View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Apollo

Outerwear through the ages: Evolutionary cell biology of vesicle coats

Joel B. Dacks[a] and Margaret S. Robinson[b]
[a] Department of Cell Biology, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, Alberta, Canada, T6G 2H7. dacks@ualberta.ca
[b] Cambridge Institute for Medical Research, University of Cambridge, Cambridge UK, CB2 0XY. msr12@cam.ac.uk

Keywords: COPI, COPII, clathrin, protocoatomer, eukaryotic evolution

Abstract

Vesicular transport was key to the evolution of eukaryotes, and is essential for eukaryotic life today. All modern eukaryotes have a set of vesicle coat proteins, which couple cargo selection to vesicle budding in the secretory and endocytic pathways. Although these coats share common features (e.g., recruitment via small GTPases, β-propeller-α-solenoid proteins acting as scaffolds), the relationships between them are not always clear. Structural studies on the coats themselves, comparative genomics and cell biology in diverse eukaryotes, and the recent discovery of the Asgard archaea and their "eukaryotic signature proteins" are helping us to piece together how coats may have evolved during the prokaryote-to-eukaryote transition.

Highlights

- Three well characterised coats have common features but different architectures.
- Studies on diverse eukaryotes reveal ancient machinery, lineage-specific innovations.
- This enabled discovery of new complexes in a clearly related set of coats (HTAC-CCs).
- Asgard archea provide clues about ancestral coats.
- Knowing the homology of vesicle coats informs both evolution and function.

Funding: This work was supported by the Natural Sciences and Engineering Research Council of Canada [RES0021028] and the Wellcome Trust [086598] to JBD and MSR respectively. JBD is the Canada Research Chair (Tier II) in Evolutionary Cell Biology.

Introduction

The eukaryotic cell is compartmentalized into membrane-bound organelles, and it is this sequestration of molecules and pathways into separate environments that most distinguishes us from Bacteria and Archaea. The biogenesis, maintenance, and functioning of these organelles all depend upon molecules shuttling from one organelle to another by vesicular trafficking. This process begins with the recruitment of vesicle coat proteins onto a particular membrane, where they collect the right cargo and shape the membrane into a vesicle. There are three types of coats for which we now have a good mechanistic understanding: COPII, COPI, and clathrin plus AP-1 or AP-2 (Figure 1).

These three coats share many features that support the idea that they are derived from a common origin: a primordial coat complex [1]. All three follow the same basic mechanism of recruitment onto the membrane via a small GTPase, followed by cargo selection by a sub-complex of the coat, and membrane deformation, usually by a different sub-complex. The fact that these roles are played by members of the same protein families, such as Arf and Sar GTPases, and the detectable sequence homology between members of the coat complexes themselves, bring robust analytical evidence for a deep ancient relationship. This idea was extended by the protocoatomer hypothesis, which was initially proposed to link the nuclear pore complex to vesicle coats, based on the common architecture of some of their subunits: a β-propeller domain fused to an α-solenoid domain [2] (see Box 1).

The idea that these vesicle coats are evolutionarily linked is now well supported (see [1,3]). This framework can be used to better understand the ways in which the coats have evolved both similarities and differences in their modern mechanisms of coupling cargo selection to vesicle formation.

The vesicle coats: Same pieces, same job, different architecture

From the point of view of a newly synthesized membrane protein, the first coat to be encountered is COPII, which selects cargo for transport from the ER to the Golgi apparatus. COPII vesicle formation is initiated by the recruitment of the small GTPase Sar1 onto the ER membrane, via the resident transmembrane protein Sec12, which acts as its guanine nucleotide exchange factor. The GTP-bound Sar1 then recruits the Sec23-Sec24 complex. Although Sec23 and Sec24 are likely homologous [4], they have diverged functionally so that Sec23 binds Sar1 and has GAP activity, while Sec24 paralogues bind to cargo proteins. The Sec23/Sec24 complex in turn recruits the Sec13/Sec31 complex. Sec13 forms a nearly complete β-propeller, while Sec31 consists of a β-propeller followed by an α-solenoid. These assemble into a cage and deform the membrane into a vesicle [5] (Figure 1 and Box 1).

The COPI coat acts downstream from the COPII coat, retrieving membrane proteins that have escaped from the ER and entered the Golgi (Figure 2). Like COPII, COPI is recruited onto membranes via a small GTPase, but the GTPase is Arf, which belongs to the same family as Sar1. The COPI coat, or coatomer, is a member of the heterotetrameric adaptor complex-containing coat (HTAC-CC) family. This is an umbrella term for vesicle coats whose core contains a complex (HTAC) that has a quaternary structure comprising two large subunits of ~100kD, a medium-sized subunit of ~50kD, and a small subunit of ~20kD. The two large subunits fall into two families: the β family and the "EGADZ" (ε/γ/α/δ/ζ) family, with each HTAC containing one of each (Figure 1 and Table 1). Both types of large subunit consist of an α-solenoid "trunk", a flexible linker, and an appendage domain that binds accessory proteins. The medium-sized subunits consist of a longin domain followed by a "μ homology domain" (MHD), while the small subunits consists of a single longin

domain. Two of the other components of the COPI coat consist of a pair of β-propellers followed by an α-solenoid, with one of the β-propellers in direct contact with the membrane [6].

