



**HAL**  
open science

## An evolutionary perspective on Arf family GTPases

Catherine L Jackson, Julie Ménétreay, Mandeep Sivia, Joel B Dacks, Marek Eliáš

► **To cite this version:**

Catherine L Jackson, Julie Ménétreay, Mandeep Sivia, Joel B Dacks, Marek Eliáš. An evolutionary perspective on Arf family GTPases. *Current Opinion in Cell Biology*, 2023, 85, pp.102268. 10.1016/j.ceb.2023.102268 . hal-04304565

**HAL Id: hal-04304565**

**<https://hal.science/hal-04304565>**

Submitted on 24 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## An evolutionary perspective on Arf Family GTPases

Catherine L. Jackson<sup>a</sup>, Julie Ménétreay<sup>b</sup>, Mandeep Sivia<sup>c</sup>, Joel B. Dacks<sup>c</sup>, Marek Eliáš<sup>d</sup>

<sup>a</sup>Université Paris Cité, CNRS, Institut Jacques Monod, F-75013 Paris, France.

E-mail : [cathy.jackson@ijm.fr](mailto:cathy.jackson@ijm.fr)

<sup>b</sup>Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC),

91198, Gif-sur-Yvette, France. E-mail: [julie.menetrey@i2bc.paris-saclay.fr](mailto:julie.menetrey@i2bc.paris-saclay.fr)

<sup>c</sup>Division of Infectious Disease, Department of Medicine, University of Alberta, Edmonton,

AB, Canada T6G 2H7. E-mail: [dacks@ualberta.ca](mailto:dacks@ualberta.ca)

<sup>d</sup>Department of Biology and Ecology, Faculty of Science, University of Ostrava, Ostrava,

Czech Republic. E-mail: [Marek.Elias@osu.cz](mailto:Marek.Elias@osu.cz)

Corresponding author : CL Jackson, E-mail: [cathy.jackson@ijm.fr](mailto:cathy.jackson@ijm.fr); Tel: +33 1 57 27 80 04

## Abstract

The Arf family GTPases are regulators of eukaryotic cellular organization, with functions in processes such as vesicle budding, membrane tethering, microtubule dynamics, actin cytoskeleton regulation and lipid dynamics. We describe the evolution of this protein family and its well-studied regulators. The last eukaryotic common ancestor had fifteen members, and the current complement of Arf GTPases has been sculpted by gene loss and gene duplications since that point. Some Arf family GTPases (such as those which recruit vesicle coats in the secretory pathway) are present in virtually all eukaryotes, whereas others (such as those functioning in cilia and flagella) have a more limited distribution. A challenge for the future is understanding the full spectrum of Arf family functions throughout eukaryotes.

## Introduction

Arf family proteins are key regulators of cellular organization in eukaryotic cells, functioning in the secretory and endocytic pathways, at mitochondria, in cilia and flagella, in actin cytoskeleton regulation, and in microtubule dynamics. This family of GTPases includes Arf, Arl, and Sar proteins, as well as Arfrp1 and ARD1/TRIM23 (Box 1). Arf family GTPases are found throughout eukaryotes, and many are highly conserved across a broad range of animals, plants, fungi and protists [1]. Here we review the latest developments in the evolution and functions of Arf family GTPases in eukaryotes.

---

### Box 1

The different names of Arf family proteins (Arf, Arl, Sar, Arfrp1, ARD1, etc.) does not indicate fundamental differences in their cellular functions. Arf1 (for ADP ribosylation factor 1), the first Arf protein identified, was discovered as a factor stimulating ADP ribosylation of the  $\alpha$  subunit of the heterotrimeric G protein Gs ( $G_s\alpha$ ) by cholera toxin, stimulating the production of the second messenger cAMP [2]. Later it was found that a major cellular function of Arf1 is in regulation of the secretory pathway [3], as well as other cellular processes [4,5], and its ADP ribosylation co-factor activity does not appear to contribute to these normal cellular functions. As more proteins related to Arf family proteins were

identified, the lack of a systematic approach led to problems in assigning names. In 2006, a consensus nomenclature was agreed upon [6]. We will refer to all of these GTPases as “Arf family proteins”. The SR $\beta$  proteins are phylogenetically related to the Arf family proteins [7], but are not included here in our definition of Arf family proteins.

---

As for small G proteins in general, Arf family proteins act as molecular switches, inactive in their GDP-bound form, and active when bound to GTP (Figure 1a). Arf1, and many other members of the family, are soluble in their GDP-bound form but are tightly membrane-bound in their active GTP-bound form [4,5]. The change in membrane association does not rely on other proteins, but is due to a second conformational change within the protein itself upon GTP binding (Figure 1a). Unique to Arf family proteins is the presence of an N-terminal amphipathic helix (AH), often myristoylated, that mediates membrane binding (Figure 1a). Although the majority of Arf family proteins have this mechanism of membrane association, there are some exceptions [7–9].

Effectors, proteins which interact with the active GTP-bound form, carry out the cellular functions of Arfs. The majority of Arf family proteins, whose GTP-bound form is membrane-bound, recruit effectors to the membrane surface, making the membrane an important component of Arf function (Figure 1a). The major classes of effectors are vesicle coats, membrane tethering proteins, lipid modifying and transfer proteins, cytoskeleton components (actin and microtubule), and microtubule motors and their adaptors.

Arf family proteins in eukaryotes require regulators, guanine nucleotide exchange factors (GEFs) and GTPase activating proteins (GAPs), to accelerate release of GDP and hydrolysis of GTP, respectively (Figure 1a). Sec12 was identified as a GEF for Sar1 on the ER membrane [10], and Sec7 domain proteins were identified as GEFs for Arf1 [4]. Whereas only one family of Sec12 GEFs has been found, multiple subfamilies of Sec7 domain proteins were identified, with activities on Arf1-6 and Arl1 [4,11,12]. It is only recently that new families of Arf GEFs have been identified. Arl13b and BART are both required for GEF activity on Arl3 in cilia [9,13], and AMPK has been found in both yeast and mammalian cells to activate Arf6 (called Arf3 in yeast) [14,15]. In both organisms, AMPK activates Arf6/ $\gamma$ Arf3 in response to glucose depletion [14,15]. Sec23, a component of the COPII coat, acts as a GAP for Sar1, and the Zn-finger domain containing proteins are GAPs for Arf1-6 and Arl1,

comprising 11 subfamilies [5,16,17]. Three additional families of Arf GAPs were subsequently identified. The ELMOD proteins have a relatively broad range of activities on Arf family proteins, including Arf1-6, Arl1, Arl2 and possibly ciliary Arl3 and Arl13b [18,19]. RP2 is a GAP that acts on Arl2 and Arl3, and has been characterized in animal cells and yeast [20]. Recently, an Arf GAP complex containing longin domains, the C9orf72 complex, was described that functions on Arf1, Arf5 and Arf6 [21,22].

### **Sculpting ancestral complexity by gene loss is a major pattern of Arf family evolution**

A recent comprehensive phylogenetic analysis of Arf family proteins has provided new insights into these important regulators of eukaryotic cellular organization (Figure 1b) [7]. The most notable result was the reconstruction of the complement of Arf family members in the last eukaryotic common ancestor (LECA), which turned out to be larger than previously suspected (Figure 2). In addition to paralogs typified by metazoan, usually well-studied proteins (Arf1 and 6, Arl1, 2, 3, 5, 6, 8, 13, and 16, Arfrp1, and Sar1), the LECA seems to have encoded three additional Arf family members that are absent from virtually all metazoans and which were denoted Arl17, Arl18 and SarB. Thus, these three proteins qualify as so-called jotrlogs, proteins that are ancient in eukaryotes yet (secondarily) missing from opisthokonts [23]. In total, the LECA was most likely endowed with 15 different Arf family genes, a complement larger than that present in many extant eukaryotes.

Indeed, gene loss has been pervasive in the evolutionary history of the Arf family, with some of the ancestral members lost many times independently across the eukaryote phylogeny. As an example, let us take you on a journey through time, following the fate of the Arf family along the organismal line of descent from the LECA to one of the most extensively studied biological model systems, the yeast *Saccharomyces cerevisiae*. We base the narrative on the previous published data [7] with updates that fill in critical gaps in the taxon sampling (Figure 2).

The first change to the LECA's Arf family gene set was the loss of Arl18 deep in the eukaryote tree (prior to the divergence of Amoebozoa). The last common ancestor of the eukaryote supergroup Opisthokonta thus inherited a nearly complete ancestral set of Arf family members. From this point on, the yeast lineage experienced a stepwise reduction of the set, with Arl16 loss having occurred immediately after the divergence of the Holozoa lineage (containing metazoans and their closest protist relatives), followed by loss of the

SarB gene after the divergence of Nucleariida (a group of amoeboid protists). As the yeast ancestors were still flagellated at this point, they had retained the flagellum-associated GTPases Arl3, Arl6, and Arl13, but these vanished with the loss of the capability to build flagella later in the evolution of fungi. Arl5 was then lost in the stem branch of the major fungal phylum Ascomycota, and Arl17 in the stem branch of its subphylum Saccharomycotina. The final loss event, removing Arl8, occurred late in the *S. cerevisiae* evolutionary lineage. As a result, budding yeast has retained only six out of the 15 ancestral Arf family members.

