine kinases, or Tyrosine Kinases (Anantharaman et al. 2007; Schaller et al. 2011; Suga et al. 2012), and these can have more important roles in the taxa where GPCR signaling is simplified.

An important conclusion from our work is the modularity of the system. We find that some species have GPCRs without G proteins and vice versa, and we also show how different parts of the GPCR signaling system evolved independently, so that different functional categories involved in the pathway can become simplified without altering the others, as has been hinted at in other studies (Wilkie and Kinch 2005; Anantharaman et al. 2011). In addition, some parts of the pathway have diversified, both in terms of gene number and domain architecture, whereas other elements remain conservative. All this evidence suggests that the system is plastic and that drastic rearrangements can occur without complete loss of functionality. This robustness of eukaryotic signaling systems has been compared with the simpler and more direct signaling systems of prokaryotes (Anantharaman et al. 2007), and indeed modularity is a key feature of eukaryotic signaling pathways, which show great diversity of signaling machineries across different lineages (Anantharaman et al. 2007; Schaller et al. 2011).

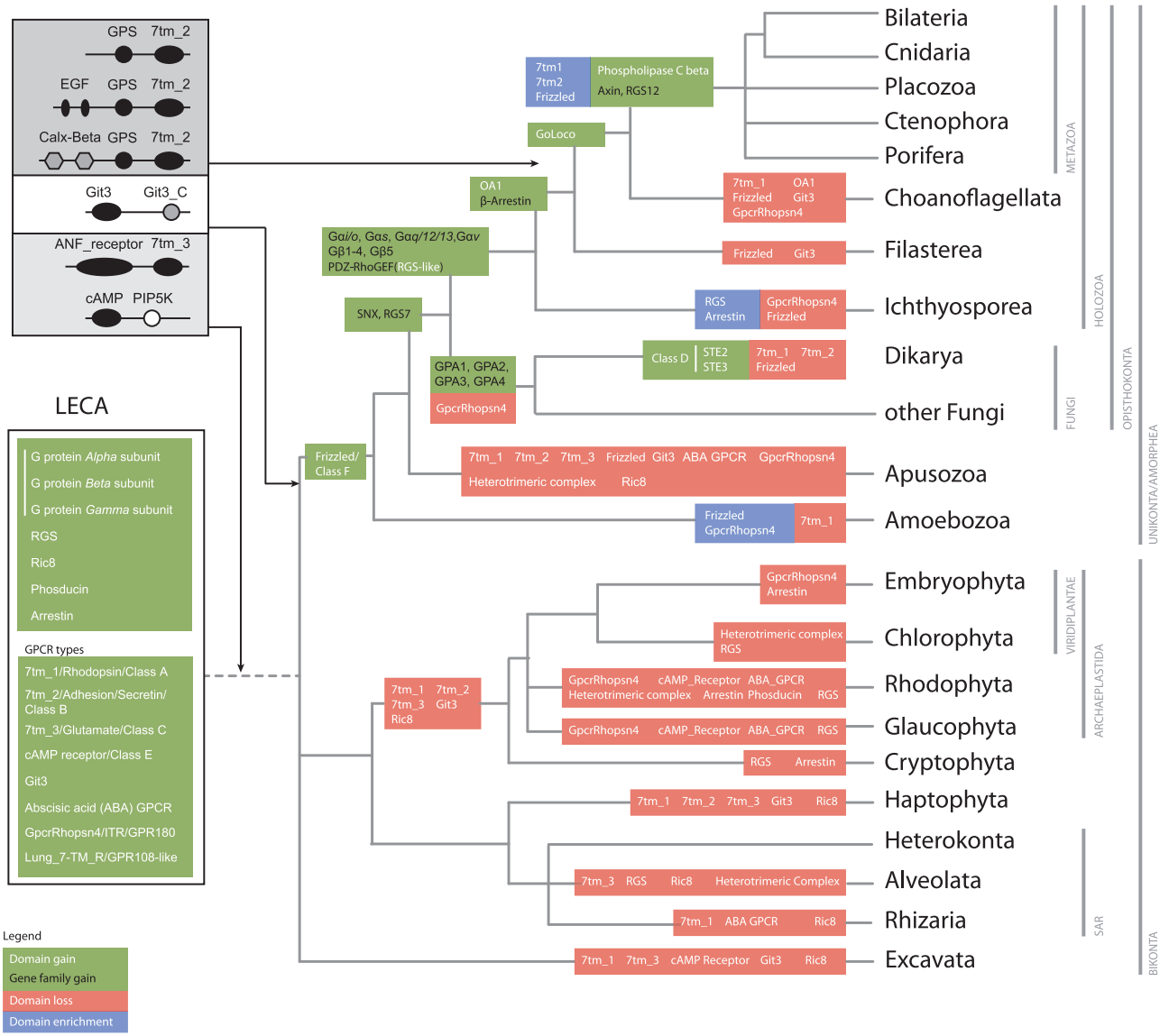

Fig. 6.—Cladogram representing the major patterns of evolution of GPCR signaling components in a eukaryotic phylogeny. Colored boxes with white text indicate specific components defined by a domain, whereas colored boxes with black text refer to specific gene family acquisitions. Green and red boxes depict gain and loss of domains, respectively, and blue boxes depict significant enrichments of the component shown, according to a Wilcoxon rank-sum test, with *P* value threshold of <0.01. Additionally, we show in the upper part a selected set of conserved GPCR architectures placed where they must have appeared according to Dollo Parsimony.