The clathrin coat has yet another arrangement of β-propellers and α-solenoids, with the solenoids wrapping around each other to form the bars of the cage, while the propellers point inwards towards the vesicle and make contact with the HTAC components of the coat, AP-1 or AP-2 (Figure 1). AP-1 and AP-2 act at intracellular membranes and at the plasma membrane respectively (Figure 2). AP-1 requires Arf to be recruited onto membranes, and uses its large subunits to interact with Arf in an identical manner to the HTAC component of the COPI coat [6-8]. In contrast, AP-2 does not appear to require a small GTPase, although it does need a lipid, phosphatidylinositol 4,5-bisphosphate (PIP2), which may be playing a similar role of "marking out" the target membrane [9]. But these interactions are not sufficient, because both Arf and PIP2 are found on other membranes that do not recruit AP-1 or AP-2. The current view is that recruitment is due to coincidence detection, involving multiple low affinity interactions, not only with small GTPases and phosphoinositides, but also with cargo, clathrin, and other peripheral membrane proteins [10].

Understanding vesicle coats through an evolutionary lens

The composition and action of these three coat complexes can be compared from a strictly functional or biophysical point of view. However, taking an evolutionary perspective provides a different dimension entirely. Knowing the relationship between the various coats allows our mechanistic understanding to be framed as conservation, divergent evolution, or convergence, giving us insight into the overall dynamics of how membrane transport might have evolved and might still be evolving.

This evolutionary perspective is already inherent in the use of common model systems (e.g., Drosophila, C. elegans, Saccharomyces). However, membrane trafficking is a feature of all eukaryotes, and while yeast to man is a large evolutionary distance, it still spans only one of the large-scale taxonomic divisions (or Supergroups) of eukaryotes (Figure 3A) [11]. Eukaryotes encompassed in the remaining diversity include the plants that feed us and the pathogens that infect them, algae and protists that are critical nodes in the global food web, and parasites that take a tremendous toll on human health world-wide [11].

Microbial genomics has allowed us to peer into the genomes of these diverse eukaryotes and assess what they have in common and what sets them apart. These same data allow us to reconstruct back in cellular evolutionary history to understand what aspects of the membrane trafficking system emerged at what times and begin to tease apart mechanisms for endomembrane organelle evolution.

Notably, the basic machinery of membrane trafficking at the protein family level, and often at the organelle-specific sub-family level, seems conserved across the span of eukaryotic diversity [12,13 and references therein]. Comfortingly, although model systems outside the animals and fungi are less prevalent and developed, the data collected for the membrane trafficking components suggest that they are performing equivalent roles at equivalent cellular locations, indicating functional homology, retained from their common evolutionary origins [14 and references therein].

Despite some informative lineage-specific innovations, the basic functional characteristics of orthologues seem to be conserved, giving confidence to the further use of sequence homology to guide our evolutionary cell biology of membrane trafficking.

The overall implications are two-fold. First, this indicates that the basic principles of membrane trafficking derived in animals and fungi can be applied to diverse eukaryotes. Second, it also means that the ancestor that gave rise to all existing eukaryotes, the Last Eukaryotic Common Ancestor or LECA, must have

been a remarkably sophisticated microbial eukaryote. In keeping with the sophistication deduced for other cellular systems such as cytoskeleton and mitochondria [12], the LECA possessed a complement of trafficking machinery that exceeds that seen in some well known eukaryotes today, including yeast [12,13 and references therein]. Now the challenge is to map the changes both in component loss and in acquisition of new complexity that have taken place in the descendents of the LECA, and to use the information for a better understanding of both evolution and function.

A sophisticated ancient set of coats has been sculpted by both loss and innovation

All three of the well-characterised coats, COPII, COPI, and clathrin plus APs, are found in nearly every eukaryote, and so must have been present in LECA [15,16]. In addition to AP-1 and AP-2, there are other AP complexes that appear to function independently of clathrin, raising the question of what (if any) scaffolding protein they use. AP-3 and AP-4 are associated with early/recycling endosomes and with the TGN respectively (Figure 2). In both cases, Arf is essential for their membrane recruitment, and both recognise YxxΦ and dileucine-containing cargo proteins, as do AP-1 and AP-2. APs 1, 2, and 3 are found in most eukaryotes, but several lineages have lost AP-3, including prominent parasitic organisms such as the Apicomplexa [17] and important oceanic algae such as the haptophytes [18]. AP-4 has been lost multiple times [15,19] (Figure 3A).

A more distantly related AP complex, AP-5, localises to late endosomes and lysosomes (Figure 2) and does not appear to use Arf, although there is a requirement for a phosphoinositide, PI3P [20,21]. AP-5 is tightly associated with two other proteins: SPG11 and SPG15. SPG11 is predicted to consist of a β-propeller followed by an α-solenoid, while SPG15 is predicted to be mainly α-solenoid but with a FYVE domain, presumably involved in PI3P binding.

The most recently discovered HTAC-containing coat, TSET, also has two subunits with β-propellers followed by an α-solenoid [15,22]. Like AP-4, AP-5 and TSET are both ancient but have been lost from multiple lineages [15]. Opisthokonts have retained only the C-terminal μ homology domain of the TSET medium subunit, which has been fused with a new N-terminal BAR domain, to generate the muniscin family. Intriguingly, although TSET is more closely related to COPI, it is associated with the plasma membrane and plays a key role in clathrin-mediated endocytosis in plants [22], while the muniscins in fungi and animals also contribute to the clathrin pathway at the cell surface [23].

In addition to the HTACs, clathrin-coated vesicles also make use of "alternative adaptors" or "CLASPs" (clathrin-associated sorting proteins), which bind to cargo proteins that cannot interact directly with AP-1 or AP-2 [24]. Some of these are ancient: for instance, the AP180/CALM family and the epsinR family are found throughout the diversity of eukaryotes, albeit with clear differences in their mechanisms having arisen [25-27]. Other alternative adaptors are more recent innovations, often arising from gene rearrangements. For example, the GGA family, found only in opisthokonts, has an N-terminal domain derived from Tom1 (Target of Myb protein 1) and a C-terminal domain derived from the flexible linker and appendage of γ-adaptin, the EGADZ family member in AP-1. Haptophytes also have novel proteins containing domains derived from EGADZ proteins [18], while kinetoplastid-specific clathrin-interacting proteins (CAPs) have also been reported [28].