### **Gene duplication also shapes the Arf family in a taxon-specific manner**

The final event in the genesis of the *S. cerevisiae* Arf family gene set was duplication of the ancestral Arf1 gene yielding a pair of lineage-specific paralogs (Arf1 and Arf2), which was part of the whole genome duplication (WGD) that occurred in the common ancestor of *Saccharomyces* and several related genera [24] (Figure 2). Although not the case for *S. cerevisiae*, taxon-specific expansions of the Arf family by gene duplication have been pervasive across eukaryotes in general. Let us illustrate this general pattern by again following the evolutionary trajectory of the Arf family in a single specific organismal lineage, this time the one leading to us, humans (Figure 2).

We start with the last common ancestor of opisthokonts, when the human lineage, as part of the clade Holozoa, diverged from that leading to yeasts. In contrast to the yeast lineage, the stem branch of Holozoa was instead evolutionarily creative, as emergence of three novel Arf family members can now be mapped to it: Arf4, Arl15 and Arl19. Arf4 and probably Arl19 emerged via duplication of the ancestral Arf1 gene, whereas the origin of Arl15 has been obscured by extreme sequence divergence. At a later point of the holozoan phylogeny but still during its “protist stage”, another paralog, Arl20, emerged again presumably from a duplicated Arf1 gene.

In contrast to numerous novel paralogs in the Rab family of GTPases evolved during the transition to multicellularity in the stem branch of Metazoa [25], the history of the Arf family is marked only with the addition of a single new gene, Arl10, and also by loss, namely of SarB. Two additional novel family members, Arl4 and the multidomain protein TRIM23, evolved in early metazoan evolution, the former from a duplicate of Arl19 and the latter perhaps from a duplicated Arf1 gene by its fusing to a region encoding additional protein

domains [7]. Around the same time, Arl17 was lost from the human ancestors, later followed by loss of Arl20 before the emergence of vertebrates.

The composition of the Arf family set was then affected by two WGD rounds in early vertebrate evolution, one most likely preceding the divergence of cyclostomes and the other specific for jawed vertebrates [26]. These events together with post-WGD losses of some of the duplicates seem to be the major cause for some of the Arf family genes inherited from the invertebrate ancestor being represented by two or three different copies in the genome of humans and other vertebrates (Figure 2). The nomenclature of the WGD-derived copies (also known as ohnologs [27]) reflects their common origin in an inconsistent way, with, eg., the Arl8a/Arl8b pair being not different in this regard than, eg., the Arf4/Arf5 pair. Another pair of ohnologs, Arl11 and Arl14, itself seems to have emerged from a vertebrate-specific parental gene that evolved before the WGD rounds from a duplicated Arl4 gene by its substantial sequence divergence. The final steps towards the Arf family gene complement as represented in the human genome were losses of the Arl19 and Arf2 genes at different nodes of mammalian evolution (Figure 2).

In general, the outcomes of the duplication events in the Arf family history span a broad gradient from augmenting the Arf family by extra copies of some of the genes that are identical or extremely similar at the protein sequence level, to the emergence of “novel” paralogs that have diverged from their parental gene to such an extent as to effectively obscure the specific origin of the gene. We will provide examples below of the functional significance of duplications, which can be more or less asymmetric. The emergence of “novel” paralogs is not exclusive for the human (holozoan) evolutionary lineage, as divergent Arf family members dissimilar to any of the 15 ancestral paralogs are common in diverse eukaryotes [7].

### **Evolution of Arf family regulators**

As compared with the Arf family of GTPases, our understanding of the diversity and evolution of its regulators is less developed. Nonetheless, the work to date brings out several key points of both contrast and parallel. Whereas the Arf family is derived from a common ancestor, the same is not true of either GEFs or GAPs, which have arisen from different non-homologous protein families. Still, the themes of protein family expansion, loss, and novel paralogs are also found in the regulators (Figure 3).

Reaching back to the LECA, we see that of the eight subfamilies of Sec7 domain Arf family GEFs in humans, three (GBF, BIG, and cytohesin) [11], along with the Sar1 GEF Sec12 [28] and AMPK [29], were already present. The Sar1 GAP Sec23 was present in at least one copy [28], seven subfamilies of Zn-finger domain Arf GAPs were present [17], and at least one (possibly two) ELMOD Arf GAPs were present in LECA [19]. ELMOD family proteins have recently been functionally characterized in *Arabidopsis* [30] and localized in *Trypanosoma brucei* (<http://tryptag.org/?id=Tb927.9.2670>) [31].

There are several novel domain architectures of both GEFs and GAPs found in various eukaryotic lineages (Figure 3). ArfGAP\_C2 has not been found in humans or other opisthokonts, making it a jotrlog, but is present in plants, haptophytes, apusomonads, and members of the SAR clade [17]. ArfGAP\_C2 possesses the canonical Zn finger Arf GAP domain, followed by a C2 lipid-binding domain. Novel cytohesin-derived Arf GEFs (ARCC) with multiple N-terminal ankyrin domains exist in amoebozoans, haptophytes, and alveolates [11]. Careful analysis showed that ARCC Arf GEFs must have evolved independently in amoebozoans and alveolates, at least [11].

Loss has also clearly shaped the complements of Arf family regulators. One of the starkest examples of reductive evolution is the regulator complement found in the metamonad lineage leading to the gut parasite *Giardia* (Figure 3). In this case, both the free-living and parasitic organisms sampled encoded only two SEC7 Arf GEF subfamilies and three Zn finger Arf GAP domain subfamilies [32]. And yet with loss also came paralog expansion. While the Arf6 subfamily has been lost from this lineage, the Arf1 subfamily has been expanded at multiple time points [32]

Jumping to the ancestor of animals and fungi (opisthokonts), we see another point of contrast between the GTPases and their regulators. While there is little emergent novelty between the LECA and the opisthokont ancestor in the Arf family, several new GEFs and GAPs have appeared on the scene by this time (Figure 3). The EFA6 subfamily (PSD in animals) is a cytohesin derivation and had clearly already evolved at this point. Similarly, though ArfGAP\_C2 has been lost, the Arf GAP ASAP emerges by this time [17]. Finally, by this time, through gene duplication and novel domain fusions, the ELMO proteins arose. However, these proteins are not Arf GAPs because their ELMO domains have lost the catalytic arginine [19].



The ancestor of animals is the border between uni- and multi-cellularity in this lineage and an area of active and exciting research into the under-appreciated diversity of small predatory flagellates that were our ancestral form [33]. Here both the GTPases and their regulators have undergone duplications to produce new members of each family, although the new regulators have not emerged to regulate only the newly evolved Arf family members. Just as several new GTPases are on scene by this time, the ArfGEFs IQSEC7 and BRAG, both cytohesin derived, had appeared by this stage [11], as did the Arf GAP subfamilies GIT and ARAP [17]. Finally, the complement of Arf regulators was shaped extensively by the WGDs around the evolution of vertebrates, with between two and four paralogs of each of the regulators appearing by the time of the vertebrate common ancestor [11,17,19].

### **Functions of Arf family proteins**

Of the 15 Arf family proteins already present in the LECA (Figure 1b), 12 have known functions [7]. These functions include coat recruitment by Sar1, Arf1, Arl6; microtubule dynamics by Arl2, Arf1, Arf6, Arl8; Golgi tethering by Arf1, Arl1, Arl5, and Arfrp1; cilia function by Arl3, Arl6, Arl13, Arl16 (in addition to Arl2, Arf4 and Arf6); regulation of the actin cytoskeleton by Arf1 and Arf6; and lipid dynamics and homeostasis by Arf1, Arf6, Arl5 and Arfrp1. A full discussion of Arf family function is beyond the scope of this review, so we will focus on core functions and key advances made in the past two years in understanding these functions, at the molecular level.

Arf1 and Sar1 were found in every one of the eukaryotic genomes examined in Vargová et al. [7]. Both Sar1 and Arf1 regulate recruitment of vesicle coats for vesicle formation in the early secretory pathway – anterograde ER-Golgi COPII vesicles in the case of Sar1, and retrograde Golgi-ER COPI vesicles in the case of Arf1. Secretion is an essential function in eukaryotic cells, and recruitment of the COPI and COPII coats by Arf1 and Sar1, respectively, has been demonstrated experimentally in all eukaryotes examined, including animals, fungi, plants and protists [34–38]. Arf1 also recruits at least four other coat complexes to *trans*-Golgi and endosomal membranes, AP-1/clathrin, GGA/clathrin, AP-3 and AP-4. Three of these vesicle coats, in addition to COPI and COPII, are all inferred to have been present in LECA [39]. Hence coat recruitment for vesicle formation represents a core function of Arf family proteins, which was likely present in the LECA.