Modularity is not only observed in how the various elements of the GPCR signaling pathway evolve but also at the level of protein domain architectures. Overall, our results on domain architectures clearly show that domain shuffling is a major mechanism of signaling system evolution. Indeed, pervasive convergent evolution of domain arrangements is a major feature of both GPCR receptors and RGS proteins (Nordström et al. 2009; Anantharaman et al. 2011; Krishnan et al. 2012). However, because not all GPCR families are equally susceptible to acquiring new domains, functional constraints might also exist that prevent this evolutionary mechanism of innovation.

A recent functional study in a subset of different G *alpha* subunits of various eukaryotes suggests that canonical GPCR signaling is restricted to amorpheans (Bradford et al. 2013). However, our results suggest some inconsistencies under that perspective. For example, the presence of Ric8 in heterokonts (including *E. siliculosus* tested in the study) may imply that in that lineage there is GEF activation of G protein *alpha* subunits and not only "self-activation." Also, the presence of both

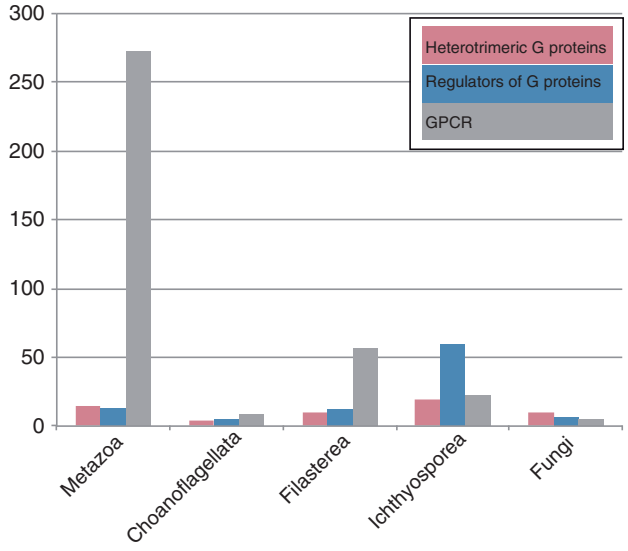

Fig. 7.—Graphic depicts the median number of GPCR signaling components in opisthokont lineages. Total numbers of G protein *alpha*, *beta,* and *gamma* subunits are comprised in the heterotrimeric G proteins category; RGS and Go-Loco-motif containing proteins are comprised in regulators of G proteins category; and GPCR types presented in figure 3 are comprised in GPCR category. Median values were obtained using all taxa of a given clade as shown in the supplementary table S1, Supplementary Material online.

7TM-RGS and canonical GPCRs in opisthokonts (filastereans, ichthyosporeans, and early branching fungi) blurs the distinction between GAP and GEF receptor-based G protein signaling, as they coexist in some lineages. Furthermore, the monophyly of lineage specific G *alpha* protein clades implies that each of those lineages had diversified their own repertoires. Thus, there is not a conserved "self-activation" subfamily.