In animals, there is an intimate relationship between signaling and endocytosis [29], and machinery has evolved to facilitate cross-talk between the two pathways. For instance, the human proteins Dab2, Numb, and ARH contain PTB domains, which are found only in holozoans [30], and which bind to NPXY motifs on

proteins involved in signalling, such as growth factor receptors and integrins. This enables these receptors to be switched off or alternatively to signal from endosomes [29].

Digging down to common origins

With clear complexity in vesicle coats already established by the time of the LECA (Figure 3A), the question becomes: how are these complexes related to one another, and from where did they come? For one subset of the vesicle coat components, molecular evolutionary analyses provide very robust answers. For over 25 years, there has been evidence of homology between the AP complexes and components of the COPI coat [31,32]. Phylogenetic analyses of the HTACs have revealed an order of emergence (Figure 2 inset), with COPI and TSET on one side of a deep evolutionary divide and the AP complexes on the other [33]. The tree of APs themselves is also resolved, with AP-5 diverging first, followed by AP-3, AP-4, and most recently AP-1 and AP-2 [15]. Indeed, many eukaryotes even share a common β subunit, meaning that the gene duplications producing separate AP-1 and AP2 complexes had not yet occurred for the β subunit by the time that the LECA diverged into its descendent lineages [19,34].

Unfortunately, while common mechanisms and the presence of proteins with the protocoatomer architecture strongly argues for homology between the HTAC-CCs and COPII, the relationship of these vesicle coats and the other protocoatomer-containing complexes, such as the nuclear pore and the intra-flagellar transport complex [1,3], is unclear. There are currently no robust data to speak to this issue. In our view, the most important unanswered questions in the area of vesicle coat evolution are how these complexes are inter-related, and based on this, what is the inferred emergence order of the corresponding organelles

Beyond the open question of how the many coats came from one, is the question of how vesicle-forming coats originated at all. This can be approached by recognizing that many parts of the vesicle coats are built from a common set of protein domains. These building blocks include not only the β-α/protocoatomer proteins, but longin domains, coiled coils, and small GTPases. Such domains are, in fact, present in other pieces of vesicle formation and fusion machinery as well (e.g., Rabs and their GEFs, tethering complexes, and SNAREs), and so tracing the phylogenetic history of proteins with these building blocks is key to understanding the history of the entire endomembrane system [33]. Because many of the vesicle trafficking proteins comprise fusions of the building blocks, understanding the point at which those fusions occurred is also crucial. Since phylogenetic analyses of many protein families involved in membrane-trafficking [13] have demonstrated that the relevant gene duplications and gene fusions predate the LECA by a considerable length of time, this question is only addressed by looking in prokaryotic genomes for homologous genes, thus pinpointing the prokaryotic lineages that contributed the various endomembrane system components.

This type of analysis received a massive boon in 2015 with the description of the Lokiarchaota [35] and again in early 2017 with the description of the larger group in which it is placed, the Asgard archaeal clade [36]. Prior to these reports, the identity of the archaeal lineage that gave rise to eukaryotes had been elusive, hampering robust analyses into the origins of membrane-trafficking machinery. The Asgard archaea appear to be the group from which eukaryotes emerged, and their genomes encode proteins that are either the closest orthologues of key membrane-trafficking components or progenitors of these important protein families. These include the first identification of a bona fide longin domain, expanded small GTPase families, BAR domain proteins, and ESCRT proteins, demonstrating that not only the vesicle coats discussed here but also unrelated trafficking proteins are derived from an archaeal origin [33,35]. Most strikingly, in the genomes of one subset of the

Asgard archaea (Thorarchaea) are encoded Sec23/Sec24 homologues and a proposed progenitor of the protocoatomer [36]. Importantly, the genomes of Asgard archaea are still several steps away from encoding the full set of components needed to make and accept a vesicle. Nonetheless, the origins of many membrane-trafficking components, including vesicle coats, can be found in the ancestor that eukaryotes share with the Asgard archaea (Figure 3A). As environmental sampling continues, and cell biological investigation in these lineages gets underway, we expect more exciting discoveries in the near future.

Sorting convergence from homology

Knowing that the major vesicle coats are homologous to one another allows us to interpret the biophysical and mechanistic data in a comparative framework. The simplest expectation, if the coats are homologous, is that they will share features retained from their ancestral state. Observations of common components (e.g., Arf use by both COPI and AP complexes) and the broader shared mechanism of a core cargo adaptor plus β-α scaffold are all interpreted as retention of an ancestral feature.

But how to interpret differences? For instance, the AP complexes undergo a conformational change upon recruitment, opening up to expose binding sites for cargo and clathrin that are inaccessible when they are in the cytosol [37,38]. COPI also appears to exist in both open and closed conformations, with the membrane-associated version seen by EM tomography appearing as a “hyper-open” form [6]. However, there is no evidence that COPI can bind to YxxΦ or dileucine motifs, the two sorting signals on cargo proteins that interact with APs 1-4. While a binding site for a different motif, $Wx_{n(1-6)}W$, has been identified on the COPI medium subunit, this binding site is accessible even in the closed conformation, and its only known binding partner is a tethering protein rather than a cargo protein [39]. This leads to a hypothesis that the evolution of cargo specificity can be seen in the divergent evolution of cargo binding sites from an ancestral mechanism of conformational change.