Arf6 functions in endosome-plasma membrane trafficking pathways and not at the Golgi as Arf1 does [4,5]. Arf1 and Arf6 were both present in LECA, indicating an ancient division of functions [7]. The four human Arf1 paralogs (Arf1-5, Arf2 is missing due to loss, see above) have been shown to promote formation of COPI vesicles *in vitro*, although only Arf1, Arf4 and Arf5 are found on COPI vesicles purified from cells [40]. Systematic knockouts of the four paralogs in HeLa cells revealed a high level of redundant function [41]. Single deletions of Arf1 or Arf4 were viable, but the double deletion mutant was inviable, indicating redundant functions essential for growth. This result illustrates the principle that after gene duplication during evolution, a certain degree of functional redundancy of the copies may persist over long periods (Figure 2). However, analysis of individual knockout phenotypes indicated that Arf1 and Arf4 have specialized functions as well: Arf4 in retrieval of resident ER proteins from the *cis*-Golgi and Arf1 in exit pathways from the Golgi mediated by COPI, AP-1 and AP-3 coats [41].

After Arf1 and Sar1, Arl2 is the next most widely distributed Arf family member among eukaryotes, with its loss so far encountered only in the genus *Entamoeba* [7]. The most universal function of Arl2 is as a chaperone for the folding of  $\alpha$ - and  $\beta$ -tubulin. As a result, Arl2 is involved in numerous microtubule-dependent processes. In addition, Arl2 shares functions with the closely related Arl3 in cilia and flagella, as well as regulating mitochondrial function [8,42,43].

Multiple Arf family GTPases regulate movement of organelles along microtubules, including movement of Golgi elements to the cell center by Arf1 [44], and bidirectional movement of endosomes at the midbody by Arf6 [45,46]. Recently, a series of papers have demonstrated the mechanisms by which Arl8 regulates bidirectional movement of lysosomes along microtubules. Arl8 interacts with SKIP on lysosomal membranes to couple them to kinesin-1, thus promoting movement of lysosomes towards the periphery of cells [47]. Arl8 can also interact with RUFY3 and RUFY4, which interact with dynein-dynactin to mediate movement of lysosomes towards the minus end of microtubules at the cell center [48,49]. In this way, Arl8 can regulate bidirectional movement along microtubules, through binding to different effectors. Together, these results show that Arf family proteins have a prominent role in microtubule dynamics.

Arl1, Arl5 and Arfrp1 are found in all of the main eukaryotic lineages, but are lost more frequently (mainly in terminal branches) compared to Sar1, Arf1 and Arl2 [7]. Arl1

shares a number of functions with Arf1, and they also share a GEF and GAP [12,16], but Arl1 and Arl5 are more closely related phylogenetically (Figure 1b) [7]. Arf1, Arl1 and Arl5 share the function of recruiting tethers to Golgi membranes: coiled-coil tethers by Arf1 and Arl1, and multisubunit tethers by Arl5 [5,50,51]. Interestingly, Arfrp1 acts upstream of both Arl1 and Arl5 to coordinate recruitment of coiled-coil golgin tethers to the Golgi by Arl1 and multisubunit tethers by Arl5 [51], and has been shown to promote secretion of insulin [52].

Four families of Arf family proteins present in LECA, Arl3, Arl6, Arl13 and Arl16, have primary functions in cilia and flagella. It was only with the extensive phylogenetic analyses of Arf family proteins that the ciliary function of Arl16 was uncovered [53]. Arl16 localizes to cilia in mammalian cells, and its deletion in mouse or human cells causes defects in cilia formation and morphology, at least in part due to a defect in trafficking of certain ciliary proteins from the Golgi to the cilium [53]. Arl3 and Arl13b have well-established functions in mediating localization of lipidated and transmembrane proteins to the cilium [13]. BBSomes are structures that are transported along the axonemal microtubules to the tip of the cilium, where they are loaded with cargo (for example, signalling molecules to be removed from the cilium) and then transported back to base of the cilium. Arl6 functions to recruit the BBSome to membranes, after which it enters the cilium, and then connects to the IFT machinery [5]. Two recent studies in *Chlamydomonas reinhardtii* report new roles for Arl3 and Arl13b in promoting BBSome-mediated export from the cilium [54,55]. Arl3-GTP binds to retrograde BBSomes shed from the IFT train, which specifically contain phospholipase D (PLD) [56], an effector of Arl3 [54,55]. In addition to the cilia/flagella-specific Arf family proteins, Arl2 and Arf4 also function in cilia [4,5,8]. Recently, Arf6 has also been shown to function with its GEF EFA6A and Arl13b in early steps of ciliogenesis [57].

Arf family proteins regulate the actin cytoskeleton [4,5]. The small GTPase Rac is particularly important in actin cytoskeleton regulation, and Arf6 recruits effectors that contain Rac regulatory domains [4,5]. In addition, both Arf1 and Arf6 cooperate with Rac to activate the WAVE regulatory complex (WRC), a crucial actin nucleation promotion factor. A recent study shows that Arf1 and Arf6 directly interact with WRC to promote actin nucleation [58]. The WAVE complex was present in the LECA [59], raising the possibility that Arf family regulation of actin nucleation is ancient.

A major class of Arf family effectors is comprised of lipid modifying and transfer proteins [4]. Arf6 regulates plasma membrane composition through recruitment of

phospholipase D1 (PLD1) and PI4P-5 kinase [4,60]. PLD1 is also an effector for Arf1, and PI4 kinases are recruited to the Golgi by Arf1 [60–62]. In a global Arf effector screen [63], Li et al. demonstrated that Arl11 and Arl14 have PLD1 as an effector. Both Arf1 and Arf6 interact with and stimulate PLD1 activity *in vitro*, but Li et al. found that PLD1 is much more strongly associated with Arl11 and Arl14 than Arf6 in co-immunoprecipitations from cells. Moreover, Arl11 and 14 colocalized with PLD1 in cells, and stimulated its activity *in vitro* [63]. Arl11 and Arl14 are closely related, as described above [7]. Similarly, although Arf1-GTP has been previously reported to interact with PI4kinase  $\beta$ , Li et al. found that Arl5a and Arl5b interact to a higher level in co-immunoprecipitations with PI4kinase  $\beta$  than Arf1 does [63]. An independent study identifying interacting partners of Arl5b did not identify PI4kinase  $\beta$ , but did find Acyl-CoA-binding domain-containing protein 3 (ACBD3), which has been shown to recruit PI4kinase  $\beta$  to the *trans*-Golgi [64].

### **Structural insights into Arf family function**

Although 3D structural analyses were essential to elucidate the mechanisms of the GDP/GTP cycle of Arf family proteins, as well as to visualize interactions with their protein partners, the crystallographic approach is limited, especially because of the absence of membranes, which are essential for the functioning of many Arf family proteins. Cryogenic electron microscopy (cryo-EM) and tomography (cryo-ET) approaches have been used to address this limitation. These approaches were used to resolve the structure of the Sar1-COPII coat complex reconstituted on membranes, permitting visualization of the Sar1 AHs embedded in the membrane and determining how the inner COPII coat deforms the membrane to form a vesicle [65]. Recently, the cryo-ET structure of Arf1 bound to AP-1 complexes on liposomes, in the absence of clathrin, was visualized [66]. AP-1 complexes form the inner layer of clathrin coats, but also play roles in clathrin-independent membrane trafficking. These data reveal that myristoylated Arf1 and AP-1 are able to form a tubular coat on membranes independently of clathrin, through an Arf1 dimer interface [66].

The seven-subunit COPI complex associated with Arf1 has been visualized on membranes, revealing that Arf1 occupies two distinct sites in very different molecular environments in the COPI coat [67]. The basic unit of the assembly is a triad of the multimeric COPI complex containing six Arf1 molecules. The complete binding site for

ArfGAP2, near the central Arf1 molecule, was predicted to form only when the coat is assembled, in a “niche” between two leaves within the triad [67], a hypothesis supported by the fitting of the recent crystal structure of ArfGAP2 bound to Arf1 [68] into the cryo-ET density. In a different, but equally exciting direction, cryo-ET can be used to visualize complexes in their native environment within cells. In the case of COPI, *Chlamydomonas* provided the ideal system to visualize coated vesicles, and allowed the structure of the COPI coat within cells to be resolved to 20 Å resolution [69]. An important conclusion from these studies is that the structure of the COPI coat is remarkably similar in distantly related organisms, including mammals, fungi and *Chlamydomonas*.