Instead, "self-activation" could have evolved as a convergent character of G *alpha* subunits. Because only the activity of a single paralog of G *alpha* subunit has been tested for most lineages, it would be interesting to test more paralogs to clarify whether self-activation is the only mechanism in bikonts (Bradford et al. 2013). Finally, the presence of many GPCR types with functionally known amorphean domain architectures and rich heterotrimeric protein complements in bikonts, such as in *Phytophthora infestans* and *B. natans,* suggest that they may have had a canonical GPCR signaling. Those species should be ideal to test different G *alpha* subunits experimentally. Overall, our results suggest that GPCR-G protein canonical signaling is older than previously hypothesized, most likely already being functional at the LECA.

Irrespectively, if the canonical GPCR signaling evolved in the root of amorpheans or before, regarding the origin of metazoans, our results show a bimodal pattern of evolution of the elements of the GPCR signaling system. Cytoplasmic transduction elements, such as G proteins, Ric-8, GoLoco motif, Arrestins, and RGS families, are largely conserved between unicellular holozoans and metazoans, both in terms of gene families and protein domain architectures (fig. 7). In contrast, receptors underwent a dramatic expansion in metazoans compared with their closest unicellular relatives, and a similar pattern has also been observed for tyrosine kinases, Hippo signaling, and Notch signaling elements (Gazave et al. 2009; Sebé-Pedrós et al. 2012; Suga et al. 2012). The signaling output of GPCRs depends on the combinatory of heterotrimeric G proteins and their regulators, and, remarkably, the combination that originated in ancient holozoans was already sufficient for transducing the huge amount of GPCR signaling inputs present in metazoans. The expansion of receptors is probably driven by metazoans' multicellularity, which co-opted the GPCR signaling system for many new functions, such as cell–cell communication, developmental control, and most importantly in the case of GPCR, complex environmental sensing, from light sensing to odor and taste. We suggest that the shift from a universal eukaryotic signaling system to a dramatic expansion and refinement in metazoans played a key role in the acquisition of complex multicellularity.

Supplementary Material

Supplementary figures S1–S13, file 1, and table S1 are available at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

This work was supported by a contract from the Institució Catalana de Recerca i Estudis Avançats, a European Research Council Starting Grant (ERC-2007-StG-206883), a grant (BFU2011-23434) from Ministerio de Economía y Competitividad (MINECO) to I.R.-T, and a pregraduate Formación del Personal Investigador and a pregraduate Formación Profesorado Universitario grant from MINECO to A.d.M. and A.S.-P.

Literature Cited

Alvarez CE. 2008. On the origins of arrestin and rhodopsin. BMC Evol Biol. 8:222.

Anantharaman V, Abhiman S, de Souza RF, Aravind L. 2011. Comparative genomics uncovers novel structural and functional features of the heterotrimeric GTPase signaling system. Gene 475:63–78.

Anantharaman V, Iyer LM, Aravind L. 2007. Comparative genomics of protists: new insights into the evolution of eukaryotic signal transduction and gene regulation. Annu Rev Microbiol. 61:453–475.

Anderson GR, Posokhova E, Martemyanov KA. 2009. The R7 RGS protein family: multi-subunit regulators of neuronal G protein signaling. Cell Biochem Biophys. 54:33–46.