This base assumption of retained function in the face of apparent differences in the coats can also point to cases where further experimental work is needed. Because the APs function in cargo selection, for many years it was assumed that the HTAC component of the COPI coat was also involved in sorting, while α and β’ COP formed an outer scaffold to deform the membrane [40]. However, recent EM tomography studies have shown that although the COPI coat is highly ordered, with individual coatomers coming together to form triads, it doesn’t have a cage-like structure sitting over an inner layer. Instead, both subcomplexes make contact with the membrane [6]. In addition, the only well-characterised interaction between the COPI coat and cargo proteins is via the β-propellers of α and β’ COP, which bind to KKXX/KXKXX motifs on escaped ER proteins [41]. Precisely how the membrane is deformed is less clear, although the inherent curvature of the triads and the insertion of the Arf helices into the lipid bilayer are thought to contribute. Indeed, the assumption that the cage formed by Sec13/Sec31 or by clathrin is the major driving force for vesicle formation has also recently been called into question [42,43].

In cases where the coats function differently, it is possible that convergent evolution is at play. Convergence, i.e., the same phenotype arrived at from independent evolutionary histories, is unlikely in a molecular system with many components or high complexity. However, it can play a role in simple traits such as fusion of two elements into a protein. Most obviously with respect to vesicle coat evolution, the possibility of convergence needs to be incorporated within the scope of the protocoatomer hypothesis [1,2]. Although the sum of data supporting the protocoatomer hypothesis is overwhelming and the fundamental theory is not in question, it has become clear that domain fusions, once held to be so rare and stable as to be immune to convergence or fission, are much more labile than we thought

[44]. Most acutely, there are now two potential, and very distantly related, prokaryotic candidates for the source of the eukaryotic protocoatomer progenitor [36,45]. Since only one of these candidates actually gave rise to the eukaryotic protocoatomer protein, the other must have arisen by convergence (Figure 3B). It has already been said that the possible role of independent fusions and fissions of protocoatomer proteins in the early history of the eukaryotic membrane-trafficking system (and beyond to the nucleus, and intra-flagellar transport system) needs to be better explored [1]. This does not undermine the fundamental protocoatomer hypothesis, it adds a layer of subtlety to an already robust theory.

Conclusions

Our understanding of vesicle coat evolution has greatly benefited from a comparative approach. Investigating diverse microbial eukaryotes has shown a pattern of ancient complexity with loss of expendable components in some modern lineages and bursts of innovation in others. Looking even further afield to our closest archaeal relatives has even shown us the earliest footsteps towards a membrane-trafficking system. This information is helping us to gain a better understanding of the forces and events that have shaped this critical system and giving us a framework for investigations into how it functions today.

Acknowledgements The authors would like to thank D. Devos, L. Barlow, J. Parrish, J. Hirst, and P. Manna for helpful comments and help with figures.

References

1. Field MC, Dacks JB: **First and last ancestors: reconstructing evolution of the endomembrane system with ESCRTs, vesicle coat proteins, and nuclear pore complexes**. *Curr. Opin. Cell Biol.* 2009, **21**:4-13.
2. Devos D, Dokudovskaya S, Alber F, Williams R, Chait BT, Sali A, Rout MP: **Components of coated vesicles and nuclear pore complexes share a common molecular architecture**. *PLoS Biol.* 2004, **2**:e380.
3. Field MC, Sali A, Rout MP: **Evolution: On a bender--BARs, ESCRTs, COPs, and finally getting your coat**. *J. Cell Biol.* 2011, **193**:963-972.
4. Tang BL, Kausalya J, Low DY, Lock ML, Hong W: **A family of mammalian proteins homologous to yeast Sec24p**. *Biochem. Biophys. Res. Commun.* 1999, **258**:679-684.
5. Lee MC, Miller EA, Goldberg J, Orci L, Schekman R: **Bi-directional protein transport between the ER and Golgi**. *Annu. Rev. Cell Dev. Biol.* 2004, **20**:87-123.

6. Dodonova SO, Diestelkoetter-Bachert P, von Appen A, Hagen WJ, Beck R, Beck M, Wieland FT, Briggs JA: **VESICULAR TRANSPORT. A structure of the COPI coat and the role of coat proteins in membrane vesicle assembly. *Science* 2015, **349**:195-198.
This paper reveals for the first time the structure of a COPI-coated vesicle, through the use of in vitro budding and cryo-EM tomography. Although it had been widely assumed that the architecture of the COPI coat would resemble that of the clathrin and the COPII coats, in fact it is completely different.

7. Yu X, Breitman M, Goldberg J: **A structure-based mechanism for Arf1-dependent recruitment of coatomer to membranes**. *Cell* 2012, **148**:530-542.
8. Ren X, Farías GG, Canagarajah BJ, Bonifacino JS, Hurley JH: **Structural basis for recruitment and activation of the AP-1 clathrin adaptor complex by Arf1**. *Cell* 2013, **152**:755-767.
9. Honing S, Ricotta D, Krauss M, Spate K, Spolaore B, Motley A, Robinson M, Robinson C, Haucke V, Owen DJ: **Phosphatidylinositol-(4,5)-bisphosphate regulates sorting signal recognition by the clathrin-associated adaptor complex AP2**. *Mol. Cell* 2005, **18**:519-531.
10. Wenk MR, De Camilli P: **Protein-lipid interactions and phosphoinositide metabolism in membrane traffic: insights from vesicle recycling in nerve terminals**. *Proc. Natl. Acad. Sci. USA* 2004, **101**:8262-8269.
11. Adl SM, Simpson AG, Lane CE, Lukeš J, Bass D, Bowser SS, Brown MW, Burki F, Dunthorn M, Hampl V, et al.: **The revised classification of eukaryotes**. *J. Eukaryot. Microbiol.* 2012, **59**:429-493.
12. Koumandou VL, Wickstead B, Ginger ML, van der Giezen M, Dacks JB, Field MC: **Molecular paleontology and complexity in the last eukaryotic common ancestor**. *Crit. Rev. Biochem. Mol. Biol.* 2013, **48**:373-396.
13. Schlacht A, Herman EK, Klute MJ, Field MC, Dacks JB: **Missing pieces of an ancient puzzle: evolution of the eukaryotic trafficking system**. *Cold Spring Harb. Perspect. Biol.* 2014, **In press**:In press.