Cryo-EM and cryo-ET have also been used to determine the 3D structures of Arf family proteins bound to very large, multidomain complexes [22] or multisubunit proteins [70], which often have disordered regions or interdomain flexibility that make crystallization difficult. The three-subunit Arf GAP complex (C9orf72:SMCR8:WDR41) described above belongs to a family of dimeric longin-DENN domain complexes [21,22]. The cryo-EM structure of Arf1 bound to the C9orf72:SMCR8:WDR41 complex showed that the longin domains of C9orf72 and SMCR8 form the binding site for Arf1 and act together to stimulate Arf1 GTP hydrolysis by an arginine finger mechanism [22]. This complex is the first multi-subunit GAP to be identified for Arf family proteins.

### **Conclusions and future directions**

In this review, we have described some key aspects of the evolution of the Arf family GTPases and their regulators in eukaryotes, and highlighted the latest advances in understanding their major functions. The last eukaryotic common ancestor (LECA) contained an unexpectedly large number of Arf family proteins, and thus many organisms have undergone significant loss of this ancient set. Even the human Arf family gene set records a history of gene loss, thus lacking three ancestral paralogs as well as two novel Arf family members (Arl19 and Arl20) that emerged more recently. Being absent also from other well-studied organisms, their function remains completely unknown. Humans and other vertebrates have nevertheless undergone duplication events through their evolutionary history, leading to a doubling of the current repertoire of Arf family proteins in extant organisms compared to LECA.

The combination of improvements in genome sequencing, and organism discovery, continues to provide fertile ground for improved evolutionary investigations of protein families. Phylogenetic analyses described here for the Arf family and the best-studied Arf family regulators can now be applied to newly identified GEFs and GAPs, as well as effectors, an exciting area for future exploration. Advances in imaging technologies have made it possible to push past previous limitations in structural biology, and coupled with cell biological and biochemical studies, is providing new insights into the mechanisms of Arf family protein function. Cryo-electron tomography can be used to obtain high resolution structures of large complexes within cells, in their native environments, an approach that has already been used to obtain structures of the Arf1-COPI coat. These approaches promise to revolutionize our understanding of cellular processes at the molecular level, in both well-studied organisms and in the diversity of less well studied eukaryotes.

#### **Author contributions**

**CLJ, JM, JBD, ME:** Conceptualization, visualisation and writing: **MS:** conceptualization and visualization

**Declarations of interest:** none

#### **Acknowledgements**

CLJ and JM were supported by grant ANR-20-CE13-0007 from the ANR, France, and the CNRS, France. JBD was supported by NSERC Canada Discovery Grants RES0043758 and RES0046091. The evolutionary studies of GTPases in the lab of ME were supported by the Czech Ministry of Education ERD Funds projects OPVVV 16\_013/0001775.

#### **Data availability**

Newly identified and analysed Arf family genes that were used to prepare Figure 2 are available at Figshare XXXXXX (URL to be added in revision).

## References

1. Li Y, Kelly WG, Logsdon Jr. JM, Schurko AM, Harfe BD, Hill-Harfe KL, Kahn RA: **Functional genomic analysis of the ADP-ribosylation factor family of GTPases: phylogeny among diverse eukaryotes and function in *C. elegans*.** *The FASEB Journal* 2004, **18**:1834–1850.
2. Kahn RA, Gilman AG: **Purification of a protein cofactor required for ADP-ribosylation of the stimulatory regulatory component of adenylate cyclase by cholera toxin.** *Journal of Biological Chemistry* 1984, **259**:6228–6234.
3. Stearns T, Willingham MC, Botstein D, Kahn RA: **ADP-ribosylation factor is functionally and physically associated with the Golgi complex.** *Proceedings of the National Academy of Sciences* 1990, **87**:1238–1242.
4. Donaldson JG, Jackson CL: **ARF family G proteins and their regulators: roles in membrane transport, development and disease.** *Nat Rev Mol Cell Biol* 2011, **12**:362–375.
5. Sztul E, Chen P-W, Casanova JE, Cherfils J, Dacks JB, Lambright DG, Lee F-JS, Randazzo PA, Santy LC, Schürmann A, et al.: **ARF GTPases and their GEFs and GAPs: concepts and challenges.** *MBoC* 2019, **30**:1249–1271.
6. Kahn RA, Cherfils J, Elias M, Lovering RC, Munro S, Schurmann A: **Nomenclature for the human Arf family of GTP-binding proteins: ARF, ARL, and SAR proteins.** *Journal of Cell Biology* 2006, **172**:645–650.
- \*\*7. Vargová R, Wideman JG, Derelle R, Klimeš V, Kahn RA, Dacks JB, Eliáš M: **A Eukaryote-Wide Perspective on the Diversity and Evolution of the ARF GTPase Protein Family.** *Genome Biology and Evolution* 2021, **13**:evab157.

Based on a sample of >2,000 manually curated genes from 114 species representing most major eukaryote lineages, the authors defined the Arf family gene complement in the last eukaryotic common ancestor, revealing the existence of previously unknown ancestral Arf family members secondarily missing from the commonly studied models including humans. The study has additionally clarified the genesis of the Arf family in metazoans and pointed to the unexpected diversity of Arf family members in terms of the membrane attachment mechanisms and domain architecture.

8. Fansa EK, Wittinghofer A: **Sorting of lipidated cargo by the Arl2/Arl3 system.** *Small GTPases* 2016, **7**:222–230.
- \*\*9. ElMaghloob Y, Sot B, McIlwraith MJ, Garcia E, Yelland T, Ismail S: **ARL3 activation requires the co-GEF BART and effector-mediated turnover.** *eLife* 2021, **10**:e64624.

The authors show that Arl3 activation requires two proteins, Arl13b and BART. As shown previously, Arl13b can stimulate nucleotide release from Arl3, but here the authors show that this reaction is greatly accelerated by addition of BART. Hence both Arl13b and BART

are required for activation of Arl3, but do not themselves form a complex, demonstrating a novel mechanism of activation for an Arf family protein.

10. Barlowe C, Schekman R: **SEC12 encodes a guanine-nucleotide-exchange factor essential for transport vesicle budding from the ER.** *Nature* 1993, **365**:347–349.
11. Pipaliya SV, Schlacht A, Klinger CM, Kahn RA, Dacks J: **Ancient complement and lineage-specific evolution of the Sec7 ARF GEF proteins in eukaryotes.** *MBoC* 2019, **30**:1846–1863.
12. Yu C-J, Lee F-JS: **Multiple activities of Arl1 GTPase in the trans-Golgi network.** *Journal of Cell Science* 2017, **130**:1691–1699.
13. Gotthardt K, Lokaj M, Koerner C, Falk N, Gießl A, Wittinghofer A: **A G-protein activation cascade from Arl13B to Arl3 and implications for ciliary targeting of lipidated proteins.** *eLife* 2015, **4**:e11859.
14. Hsu J-W, Chen K-J, Lee F-JS: **Snf1/AMP-activated protein kinase activates Arf3p to promote invasive yeast growth via a non-canonical GEF domain.** *Nat Commun* 2015, **6**:7840.
- \*15. Chen K-J, Hsu J-W, Lee F-JS: **AMPK promotes Arf6 activation in a kinase-independent manner upon glucose starvation.** *Journal of Cell Science* 2022, **135**:jcs259609.

This paper reports that AMPK is a GEF for Arf6 in mammalian cells. AMPK was shown to be a GEF for budding yeast Arf6 (yArf3) [18], hence GEF activity on Arf6 proteins is a conserved function for AMPK, at least in opisthokonts. In both cases, AMPK GEF activity is stimulated in response to glucose depletion. AMPK is one of only two non-Sec7 domain GEFs for Arf family proteins that have been reported to date.

- \*16. Feng H-P, Cheng H-Y, Hsiao T-F, Lin T-W, Hsu J-W, Huang L-H, Yu C-J: **ArfGAP1 acts as a GTPase-activating protein for human ADP-ribosylation factor-like 1 protein.** *The FASEB Journal* 2021, **35**:e21337.

This study shows that a mammalian Zn finger Arf GAP, that has been shown to function on Arf1, can also act on Arl1. ArfGAP/Gcs1 was shown previously to act on Arl1 in yeast, so activity of a Zn finger Arf GAP on multiple members of the Arf family appears to be a conserved function, at least in opisthokonts.

17. Schlacht A, Mowbrey K, Elias M, Kahn RA, Dacks JB: **Ancient Complexity, Opisthokont Plasticity, and Discovery of the 11th Subfamily of Arf GAP Proteins.** *Traffic* 2013, **14**:636–649.
18. Ivanova AA, East MP, Yi SL, Kahn RA: **Characterization of Recombinant ELMOD (Cell Engulfment and Motility Domain) Proteins as GTPase-activating Proteins (GAPs) for ARF Family GTPases.** *Journal of Biological Chemistry* 2014, **289**:11111–11121.