Blaauw M, et al. 2003. Phosducin-like proteins in *Dictyostelium discoideum*: implications for the phosducin family of proteins. EMBO J. 22:5047–5057.
Bockaert J, Pin J. 1999. Molecular tinkering of G protein-coupled receptors: an evolutionary success. EMBO J. 18:1723–1729.
Bommert P, Je BI, Goldshmidt A, Jackson D. 2013. The maize Gα gene COMPACT PLANT2 functions in CLAVATA signalling to control shoot meristem size. Nature 502:555–558.
Bradford W, et al. 2013. Eukaryotic G protein signaling evolved to require G protein-coupled receptors for activation. Sci Signal. 6:ra37.
Brown MW, Kolisko M, Silberman JD, Roger AJ. 2012. Aggregative multicellularity evolved independently in the eukaryotic supergroup Rhizaria. Curr Biol. 22:1–5.
Burki F, Okamoto N, Pombert J-F, Keeling PJ. 2012. The evolutionary history of haptophytes and cryptophytes: phylogenomic evidence for separate origins. Proc Biol Sci. 279:2246–2254.
Chan P, Thomas CJ, Sprang SR, Tall GG. 2013. Molecular chaperoning function of Ric-8 is to fold nascent heterotrimeric G protein. Proc Natl Acad Sci U S A. 110:3794–3799.
Crespi BJ. 2001. The evolution of social behavior in microorganisms. Trends Ecol Evol. 16:178–183.
Derelle R, Lang FB. 2012. Rooting the eukaryotic tree with mitochondrial and bacterial proteins. Mol Biol Evol. 29:1277–1289.
Derelle R, Lopez P, Le Guyader H, Manuel M. 2007. Homeodomain proteins belong to the ancestral molecular toolkit of eukaryotes. Evol Dev. 9:212–219.
DeWire SM, Ahn S, Lefkowitz RJ, Shenoy SK. 2007. β-Arrestins and cell signaling. Annu Rev Physiol. 69:483–510.
Fredriksson R, Lagerström MC, Lundin L-G, Schiöth HB. 2003. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. Mol Pharmacol. 63:1256–1272.
Fredriksson R, Schiöth HB. 2005. The repertoire of G-protein–coupled receptors in fully sequenced genomes. Mol Pharmacol. 67:1414.
Fritz-Laylin LK, et al. 2010. The genome of *Naegleria gruberi* illuminates early eukaryotic versatility. Cell 140:631–642.
Gazave E, et al. 2009. Origin and evolution of the Notch signalling pathway: an overview from eukaryotic genomes. BMC Evol Biol. 9:249.
Gerhart J. 1999. 1998 Warkany lecture: signaling pathways in development. Teratology 60:226–239.
Grau-Bové X, Sebé-Pedrós A, Ruiz-Trillo I. 2013. A genomic survey of HECT ubiquitin ligases in eukaryotes reveals independent expansions of the HECT system in several lineages. Genome Biol Evol. 5: 833–847.
Gurevich VV, Gurevich EV. 2006. The structural basis of arrestin-mediated regulation of G-protein-coupled receptors. Pharmacol Ther. 110: 465–502.
Hinrichs MV, Torrejón M, Montecino M, Olate J. 2012. Ric-8: different cellular roles for a heterotrimeric G-protein GEF. J Cell Biochem. 113: 2797–2805.
Ingham PW, Nakano Y, Seger C. 2011. Mechanisms and functions of Hedgehog signalling across the metazoa. Nat Rev Genet. 12: 393–406.
Kataria R, et al. 2013. *Dictyostelium* Ric8 is a nonreceptor guanine exchange factor for heterotrimeric G proteins and is important for development and chemotaxis. Proc Natl Acad Sci U S A. 110: 6424–6429.
Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30:772–780.
King N. 2004. The unicellular ancestry of animal development. Dev Cell. 7: 313–325.
King N, et al. 2008. The genome of the choanoflagellate *Monosiga brevicollis* and the origin of metazoans. Nature 451:783–788.
Krishnan A, Almén MS, Fredriksson R, Schiöth HB. 2012. The origin of GPCRs: identification of mammalian like rhodopsin, adhesion, glutamate and frizzled GPCRs in fungi. PLoS One 7:e29817.
Krogh A, Larsson B, von Heijne G, Sonnhammer ELL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol. 305:567–580.
Lagerström MC, Schiöth HB. 2008. Structural diversity of G protein-coupled receptors and significance for drug discovery. Nat Rev Drug Discov. 7:339–357.
Lartillot N, Lepage T, Blanquart S. 2009. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. Bioinformatics 25:2286–2288.
Li L, Wright SJ, Krystofova S, Park G, Borkovich KA. 2007. Heterotrimeric G protein signaling in filamentous fungi. Annu Rev Microbiol. 61: 423–452.
Liggett SB. 2011. Phosphorylation barcoding as a mechanism of directing GPCR signaling. Sci Signal. 4:pe36.
Maddison WP, Maddison DR. 2011. Mesquite: a modular system for evolutionary analysis. Version 2.75. [cited 2013 June]. Available from: http://mesquiteproject.org.
Maeda Y, Ide T, Koike M, Uchiyama Y, Kinoshita T. 2008. GPHR is a novel anion channel critical for acidification and functions of the Golgi apparatus. Nat Cell Biol. 10:1135–1145.
Manning G, Young SL, Miller WT, Zhai Y. 2008. The protist, *Monosiga brevicollis*, has a tyrosine kinase signaling network more elaborate and diverse than found in any known metazoan. Proc Natl Acad Sci U S A. 105:9674–9679.
Milligan G, Kostenis E. 2006. Heterotrimeric G-proteins: a short history. Br J Pharmacol. 147(Suppl): S46–S55.
Mushegian A, Gurevich VV, Gurevich EV. 2012. The origin and evolution of G protein-coupled receptor kinases. PLoS One 7:e33806.
Niehrs C. 2012. The complex world of WNT receptor signalling. Nat Rev Mol Cell Biol. 13:767–779.
Nordström KJ, Almén MS, Edstam MM, Fredriksson R, Schiöth HB. 2011. Independent HHsearch, Needleman-Wunsch-based and motif analyses reveals the overall hierarchy for most of the G protein-coupled receptor families. Mol Biol Evol. 28:2471–2480.
Nordström KJ, Lagerström MC, Wallér LM, Fredriksson R, Schiöth HB. 2009. The Secretin GPCRs descended from the family of Adhesion GPCRs. Mol Biol Evol. 26:71–84.
Oka Y, Saraiva LR, Kwan YY, Korsching SI. 2009. The fifth class of Gα proteins. Proc Natl Acad Sci U S A. 106:1484–1489.
Oldham WM, Hamm HE. 2008. Heterotrimeric G protein activation by G-protein-coupled receptors. Nat Rev Mol Cell Biol. 9:60–71.
Pierce KL, Premont RT, Lefkowitz RJ. 2002. Seven-transmembrane receptors. Nat Rev Mol Cell Biol. 3:639–650.
Pires-daSilva A, Sommer RJ. 2003. The evolution of signalling pathways in animal development. Nat Rev Genet. 4:39–49.
Plakidou-Dymock S, Dymock D, Hooley R. 1998. A higher plant seven-transmembrane receptor that influences sensitivity to cytokinins. Curr Biol. 8:315–324.
Punta M, et al. 2012. The Pfam protein families database. Nucleic Acids Res. 40:D290–D301.
Reiter E, Lefkowitz RJ. 2006. GRKs and beta-arrestins: roles in receptor silencing, trafficking and signaling. Trends Endocrinol Metab. 17: 159–165.
Rensing SA, et al. 2008. The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. Science 319:64–69.
Rokas A. 2008. The origins of multicellularity and the early history of the genetic toolkit for animal development. Annu Rev Genet. 42:235–251.
Rosenbaum DM, Rasmussen SGF, Kobilka BK. 2009. The structure and function of G-protein-coupled receptors. Nature 459:356–363.