14. Klinger CM, Ramirez-Macias I, Herman EK, Turkewitz AP, Field MC, Dacks JB: **Resolving the homology-function relationship through comparative genomics of membrane-trafficking machinery and parasite cell biology. *Mol. Biochem. Parasitol.* 2016, **209**:88-103.
Evolutionary cell biology across eukaryotes has been based extensively on genomics, leaving open the question of whether the same gene in the same organisms performs the same function. This meta-analysis addresses this

question for proteins of the endocytic system showing that in emerging model organisms from across the eukaryotic tree functional homology seems to be largely preserved. The analysis is also a compendium of the molecular cell biological characterizations beginning to take place in non-opisthokont model organisms.

15. Hirst J, Schlacht A, Norcott JP, Traynor D, Bloomfield G, Antrobus R, Kay RR, Dacks JB, Robinson MS: **Characterization of TSET, an ancient and widespread membrane trafficking complex**. *eLife* 2014, **3**:e02866.
16. Schlacht A, Dacks JB: **Unexpected ancient paralogs and an evolutionary model for theCOPII coat complex**. *Genome Biol. Evol.* 2015, **5**:1098-1109.
17. Nevin WD, Dacks JB: **Repeated secondary loss of adaptin complex genes in the Apicomplexa**. *Parasitol. Int.* 2009, **58**:86-94.
18. Lee LJ, Klute MJ, Herman EK, Read B, Dacks JB: **Losses, expansions, and novel subunit discovery of adaptor protein complexes in Haptophyte algae**. *Protist* 2015, **166**:585-597.
19. Boehm M, Bonifacino JS: **Adaptins: the final recount**. *Mol. Biol. Cell* 2001, **12**:2907-2920.
20. Hirst J, Barlow LD, Francisco GC, Sahlender DA, Seaman MNJ, Dacks JB, Robinson MS: **The fifth adaptor protein complex**. *PLoS Biol.* 2011, **9**:e1001170.
21. Hirst J, Borner GHH, Edgar J, Hein MY, Mann M, Buchholz F, Antrobus R, Robinson MS: **Interaction between AP-5 and the hereditary spastic paraplegia proteins SPG11 and SPG15**. *Mol. Biol. Cell* 2013, **24**:2558-2569.
22. Gadeyne A, Sanchez-Rodriguez C, Vanneste S, Di Rubbo S, Zauber H, Vanneste K, Van Leene J, De Winne N, Eeckhout D, Persiau G, et al.: **The TPLATE adaptor complex drives clathrin-mediated endocytosis in plants**. *Cell* 2014, **156**:691-704.
23. Umasankar PK, Ma L, Thieman JR, Jha A, Doray B, Watkins SC, Traub LM: **A clathrin coat assembly role for the muniscin protein central linker revealed by TALEN-mediated gene editing**. *eLife* 2014, **3**:e04137.
24. Traub LM: **Tickets to ride: selecting cargo for clathrin-regulated internalization**. *Nat. Rev. Mol. Cell Biol.* 2009, **10**:583-596.
25. Miller SE, Sahlender DA, Graham SC, Höning S, Robinson MS, Peden AA, Owen DJ: **The molecular basis for the endocytosis of small R-SNAREs by the clathrin adaptor CALM**. *Cell* 2011, **147**:1118-1131.
26. Wang J, Gossing M, Fang P, Zimmermann J, Li X, von Mollard GF, Niu L, Teng M: **Epsin N-terminal homology domains bind on opposite sides of two SNAREs.** *Proc. Natl. Acad. Sci. USA* 2011, **108**:12277-12282.
27. Gabernet-Castello C, Dacks JB, Field MC: **The single ENTH-domain protein of trypanosomes; endocytic functions and evolutionary relationship with epsin**. *Traffic* 2009, **10**:894-911.
28. Adung'a VO, Gadelha C, Field MC: **Proteomic analysis of clathrin interactions in trypanosomes reveals dynamic evolution of endocytosis**. *Traffic* 2013, **14**:440-457.
29. Di Fiore PP, von Zastrow M: **Endocytosis, signaling, and beyond**. *Cold Spring Harb. Perspect. Biol.* 2014, **6**:a016865.
30. Suga H, Torruella G, Burger G, Brown MW, Ruiz-Trillo I: **Earliest Holozoan expansion of phosphotyrosine signaling**. *Mol. Biol. Evol.* 2014, **31**:517-528.
31. Duden R, Griffiths G, Frank R, Argos P, Kreis TE: **b-COP, a 100 kD protein associated with non-clathrin-coated vesicles and the Golgi complex, shows homology to b-adaptin**. *Cell* 1991, **64**:649-665.
32. Serafini T, Stenbeck G, Brecht A, Lottspeich F, Orci L, Rothman JE, Wieland FT: **A coat subunit of Golgi-derived non-clathrin-coated vesicles with homology to the clathrin-coated vesicle coat protein b-adaptin**. *Nature*

1991, **349**.

33. Klinger CM, Spang A, Dacks JB, Ettema TJ: **Tracing the archaeal origins of eukaryotic membrane-trafficking system building blocks**. *Mol. Biol. Evol.* 2016, **33**:528-541.
34. Dacks JB, Poon PP, Field MC: **Phylogeny of endocytic components yields insight into the process of nonendosymbiotic organelle evolution**. *Proc. Natl. Acad. Sci. USA* 2008, **105**:588-593.