19. East MP, Bowzard JB, Dacks JB, Kahn RA: **ELMO Domains, Evolutionary and Functional Characterization of a Novel GTPase-activating Protein (GAP) Domain for Arf Protein Family GTPases.** *Journal of Biological Chemistry* 2012, **287**:39538–39553.
20. Veltel S, Gasper R, Eisenacher E, Wittinghofer A: **The retinitis pigmentosa 2 gene product is a GTPase-activating protein for Arf-like 3.** *Nat Struct Mol Biol* 2008, **15**:373–380.
21. Su M-Y, Fromm SA, Zoncu R, Hurley JH: **Structure of the C9orf72 ARF GAP complex that is haploinsufficient in ALS and FTD.** *Nature* 2020, **585**:251–255.
- \*\*22. Su M-Y, Fromm SA, Remis J, Toso DB, Hurley JH: **Structural basis for the ARF GAP activity and specificity of the C9orf72 complex.** *Nat Commun* 2021, **12**:3786.

Along with Su et al. 2020, this paper demonstrates that a complex made up of three proteins, C9orf72, SMCR8 and WDR41, has potent GAP activity on Arf1, Arf5 and Arf6. These studies provide one of the first demonstrations at the molecular level of a cellular function for the C9orf72 protein, one of the major causative agents of the neurodegenerative disease amyotrophic lateral sclerosis (ALS). Interestingly, this complex localizes to lysosomes, which have not been shown to have resident Arf proteins, suggesting that this complex might either act in trans or be sequestered on lysosomal membranes to regulate its activity.

23. More K, Klinger CM, Barlow LD, Dacks JB: **Evolution and Natural History of Membrane Trafficking in Eukaryotes.** *Current Biology* 2020, **30**:R553–R564.
24. Purkanti R, Thattai M: **Genome doubling enabled the expansion of yeast vesicle traffic pathways.** *Sci Rep* 2022, **12**:11213.
25. Elias M, Brighthouse A, Castello CG, Field MC, Dacks JB: **Sculpting the endomembrane system in deep time: High resolution phylogenetics of Rab GTPases.** *Journal of Cell Science* 2012, doi:10.1242/jcs.101378.
26. Simakov O, Marlétaz F, Yue J-X, O’Connell B, Jenkins J, Brandt A, Calef R, Tung C-H, Huang T-K, Schmutz J, et al.: **Deeply conserved synteny resolves early events in vertebrate evolution.** *Nat Ecol Evol* 2020, **4**:820–830.
27. Singh PP, Isambert H: **OHNOLOGS v2: a comprehensive resource for the genes retained from whole genome duplication in vertebrates.** *Nucleic Acids Research* 2020, **48**:D724–D730.
28. Schlacht A, Dacks JB: **Unexpected Ancient Paralogs and an Evolutionary Model for the COPII Coat Complex.** *Genome Biology and Evolution* 2015, **7**:1098–1109.
29. Roustan V, Jain A, Teige M, Ebersberger I, Weckwerth W: **An evolutionary perspective of AMPK–TOR signaling in the three domains of life.** *Journal of Experimental Botany* 2016, **67**:3897–3907.

- \*30. Zhou Y, Amom P, Reeder SH, Lee BH, Helton A, Dobritsa AA: **Members of the ELMOD protein family specify formation of distinct aperture domains on the Arabidopsis pollen surface.** *eLife* 2021, **10**:e71061.

This paper represents the first functional characterization of ELMOD proteins in plants. ELMOD proteins in mammalian cells have a highly conserved region with an invariant arginine that is essential for GAP activity on Arf family proteins *in vitro*. The *Arabidopsis* ELMOD family proteins investigated in this study (MCR and ELMOD\_A) have the same conserved region with the invariant arginine, mutation of which abolished their capacity to complement phenotypes of mutants in *Arabidopsis*. These results support the conclusion that ELMOD proteins are Arf GAPs in plants.

- \*31. Billington K, Halliday C, Madden R, Dyer P, Barker AR, Moreira-Leite FF, Carrington M, Vaughan S, Hertz-Fowler C, Dean S, et al.: **Genome-wide subcellular protein map for the flagellate parasite Trypanosoma brucei.** *Nat Microbiol* 2023, **8**:533–547.

This open resource reports the subcellular localization of 89% of the encoded proteome (7766 proteins) of *Trypanosoma brucei*, one of the best developed molecular cell biological systems outside of animals, plants and fungi. A searchable database allows localization data to be obtained for proteins of interest for comparative cell biology.

32. Pipaliya SV, Thompson LA, Dacks JB: **The reduced ARF regulatory system in Giardia intestinalis pre-dates the transition to parasitism in the lineage Fornicata.** *International Journal for Parasitology* 2021, **51**:825–839.
33. Tikhonenkov DV, Mikhailov KV, Hehenberger E, Karpov SA, Prokina KI, Esaulov AS, Belyakova OI, Mazei YA, Mylnikov AP, Aleoshin VV, et al.: **New Lineage of Microbial Predators Adds Complexity to Reconstructing the Evolutionary Origin of Animals.** *Current Biology* 2020, **30**:4500-4509.e5.
34. Gomez-Navarro N, Miller EA: **COP-coated vesicles.** *Current Biology* 2016, **26**:R54–R57.
35. Yorimitsu T, Sato K, Takeuchi M: **Molecular mechanisms of Sar/Arf GTPases in vesicular trafficking in yeast and plants.** *Frontiers in Plant Science* 2014, **5**:411.
36. Herman EK, Yiangou L, Cantoni DM, Miller CN, Marciano-Cabral F, Anthonyrajah E, Dacks JB, Tsoulos AD: **Identification and characterisation of a cryptic Golgi complex in Naegleria gruberi.** *Journal of Cell Science* 2018, **131**:jcs213306.
37. Brandizzi F, Barlowe C: **Organization of the ER–Golgi interface for membrane traffic control.** *Nat Rev Mol Cell Biol* 2013, **14**:382–392.
38. Singh MK, Richter S, Beckmann H, Kientz M, Stierhof Y-D, Anders N, Fäßler F, Nielsen M, Knöll C, Thomann A, et al.: **A single class of ARF GTPase activated by several pathway-specific ARF-GEFs regulates essential membrane traffic in Arabidopsis.** *PLoS Genetics* 2018, **14**:e1007795.
39. Dacks JB, Robinson MS: **Outerwear through the ages: evolutionary cell biology of vesicle coats.** *Current Opinion in Cell Biology* 2017, **47**:108–116.

40. Popoff V, Langer JD, Reckmann I, Hellwig A, Kahn RA, Brügger B, Wieland FT: **Several ADP-ribosylation Factor (Arf) Isoforms Support COPI Vesicle Formation.** *Journal of Biological Chemistry* 2011, **286**:35634–35642.

\*41. Pennauer M, Buczak K, Prescianotto-Baschong C, Spiess M: **Shared and specific functions of Arfs 1–5 at the Golgi revealed by systematic knockouts.** *Journal of Cell Biology* 2021, **221**:e202106100.

This paper reports the characterization of CRISPR/Cas9-mediated genomic deletion for Arf1 and the three human proteins the most closely related to it, Arf3-5, comprising the Class I and Class II Arf proteins. This study reveals a high level of functional redundancy among these proteins, but also defines specific functions.

42. Newman LE, Schiavon CR, Zhou C, Kahn RA: **The abundance of the ARL2 GTPase and its GAP, ELMOD2, at mitochondria are modulated by the fusogenic activity of mitofusins and stressors.** *PLOS ONE* 2017, **12**:e0175164.

43. Turn RE, Linnert J, Gigante ED, Wolfrum U, Caspary T, Kahn RA: **Roles for ELMOD2 and Rootletin in ciliogenesis.** *MBoC* 2021, **32**:800–822.

44. Yadav S, Puthenveedu MA, Linstedt AD: **Golgin160 Recruits the Dynein Motor to Position the Golgi Apparatus.** *Developmental Cell* 2012, **23**:153–165.

45. Isabet T, Montagnac G, Regazzoni K, Raynal B, Khadali FE, England P, Franco M, Chavrier P, Houdusse A, Ménétrey J: **The structural basis of Arf effector specificity: the crystal structure of ARF6 in a complex with JIP4.** *The EMBO Journal* 2009, **28**:2835–2845.

46. Montagnac G, Sibarita J-B, Loubéry S, Daviet L, Romao M, Raposo G, Chavrier P: **ARF6 Interacts with JIP4 to Control a Motor Switch Mechanism Regulating Endosome Traffic in Cytokinesis.** *Current Biology* 2009, **19**:184–195.

47. Keren-Kaplan T, Bonifacino JS: **ARL8 Relieves SKIP Autoinhibition to Enable Coupling of Lysosomes to Kinesin-1.** *Current Biology* 2021, **31**:540-554.e5.

\*\*48. Keren-Kaplan T, Sarić A, Ghosh S, Williamson CD, Jia R, Li Y, Bonifacino JS: **RUFY3 and RUFY4 are ARL8 effectors that promote coupling of endolysosomes to dynein-dynactin.** *Nat Commun* 2022, **13**:1506.