Schaller GE, Shiu S-H, Armitage JP. 2011. Two-component systems and their co-option for eukaryotic signal transduction. Curr Biol. 21: R320–R330.
Sebé-Pedrós A, Zheng Y, Ruiz-Trillo I, Pan D. 2012. Premetazoan origin of the hippo signaling pathway. Cell Rep. 1:1–8.
Shenoy S, Lefkowitz R. 2011. β-arrestin-mediated receptor trafficking and signal transduction. Trends Pharmacol Sci. 32:521–533.
Siderovski DP, Willard FS. 2005. The GAPs, GEFs, and GDIs of heterotrimeric G-protein alpha subunits. Int J Biol Sci. 1:51–66.
Sondek J, Siderovski DP. 2001. Ggamma-like (GGL) domains: new frontiers in G-protein signaling and beta-propeller scaffolding. Biochem Pharmacol. 61:1329–1337.
Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688–2690.
Suga H, et al. 2012. Genomic survey of premetazoans shows deep conservation of cytoplasmic tyrosine kinases and multiple radiations of receptor tyrosine kinases. Sci Signal. 5:ra35.
Torruella G, et al. 2012. Phylogenetic relationships within the Opisthokonta based on phylogenomic analyses of conserved single-copy protein domains. Mol Biol Evol. 29:531–544.
Tuteja N. 2009. Signaling through G protein coupled receptors. Plant Signal Behav. 4:942–947.
Urano D, et al. 2012. G protein activation without a GEF in the plant kingdom. PLoS Genet. 8:e1002756.
Wickstead B, Gull K, Richards TA. 2010. Patterns of kinesin evolution reveal a complex ancestral eukaryote with a multifunctional cytoskeleton. BMC Evol Biol. 10:110.
Wilkie TM, Kinch L. 2005. New roles for Galpha and RGS proteins: communication continues despite pulling sisters apart. Curr Biol. 15: R843–R854.
Willardson B, Howlett A. 2007. Function of phosducin-like proteins in G protein signaling and chaperone-assisted protein folding. Cell Signal. 19:2417–2427.

Associate editor: Yves Van De Peer