35. Spang A, Saw JH, Jørgensen SL, Zaremba-Niedzwiedzka K, Martijn J, Lind AE, van Eijk R, Schleper C, Guy L, Ettema TJ: **Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* 2015, **521**:173-179.
This paper identifies for the first time the archael lineage that gave rise to eukaryotes, massively enhancing our ability to delve into the origins of the eukaryotic cell.

36. Zaremba-Niedzwiedzka K, Caceres EF, Saw JH, Bäckström D, Juzokaite L, Vancaester E, Seitz KW, Anantharaman K, Starnawski P, Kjeldsen KU, et al.: **Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature* 2017, **54**:353-358.
This paper builds upon Spang et al. in identifying the larger entire clade of organisms from which eukaryotes arose. This showed that these archaea are not necessarily extremophiles. It also showed the presence of proteins that are progenitors for key components of the membrane-trafficking system, including a candidate for the protocoatomer upon which the main three vesicle coats are based.

37. Jackson LP, Kelly BT, McCoy AJ, Gaffry T, James LC, Collins BM, Höning S, Evans PR, Owen DJ: **A large-scale conformational change couples membrane recruitment to cargo binding in the AP2 clathrin adaptor complex**. *Cell* 2010, **141**:1220-1229.
38. Kelly BT, Graham SC, Liska N, Dannhauser PN, Höning S, Ungewickell EJ, Owen DJ: **AP2 controls clathrin polymerization with a membrane-activated switch**. *Science* 2014, **345**:459-463.
39. Suckling RJ, Poon PP, Travis SM, Majoul IV, Hughson FM, Evans PR, Duden R, Owen DJ: **Structural basis for the binding of tryptophan-based motifs by δ-COP**. *Proc. Natl. Acad. Sci. USA* 2015, **112**:14242-14247.
40. Lee C, Goldberg J: **Structure of coatomer cage proteins and the relationship among COPI, COPII, and clathrin vesicle coats**. *Cell* 2010, **142**:123-132.
41. Jackson LP, Lewis MJ, Kent HM, Edeling MA, Evans PR, Duden R, Owen DJ: **Molecular basis for recognition of dilysine trafficking motifs by COPI**. *Dev. Cell* 2012, **11**:1255-1262.
42. Kirchhausen T: **Bending membranes**. *Nat. Cell Biol.* 2012, **14**:906-908.
43. Stachowiak JC, Brodsky FM, Miller EA: **A cost-benefit analysis of the physical mechanisms of membrane curvature**. *Nat. Cell Biol.* 2013, **15**:1019-1027.
44. Leonard G, Richards TA: **Genome-scale comparative analysis of gene fusions, gene fissions, and the fungal tree of life**. *Proc. Natl. Acad. Sci. USA* 2012, **109**:21402-21407.
45. Santarella-Mellwig R, Franke J, Jaedicke A, Gorjanacz M, Bauer U, Budd A, Mattaj IW, Devos DP: **The compartmentalized bacteria of the planctomycetes-verrucomicrobia-chlamydiae superphylum have membrane coat-like proteins**. *PLoS Biol.* 2010, **8**:e1000281.
46. Burki F, Kaplan M, Tikhonenkov DV, Zlatogursky V, Minh BQ, Radaykina LV, Smirnov A, Mylnikov AP, Keeling PJ: **Untangling the early diversification of eukaryotes: a phylogenomic study of the evolutionary origins of Centrohelida, Haptophyta and Cryptista**. *Proc. Biol. Sci.* 2016, **283**:20152802.
47. Field MC, Gabernet-Castello C, Dacks JB: **Reconstructing the evolution of the endocytic system: insights from genomics and molecular cell biology**.

Adv. Exp. Med. Biol. 2007, **607**:84-96.
48. Forterre P: **A new fusion hypothesis for the origin of Eukarya: better than previous ones, but probably also wrong**. *Res. Microbiol.* 2011, **162**:77-91.
49. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ: **The Phyre2 web portal for protein modeling, prediction and analysis**. *Nat. Protoc.* 2015, **10**:845-858.
50. Alva V, Nam SZ, Söding J, Lupas AN: **The MPI bioinformatics Toolkit as an integrative platform for advanced protein sequence and structure analysis**. *Nucleic Acids Res.* 2016, **44**:W410-415.

Figure Legends

Figure 1. Architecture of the three well-characterised types of vesicle coats. All three coats are recruited onto the membrane by a small GTPase, where they collect cargo and form a vesicle. However, different types of proteins contribute to cargo selection. In addition, the β-propeller plus α-solenoid components of the coat assemble in different ways, and whereas COPII and clathrin coats have an outer polyhedral layer, COPI does not.

Figure 2. The seven HTAC-containing coats plus COPII. Diagram of the endomembrane system of a eukaryotic cell and the coats that facilitate vesicle budding. The functions of the AP-4 and AP-5 coats are still unclear. Inset: Known relationships between the 7 HTACs, based on [15] and [33].