This paper shows that Arl8 promotes the localization of endolysosomes towards the cell center in different cell types, and identifies RUFY3 and RUFY4 as effectors of Arl8 that mediate this effect through interaction with the dynein microtubule motor. The authors show that RUFY3 and RUFY4 interact with dynein-dynactin in cells, and using purified proteins, they demonstrate that RUFY3 interacts directly with dynein light intermediate chain 1 (DLIC1). The authors showed that RUFY3 is required for movement of endolysosomes toward the cell center upon cytosol alkalinization. This work and [54] together show the mechanism by which Arl8 mediates retrograde movement of endolysosomes along microtubules towards the cell center.

- \*\*49. Kumar G, Chawla P, Dhiman N, Chadha S, Sharma S, Sethi K, Sharma M, Tuli A: **RUFY3 links Arl8b and JIP4-Dynein complex to regulate lysosome size and positioning.** *Nat Commun* 2022, **13**:1540.

This paper identifies RUFY3 as an effector of Arl8 that promotes the juxtannuclear localization of endolysosomes. The authors show that RUFY3 physically interacts with JIP4, an adaptor protein that has previously been shown to link dynein-dynactin to organelle membranes through interaction with Arf6-GTP. They also show that upon nutrient starvation, RUFY3 depletion blocks the redistribution of endolysosomes towards the cell center. This work shows that two dynein-adaptors can interact to regulate lysosome positioning, and along with [53] demonstrates the mechanism by which Arl8 mediates retrograde movement of endolysosomes along microtubules towards the cell center.

50. Gillingham AK, Munro S: **Finding the Golgi: Golgin Coiled-Coil Proteins Show the Way.** *Trends in Cell Biology* 2016, **26**:399–408.
51. Ishida M, Bonifacino JS: **ARFRP1 functions upstream of ARL1 and ARL5 to coordinate recruitment of distinct tethering factors to the trans-Golgi network.** *Journal of Cell Biology* 2019, **218**:3681–3696.
52. Wilhelmi I, Grunwald S, Gimber N, Popp O, Dittmar G, Arumughan A, Wanker EE, Laeger T, Schmoranzler J, Daumke O, et al.: **The ARFRP1-dependent Golgi scaffolding protein GOPC is required for insulin secretion from pancreatic  $\beta$ -cells.** *Molecular Metabolism* 2021, **45**:101151.
- \*53. Dewees SI, Vargová R, Hardin KR, Turn RE, Devi S, Linnert J, Wolfrum U, Caspary T, Eliáš M, Kahn RA: **Phylogenetic profiling and cellular analyses of ARL16 reveal roles in traffic of IFT140 and INPP5E.** *MBoC* 2022, **33**:ar33.

By studying the phylogenetic profile (i.e. taxonomic distribution) of the previously unstudied Arf family member Arl16, the authors hypothesized that the protein is functionally connected to the cilium. The hypothesis has subsequently been confirmed by experimental work on mammalian cell lines, demonstrating the role of Arl16 in the delivery of specific proteins to the cilium and assigning thus a cellular function to one of the last functionally uncharacterized ancestral Arf family members.

- \*54. Liu Y-X, Sun W-Y, Xue B, Zhang R-K, Li W-J, Xie X, Fan Z-C: **ARL3 mediates BBSome ciliary turnover by promoting its outward movement across the transition zone.** *Journal of Cell Biology* 2022, **221**:e202111076.
- \*55. Dai J, Zhang G, Alkhofash RA, Mekonnen B, Saravanan S, Xue B, Fan Z-C, Betleja E, Cole DG, Liu P, et al.: **Loss of ARL13 impedes BBSome-dependent cargo export from *Chlamydomonas* cilia.** *Journal of Cell Biology* 2022, **221**:e202201050.

Using *arl13* mutants in *Chlamydomonas reinhardtii*, Dai et al. show that loss of Arl13 results in accumulation of membrane proteins in cilia, and provide evidence that Arl13 is involved in their exit from cilia. Arl3-GTP is localized near the transition zone between the cilium and cell body due to localized Sec13/BART GEF activity. Also working in in *Chlamydomonas*, Liu

et al. show that Arl3-GTP binds to retrograde BBSomes shed from the IFT train through specific interaction with phospholipase D (PLD), an effector of Arl3 that is present on shed BBSomes, thus providing a mechanism for their export from cilia. These two papers are the first to report new roles for Arl3 and Arl13 in promoting BBSome-mediated export of membrane proteins from the cilium.

56. Liu Y-X, Xue B, Sun W-Y, Wingfield JL, Sun J, Wu M, Lechtreck KF, Wu Z, Fan Z-C: **Bardet-Biedl syndrome 3 protein promotes ciliary exit of the signaling protein phospholipase D via the BBSome.** *eLife* 2021, **10**:e59119.
57. Partisani M, Baron CL, Ghossoub R, Fayad R, Pagnotta S, Abélanet S, Macia E, Brau F, Lacas-Gervais S, Benmerah A, et al.: **EFA6A, an exchange factor for Arf6, regulates early steps in ciliogenesis.** *Journal of Cell Science* 2021, **134**:jcs249565.
58. Yang S, Tang Y, Liu Y, Brown AJ, Schaks M, Ding B, Kramer DA, Mietkowska M, Ding L, Alekhina O, et al.: **Arf GTPase activates the WAVE regulatory complex through a distinct binding site.** *Science Advances* 2022, **8**:eadd1412.
59. Kollmar M, Lbik D, Enge S: **Evolution of the eukaryotic ARP2/3 activators of the WASP family: WASP, WAVE, WASH, and WHAMM, and the proposed new family members WAWH and WAML.** *BMC Res Notes* 2012, **5**:1–23.
60. Jang J-H, Lee CS, Hwang D, Ryu SH: **Understanding of the roles of phospholipase D and phosphatidic acid through their binding partners.** *Progress in Lipid Research* 2012, **51**:71–81.
61. Venditti R, Masone MC, Wilson C, De Matteis MA: **PI(4)P homeostasis: Who controls the controllers?** *Advances in Biological Regulation* 2016, **60**:105–114.
62. Highland CM, Fromme JC: **Arf1 directly recruits the Pik1-Frq1 PI4K complex to regulate the final stages of Golgi maturation.** *MBoC* 2021, **32**:1064–1080.
- \*\*63. Li F-L, Wu Z, Gao Y-Q, Bowling FZ, Franklin JM, Hu C, Suhandynata RT, Frohman MA, Airola MV, Zhou H, et al.: **Defining the proximal interaction networks of Arf GTPases reveals a mechanism for the regulation of PLD1 and PI4KB.** *The EMBO Journal* 2022, **41**:e110698.

The authors use a proximity labelling and mass spectrometry approach to identify effectors of 25 of the 29 human Arf family proteins. The approach was validated by the identification of many of the known Arf family effectors, and hence will provide an important resource for the community. Remarkably, only 2% of the total number of interactors identified in this study have been previously reported as effectors of Arf family proteins, indicating that there is much to be discovered about the functions of Arf family GTPases.

64. Houghton FJ, Makhoul C, Cho EH-J, Williamson NA, Gleeson PA: **Interacting partners of Golgi-localized small G protein Arl5b identified by a combination of in vivo proximity labelling and GFP-Trap pull down.** *FEBS Letters* 2022, **596**:2382–2399.

65. Hutchings J, Stancheva V, Miller EA, Zanetti G: **Subtomogram averaging of COPII assemblies reveals how coat organization dictates membrane shape.** *Nat Commun* 2018, **9**:4154.

\*66. Hooy RM, Iwamoto Y, Tudorica DA, Ren X, Hurley JH: **Self-assembly and structure of a clathrin-independent AP-1:Arf1 tubular membrane coat.** *Science Advances* 2022, **8**:eadd3914.

This paper reports the structure of an Arf1:AP-1 coat without clathrin. The adaptor complexes AP-1, AP-2 and in some cases AP-3 form the inner layer of a clathrin coat, but AP-4 does not generally form a coat with clathrin, and both AP-1 and AP-3 can form clathrin-independent coats. All of these coats are recruited or stabilized on membranes by Arf1. Clathrin coats form a rounded vesicular coat, whereas AP-1 and AP-3 can be found on tubular structures. Using cryo-electron tomography and subtomogram averaging, the authors determined the structure of a tubular Arf1:AP-1 complex, and provide a paradigm for how Arf1 and adaptor complexes can form a coat on membranes in the absence of clathrin.

67. Dodonova SO, Aderhold P, Kopp J, Ganeva I, Röhling S, Hagen WJH, Sinning I, Wieland F, Briggs JAG: **9Å structure of the COPI coat reveals that the Arf1 GTPase occupies two contrasting molecular environments.** *eLife* 2017, **6**:e26691.

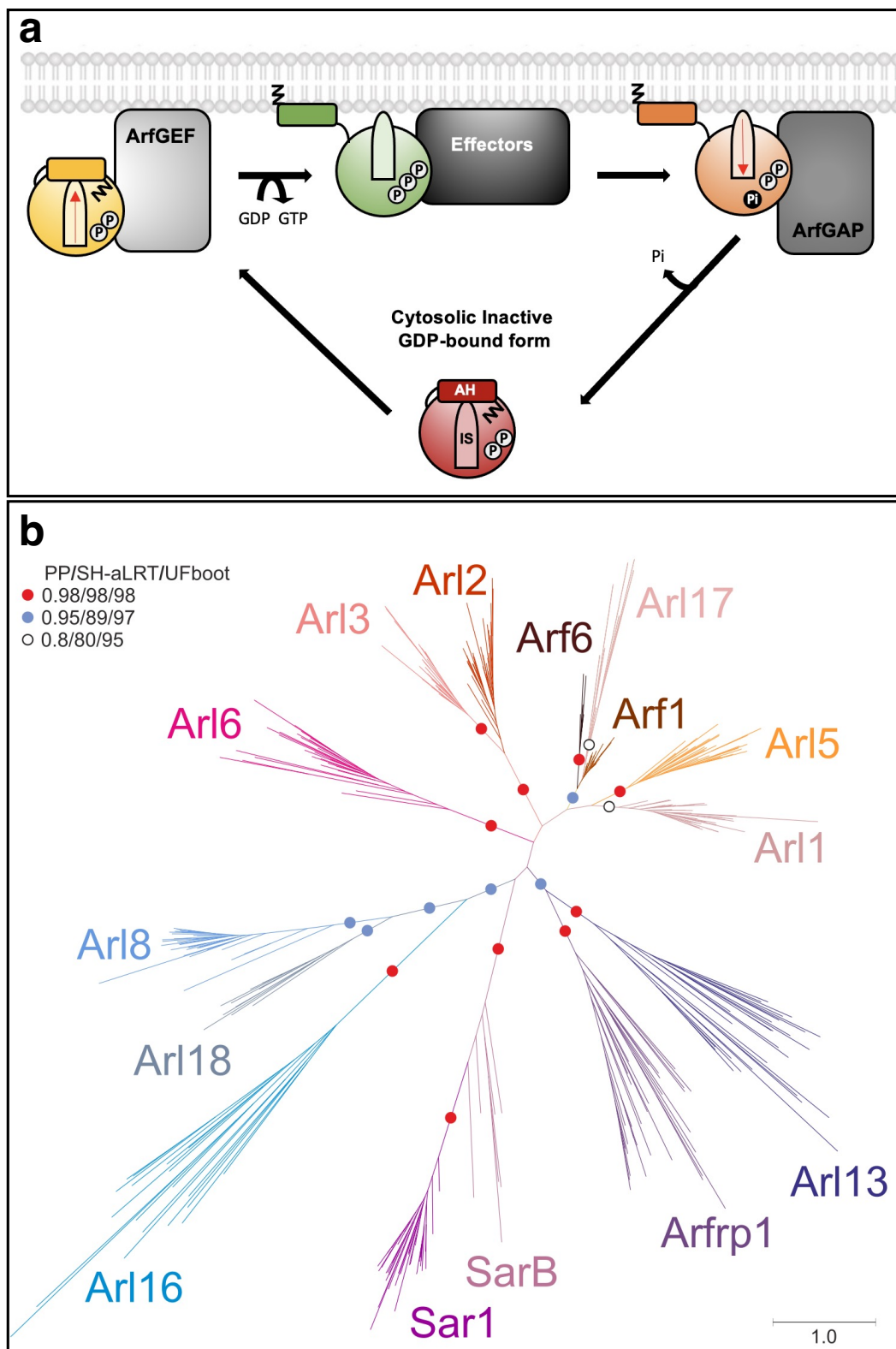
68. Xie B, Jung C, Chandra M, Engel A, Kendall AK, Jackson LP: **The Glo3 GAP crystal structure supports the molecular niche model for ArfGAPs in COPI coats.** *Advances in Biological Regulation* 2021, **79**:100781.

69. Bykov YS, Schaffer M, Dodonova SO, Albert S, Plitzko JM, Baumeister W, Engel BD, Briggs JA: **The structure of the COPI coat determined within the cell.** *eLife* 2017, **6**:e32493.

\*70. Muccini AJ, Gustafson MA, Fromme JC: **Structural basis for activation of Arf1 at the Golgi complex.** *Cell Reports* 2022, **40**:111282.

A cryo-EM structural analysis of the Gea2 Arf GEF, a member of the GBF1 Sec7 domain GEF subfamily, alone and in complex with Arf1. This study reveals a mechanism for membrane association of the GEF, and also reveals a conformational switch of the GEF during nucleotide exchange that promotes membrane insertion of its substrate Arf1.

Figures



**Figure 1. General overview of the function and phylogenetic diversity of the Arf family.**

a) Activation cycle of Arf family GTPases on membranes. The inactive form of an Arf family protein bound to GDP (red) is cytosolic with its amphipathic helix (AH) (often myristoylated,

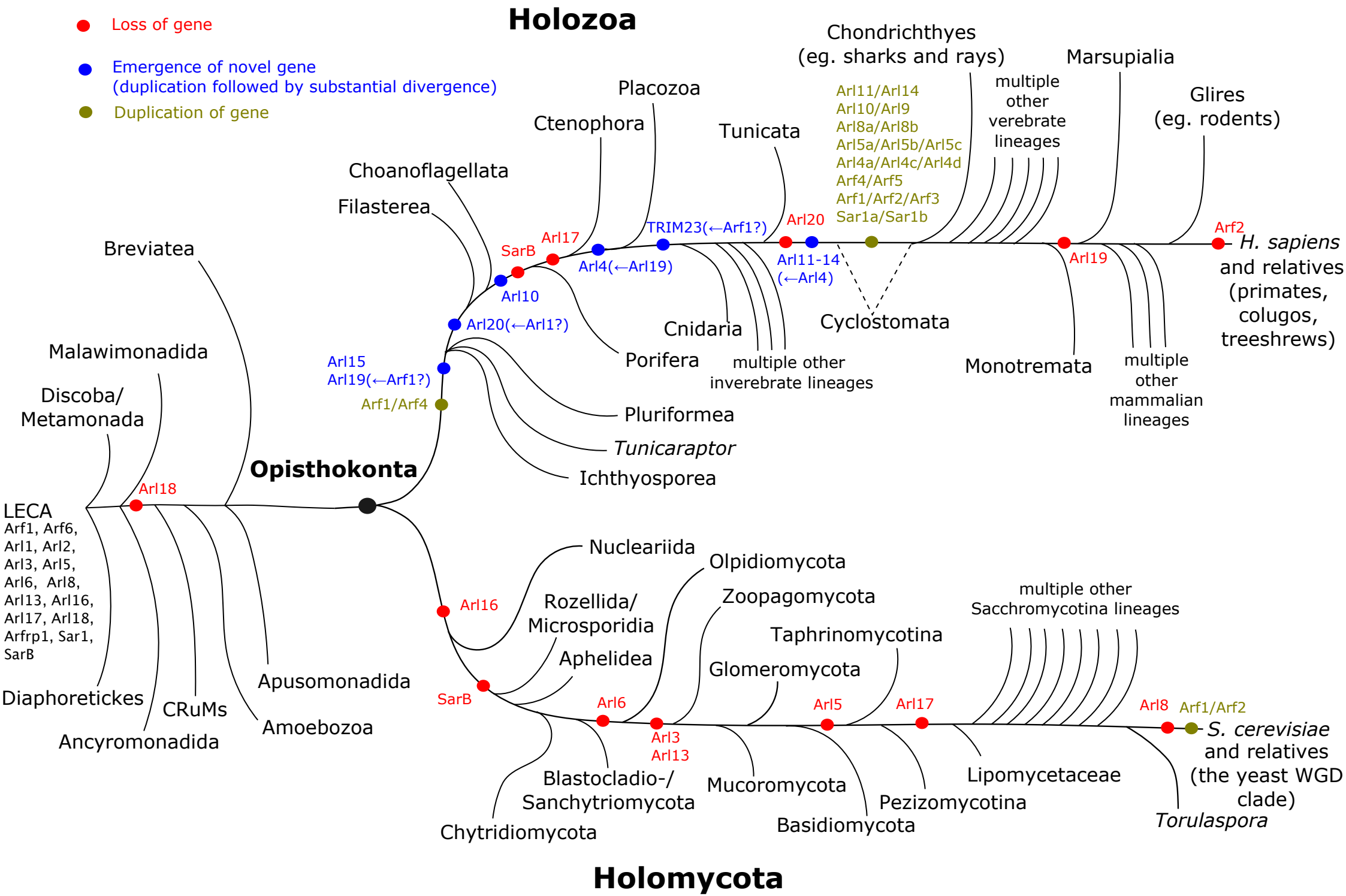
wavy black line) folded back on the core protein and its interswitch region retracted. Guanine nucleotide exchange factors (GEFs) (light grey) interact with Arf proteins (yellow) on the membrane, and trigger GDP to GTP exchange and the interswitch toggle (red arrow). This change in conformation of the interswitch region results in its movement into the hydrophobic pocket, displacing the AH. The active Arf-GTP (green) is bound to the membrane with the AH inserted into the cytosolic leaflet of the bilayer. Arf-GTP specifically interacts with other proteins, called effectors (dark grey), which carry out their downstream functions. GTPase activating proteins (GAPs) (grey) interact with Arf-GTP on the membrane and stimulate GTP hydrolysis and the interswitch toggle (red arrow). Arf-GDP is released to the cytosol with the AH folded back on the core protein. As shown here, Arf1 and several other Arf family proteins have been demonstrated to be cytosolic in their GDP-bound form, but others can bind membranes due to interaction of their GDP-bound form with membrane proteins.