Figure 3. Evolution of vesicle coats.
A) Tree of eukaryotes with cartoon depictions of recognizable members of each major grouping showing the diversity of eukaryotes. The eight coats discussed were all present in the Last Eukaryotic Common Ancestor, as shown by the blue circle at the base of the eukaryotic tree. The blue circles also denote instances of lineage-specific innovations, while red circles show instances of AP-3 and AP-4 loss. In all cases but AP-3 in Haptophyta, the loss is incomplete, with some members of the grouping having lost the complex but others retaining it. The contribution of the robustly validated building block domains of the trafficking machinery from the Asgard archaea (Lokiarchaeota, Thorarchaeota, Odinarchaeota, Heimdallarchaeota) is also shown. The tree topology is based on a combination of molecular phylogenetic and ultrastructural data as detailed in [11] and [46]. This image is modified from [33] and [47] with permission from the authors.
B) Cartoon illustrating the possible role of convergence in the evolution of eukaryotic vesicle coats. There are two prominent options for the source of the progenitor eukaryotic protocoatomer. 1) The gene clusters found in the Thorarchaeote AB_25 and WOR_45 genomes encode a WD-40 (β-propeller domain) adjacent to a TPR (α-solenoid domain) (eg. AM324_13180 AND AM324_13175) [36]. Although the domains are unfused, the evolutionary link is based on the Asgard archaea as the clear source of other important components of the membrane trafficking machinery. 2) The only example of fused proteins containing β-propeller-α-solenoid domains in prokaryotes are present in the Planctomycetes-Verrucomicrobia-Chlamydiae (PVC) bacterial superphylum (eg. WP_010038441)[45]. Planctomycetes are known to possess internal membranes with which the fused β-α are associated. In this case either horizontal gene transfer of the protein or a cryptic fusion event between planctomycetes and an archaeal lineage would need to be invoked [48]. Under scenario 1, the planctomycete fusion must have occurred independently of the one that gave rise to the protocoatomer in eukaryotes. In this case the fact that these fused proteins are also associated with internal membrane deformations would also be indicating the biophysical necessity for fusion of the domains in order to produce the membrane-deformation result. Under scenario 2, the arrangement of the domains must have arisen independently in the Asgard archaea. The lozenges represent genes encoding the proteins whose accessions are written within. Note the lozenges are not to scale. The ribbon diagrams were produced either by structure prediction using the Phyre2 Server ([49] or using the HHPred algorithm [50] in the case of the G. obscuriglobus WP_010038441 protein sequence. For this protein the N- and C-terminal structures were predicted separately and the grey dotted line denotes the ~20AA linker between the segments that is not included in either structure. For the

AM324_13175 prediction, the model covered only 20% of the sequences, explaining why it does not show as clear a solenoid structure as the other domains.

Box 1. Structures commonly found in vesicle coat components.

β-propeller. A ~300-residue structure consisting of β strands resembling the propeller of an airplane.

α-solenoid. A structure consisting of repeating α-helices forming a zigzag, also known as a helix-turn-helix motif.

Longin domain. A ~200-residue domain with an α-β-α sandwich architecture, found not only in HTAC medium and small subunits, but also in several other proteins involved in membrane traffic, such as R-SNAREs and Rab guanine nucleotide exchange factors.

μ homology domain (MHD). A ~280-residue banana-shaped structure consisting of β strands, found at the C-terminal end of HTAC medium subunits.

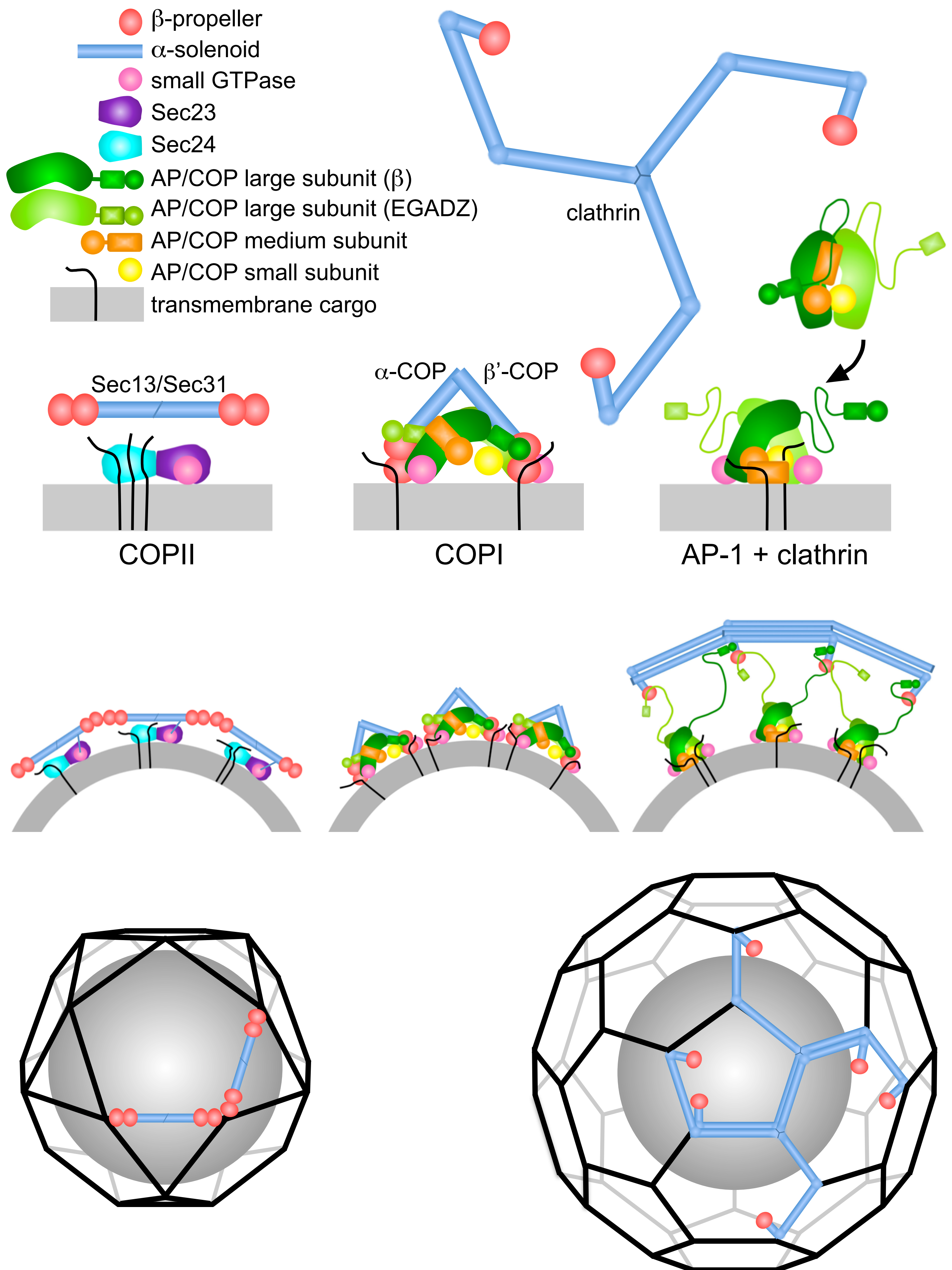

Figure 1

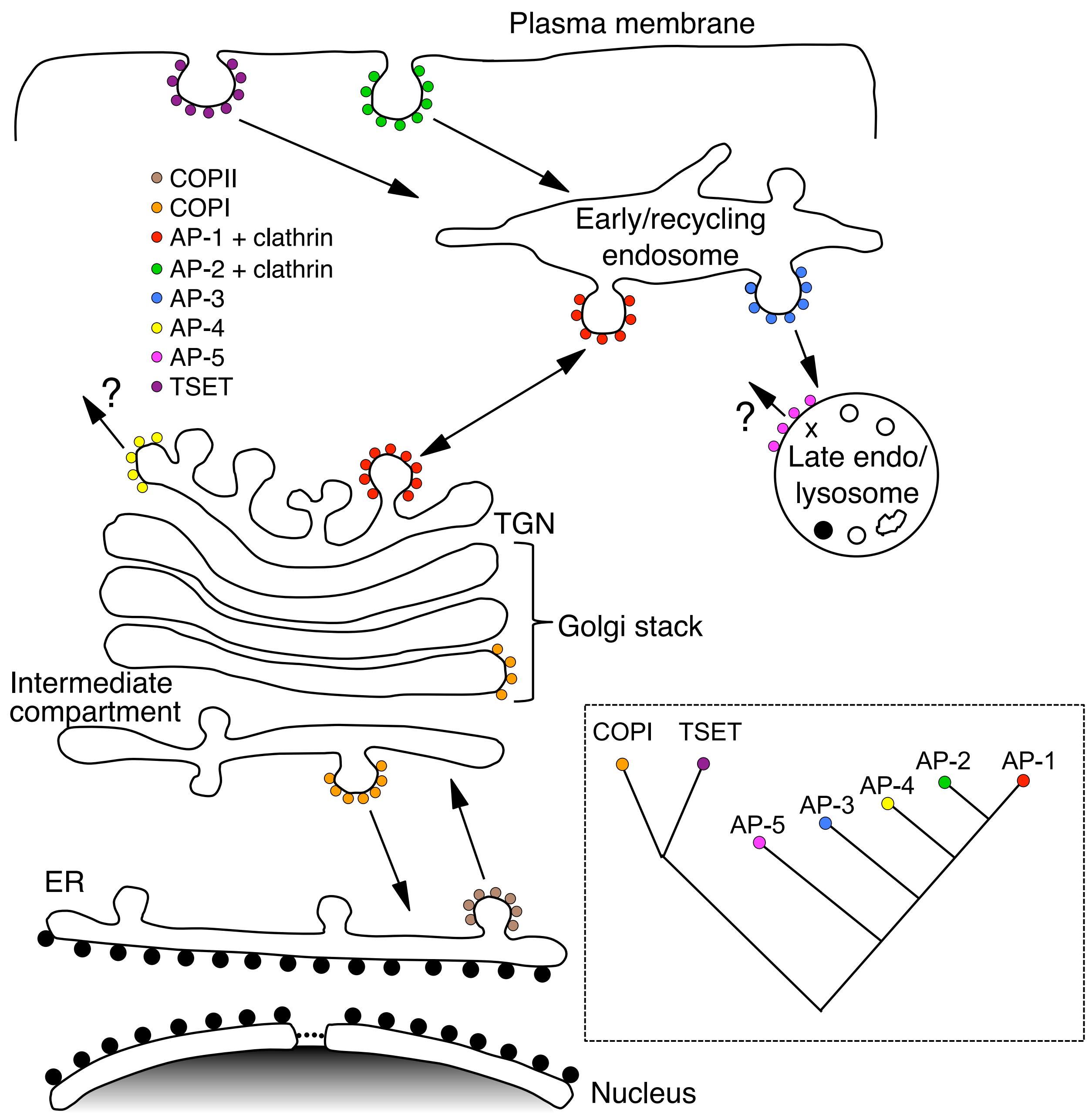

Figure 2

A

Excavata

Archaeplastida

Opisthokonta

Discoba

Metamonada

Fungi

Chlorophyta

Embryophyta

TbCAPs

AP-4

AP-3

AP-4

Stramenopila

Cryptophyta

AP-4

AP-4

Holozoa

RASH

Alveolata

Glaucophyta

Dab2

GGA

Amoebozoa

AP-3

AP-EAR

AP-3

Rhizaria

COPI AP-1
TSET AP-2
COPII AP-3
EpsinR AP-4
AP180 AP-5

Asgard archaea

Haptophyta

Sec23/24
Longin Domain
Arf/Sar-like GTPases
Ras Superfamily-like GTPases

Inferred point of origin
Inferred point of secondary loss in some members of group

B

AM324_13180 AM324_13175

Thorarchaeote WOR83

Scenario 1=Fusion and vertical inheritance in Asgard archaea: convergence in PVC Bacteria

CAA98772

S. cerevisiae-Sec31

WP_010038441

G. obscuriglobus

Scenario 2=Horizontal Gene Transfer or cryptic fusion in eukaryogenesis with PVC Bacteria: convergence in Asgard archaea