b) Maximum likelihood phylogenetic tree (IQ-TREE, LG+I+G4 model) inferred from a reference set of Arf family protein sequences from diverse eukaryotes (adopted with modifications from Vargová et al. [7]). Clades corresponding to the 15 ancestral eukaryotic paralogs are annotated and distinguished by different colors. The SarB group appears as a paraphyletic grade basal to Sar1, but the unity of SarB as a separate paralog is supported by sharing specific intron positions and sequence signatures. Branches separating the different ancestral paralogs are marked to indicate branch support values according to the legend in the bottom right corner. The support values correspond to posterior probabilities (calculated with MrBayes), SH-aLRT test values, and ultrafast bootstrap values (both calculated with IQ-TREE).

Reproduced from Vargova et al. 2021 [7],

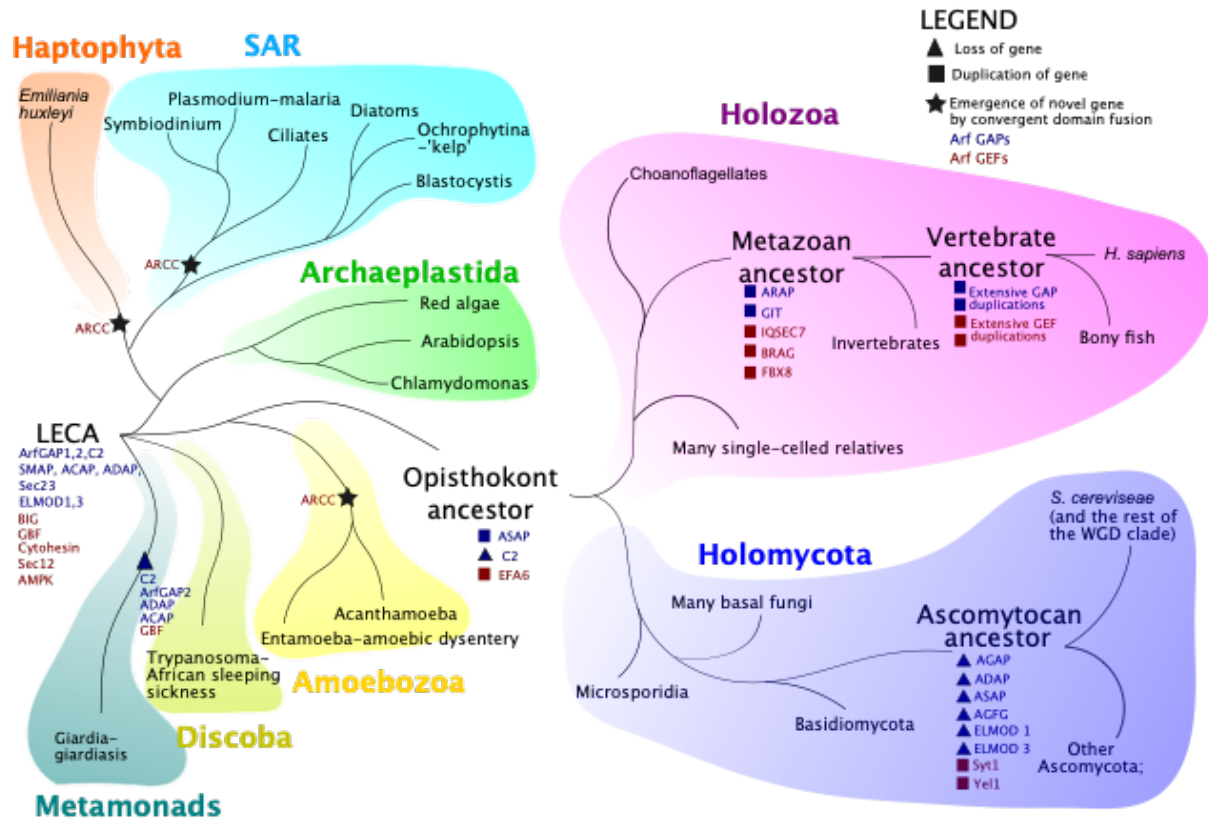
<https://academic.oup.com/gbe/article/13/8/evab157/6319025>. Licensed under [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/).





**Figure 2. The evolutionary trajectory of the Arf family from the last eukaryotic common ancestor (LECA) to budding yeast (*S. cerevisiae*) and humans.** The schematic phylogenetic tree is based on a synthesis of recent phylogenomic literature and shows the consecutive divergence of “side” lineages along the evolutionary paths towards the two focal taxa. Three types of evolutionary events shaping the Arf family are mapped onto the paths: loss of a particular gene; emergence of a “novel” Arf family member, which equals a duplication of one of the pre-existing genes followed by substantial divergence of one of the duplicates (when known, the parental gene is indicated in brackets, with an arrow pointing to the “novel” descendant gene); gene duplication, which corresponds to situations where the fate of the duplicates was more symmetrical (in terms of sequence divergence) than in the previous case. For the duplications, the resulting genes are listed (separated with a slash), with one of the descendants retaining the name of the parental gene (then listed first; eg. the Arf1/Arf4 pair in early Holozoa emerged by duplication of the ancestral Arf1 gene) or with the descendants designated by the name of the parental name extended by a suffix ( Sar1a and Sar1b resulting from duplication of the parental Sar1 gene). The history of the Arf family as presented in the figure is based on results reported by Vargová et al. [7] and Dewees et al. [53], with substantial updates based on checking the occurrence of the different Arf family members in additional eukaryote lineages that were not previously analysed. Briefly, for Arf family genes previously inferred to have been lost at a particular node of the yeast or human evolutionary lineages, the possible presence of putative orthologs was probed in previously untested taxa descendant to the respective node. Analogously, for Arf family members previously inferred to be evolutionary novelties of a particular organismal clade, the possible presence of orthologs was checked in the previously unsampled outgroup taxa. As a result, losses or origins of certain Arf family genes were thus pushed to more recent or older nodes, respectively. For instance, SarB, which was previously thought to have been lost already prior to the emergence of opisthokonts [7], was now newly identified in several opisthokont taxa (Nucleariida, Filasterea, Choanoflagellata), indicating it was lost independently in the early history of the Holomycota and Holozoa lineages. Similar re-evaluations implying a more recent loss concern Arl17 (not previously known to occur in basal holozoan lineages including sponges) and Arl19 (now identified in multiple vertebrate lineages including monotreme mammals). On the other hand, Arl15 and Arl19 now seem to have emerged already in the stem branch

of Holozoa, i.e. earlier than reconstructed previously, owing to their identification in previously untested representatives of the basal holozoan lineages. The assignment of Arf family genes in the newly analysed taxa to specific Arf family members was primarily based on BLAST searches against the curated sequence set reported by Vargová et al. [7], combined in some cases with the inspection of intron positions (which proved instrumental in defining orthologous relationships in some cases where sequence comparisons or standard phylogenetic analyses did not provide sufficient resolution [7]). The newly identified Arf family genes with details on the sequence sources are listed in supplementary Table S1 available XXXXX (URL to be added in revision). The set of gene duplications mapped to early vertebrate evolution (prior to the divergence of Chondrichthyes) presumably results from two WGD rounds, as follows from consulting the OHNOLOGS version 2.0 database (<http://ohnologs.curie.fr/>; [27]). The relation of the duplications with respect to the divergence of jawless vertebrates (cyclostomes) remains unclear, as indicated in the figure by the dashed lines connecting cyclostomes to the vertebrate main trunk at two alternative positions.



**Figure 3. Evolution of Arf GAPs and GEFs across eukaryotes.** This cartoon illustrates a highly simplified depiction of the eukaryotic tree. Some of the major groups of eukaryotes, and their interrelationships are shown, with emphasis on specific examples from the text or on major medically relevant organisms, with the diseases caused given after the hyphen in such cases. The diagram is not to scale. Note that the different groups highlighted are not at equivalent levels, and as in Figure 2, the Holomycota and Holozoa lineages are highlighted (left). Losses, duplications, and convergent gene fusions of GAP and GEF families are mapped, noting that only the losses explicitly stated in the text are mapped on the tree. Symbols, and colouration is as in the inset legend.

**Declaration of interests**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: