The International Journal of Biochemistry & Cell Biology 41 (2009) 330–340

Contents lists available at ScienceDirect

The International Journal of Biochemistry & Cell Biology

journal homepage: www.elsevier.com/locate/biocel

Review

Evolution of specificity in the eukaryotic endomembrane system

Joel B. Dacks[a,b,*], Andrew A. Peden[c], Mark C. Field[a]

[a] The Molteno Building, Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QP, UK
[b] Department of Cell Biology, Medical Sciences Building, University of Alberta, Edmonton, Alberta, Canada T6G 2H7
[c] Department of Clinical Biochemistry, Cambridge Institute for Medical Research, University of Cambridge, UK

ARTICLE INFO

Article history:
Available online 16 September 2008

Keywords:
SNARE
Rab
Vesicle coat
Tether
ESCRT

ABSTRACT

Two hundred years after Darwin's birth, our understanding of genetic mechanisms and cell biology has advanced to a level unimaginable in the 19th century. We now know that eukaryotic cells contain a huge variety of internal compartments, each with their own function, identity and history. For the compartments that together form the membrane-trafficking system, one of the central questions is how that identity is encoded and how it evolved. Here we review the key components involved in membrane-trafficking events, including SNAREs, Rabs, vesicle coats, and tethers and what is known about their evolutionary history. Our current understanding suggests a possible common mechanism by which the membrane-trafficking organelles might have evolved. This model of increased organellar complexity by gene duplication and co-evolution of multiple, interacting, specificity-encoding proteins could well be applicable to other non-endosymbiotic organelles as well. The application of basic evolutionary principles well beyond their original scope has been exceedingly powerful not only in reconstructing the history of cellular compartments, but for medical and applied research as well, and underlines the contributions of Darwin's ideas in modern biology.

Contents

1. Introduction

Charles Darwin's ideas on evolution by natural selection have profoundly influenced every aspect of biology. While Darwin wrote mostly on macroevolution in plants and animals, he was certainly aware that evolutionary theory could be extended to single-celled organisms. Examples of infusoria and rhizopods (protists) were included in his 1861 edition of The Origin of the Species (Darwin, 1861). The organism was still the focus, but cell theory had been formulated and Virchow had already coined his famous dictum that 'all cells come from cells' (reviewed in Mazzarello, 1999). The stage for study into the evolution of cellular compartments had been set. Over 100 years later, Stanier (1974) made clear the profound divisions between prokaryotic and eukaryotic cellular organisations. Since then, the question of how eukaryotes arose, presumably from simpler cellular ancestors, has been a central issue in biology. The basic question can be rephrased: how did eukaryotic organelles

Abbreviations: LCEA, last Common Eukaryotic Ancestor; SM, Sec1/Munc18; GAP, GTPase-activating protein; GEF, guanine exchange factor; TGN, trans-Golgi network; AP, adaptin; NPC, nuclear pore complex; MVB, multivesicular body; ENTH, epsin N-terminal homology; REP, Rab escort protein.

* Corresponding author at: Department of Cell Biology, 6-30 Medical Sciences Building, University of Alberta, Edmonton, Alberta, Canada T6G 2H7.
Tel.: +1 780 248 1493; fax: +1 780 492 0450.
E-mail address: joel.dacks@ualberta.ca (J.B. Dacks).

1357-2725/$ – see front matter.
doi:10.1016/j.biocel.2008.08.041

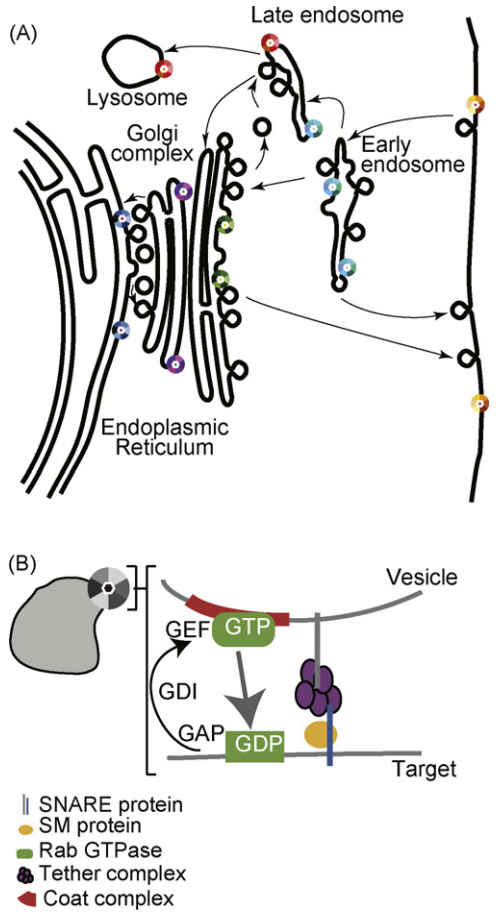

Fig. 1. Eukaryotic cell and vesicle formation/fusion machinery. (A) Simplified depiction of a generalised eukaryotic cell emphasising the main trafficking pathways and the concept that a common set of proteins are involved in vesicle formation and fusion events between the distinct intracellular compartments. The different coloured sectored 'identity' circles indicate that many of the proteins involved have specific paralogues associated with distinct organelles or transport steps and have been proposed as encoding the 'identity' of the organelle. This figure is redrawn and modified from Koumandou et al. (2007) with permission of the authors. (B) Proteins included in an 'identity circle' are vesicle coats that select cargo and allow transport carrier formation, as well as Rab proteins that initiate vesicle attachment at the target organelle and SNARE-associated machinery that drives membrane fusion. The hexagons in the center represent factors such as tethers that interact with the paralogously derived machinery. The colours of the individual shapes in panel B are not correlated to the colours in panel A.

originate and how did they subsequently evolve? Studies in the 1970s and since of endosymbiotic processes have held answers for mitochondria (van der Giezen, 2006) and plastids (Lane and Archibald, 2008); but many eukaryotic organelles more likely have autogenous origins, arising from components and structures that preexisted within the proto-eukaryote itself.

The endomembrane system is one such case. This fundamental set of cellular machinery consists of dynamic membrane-bound organelles, with each having a specific protein and lipid composition and thus a discrete function (Fig. 1). For example, the endoplasmic reticulum (ER) is the site of polypeptide entry to the secretory system and also important in lipid biosynthesis. The Golgi apparatus performs post-translational modifications on newly synthesised proteins, and sorts them for transport to other endomembrane organelles. This organelle underlies both the endocytic and exocytic pathways and is particularly important for extracellular matrix formation. The plasma membrane is the interface between the cell and environment, a site of cell signaling as well as the place where material is internalized. Finally, endosomes sort material, recycling back to the cell surface or targeting into terminal degradative organelles variously termed lysosomes, vacuoles, and reservosomes. Proteins and lipids are transported between these compartments, in part, *via* small membrane-bounded transport vesicles and tubules, which bud from one compartment and fuse with another, thereby delivering their contents and membrane (Bonifacino and Glick, 2004). The generation, targeting and fusion of these transport intermediates requires four major classes of protein (coat proteins, Rab GTPases, tethers and SNAREs). The coat proteins perform an early step in this process where they select protein cargo for transport and deform the underlying membrane into a vesicle. Rabs subsequently recruit a diverse range of effector molecules that play multiple roles in the transport, targeting and tethering of the vesicle. The tethers are a diverse class of protein complexes that promote attachment between the transport vesicle and its target organelle. Finally SNAREs drive vesicle fusion (Fig. 1) (Bonifacino and Glick, 2004). All of these factors act in a collective manner to generate specificity in membrane transport, and it remains unclear if any one is paramount in defining specificity (Cai et al., 2007a).

Both a tremendous growth in the volume of data and a new generation of informatics tools have vastly expanded the scope of viable molecular-level evolutionary studies. With an increasingly detailed knowledge of the molecular machinery underlying membrane-trafficking events, it is now possible to consider the ubiquity, or otherwise, of individual components and assess when in evolution particular pieces of the endomembrane system could have originated. While it is possible to find occasional direct prokaryotic homologues, by far the most tractable approach at present is to reconstruct the endomembrane system present in the Last Common Eukaryotic Ancestor (LCEA). This is partly achieved by searching for components of the protein machinery in genomes of organisms representing the broadest available sample of eukaryotic diversity. If a factor is found in all, or most, of the genomes, then it was most likely anciently derived and hence present in their common ancestor, a conceptual proxy for the LCEA. On the other hand if the distribution is restricted to a narrow range of organisms, e.g. yeast to mammals, then the factor most likely evolved more recently. While there are certainly caveats and limitations to such studies as discussed elsewhere (Dacks and Field, 2004, 2007), these issues can, and in many cases have been, redressed by additional phylogenetic and experimental analyses.

These reconstructions suggest that the LCEA possessed a complex repertoire of endomembrane organelles and a remarkably complete complement of protein factors for membrane-trafficking. The endoplasmic reticulum, plasma membrane, multi-vesicular body (MVB) and a stacked Golgi apparatus were all most likely present (Roger, 1999; Dacks et al., 2003; Leung et al., 2008). The degree of differentiation of the endocytic organelles remains incompletely defined, but even here some division into recycling, and degradation pathways was clearly established (Dacks et al., 2008). At the molecular level, there was also significant complexity. All of the major protein families involved in vesicle formation and vesicle fusion, as described for metazoa and fungi, appear to have been on board by the time of the LCEA (Hartman and Fedorov, 2002; Dacks and Field, 2004). The ancestral eukaryote may, therefore, have been ancient but it was not simple and exhibited a surprising level of complexity that is recognisable in modern cell systems.

The evolutionary connection between these diverse organelles and complex machinery is not immediately obvious at first inspection. More puzzling is the question of how such cellular features originated. It appears unlikely that the organelles of the endomembrane system are of direct endosymbiotic origin. Each of them lacks the major endosymbiotic hallmarks of closely adpressed

double-membranes, associated genomes, or autonomous replication. While there are examples of prokaryotic homologues for a very few proteins associated with the membrane-trafficking system (Collins et al., 2005; Obita et al., 2007; Podar et al., 2008), there are not the obvious numbers and homologies seen between mitochondria and alpha proteobacteria (van der Giezen, 2006), plastids and cyanobacteria (Lane and Archibald, 2008). Rather, this suggests that some of the factors associated with the endomembrane system simply have origins that predate eukaryogenesis, and hence are shared with prokaryotes for that reason alone.

If the endomembrane organelles are indeed derived autogenously, then what mechanisms could account for their origins? Similarity between the mechanisms of trafficking and the protein families involved in distinct transport steps hints at a common origin. We, and others, have speculated that one key to unraveling the evolution of membrane-trafficking organelles lies in the molecular machinery that participates in encoding organellar identity, regulating the specificity of sorting into distinct transport carriers and also delivery to a specific destination (Cavalier-Smith, 2002; Dacks and Doolittle, 2002; Jekely, 2003; Arac et al., 2005; Dacks, 2007). In the remainder of this article, we will examine the evolution of protein factors associated with encoding specificity and identity. We will begin with the three most prominent families, the SNAREs, the Rabs and the vesicle coats, as well as some of their associated machinery. Evolutionary analyses, primarily of these three sets of machinery have contributed to a theory on the process of autogenous organelle evolution, driven by paralogous gene family expansion and co-evolution. This theory has even been bolstered by an example of the process caught mid-stream but is here enhanced to incorporate the evolution of machinery such as the tethering complexes that act to cement the differences between the newly evolving organelles. This theory is best encapsulated in the evolution of the ubiquitylation machinery which shows both paralogy-driven expansion and accretion of non-homologous components. We will conclude with a few thoughts on the future of evolutionary studies of the membrane-trafficking system and the application of this theory to other autogenous organelles.

2. SNAREs and Sec1/Munc18 proteins

SNAREs (soluble N-ethylmaleimide sensitive factor attachment protein receptors) were first identified in the early 1980s and were shown to be required for vesicle fusion by a combination of genetic and biochemical reconstitution experiments (Novick et al., 1980; Sollner et al., 1993). A model for SNARE function, termed the SNARE hypothesis, initially proposed that a transport vesicle-located SNARE (v-SNARE or VAMP) and a second SNARE on the target membrane (t-SNARE or syntaxin) would interact to form a stable complex. This *trans*-membrane complex would potentially bring the two membranes in close enough proximity to drive lipid bilayer fusion. In addition, it was proposed that for each transport step within the cell there would be a specific set of SNAREs. Hence these molecules and their selective interactions would provide specificity to membrane transport (Sollner et al., 1993). On the whole, this model has proven to be remarkably insightful although it is now apparent that SNAREs cannot be the sole determinants of specificity in vesicle transport (Fig. 1).

The majority of SNARE proteins are small (14–40 kDa) type-2 transmembrane polypeptides that contain 60–70 highly conserved amino acids, arranged into heptad repeats (the SNARE motif) that are required for SNARE complex formation. Based on this shared homologous motif, all SNAREs are presumed to be derived from a single ancestral protein by gene duplications giving rise to the SNARE families and then subfamilies. The timing and order of these gene duplications are obviously central to understanding the evolution of the SNARE proteins. The crystal structure of several intracellular SNARE complexes revealed a parallel four-helical coiled-coil bundle with each SNARE contributing one coil (Antonin et al., 2002; Zwilling et al., 2007), confirming the results of earlier studies on the neuronal SNARE complex required for the fusion of synaptic vesicles at the plasma membrane (Sutton et al., 1998). The majority of side-chain interactions between the coiled-coils are hydrophobic, but at the centre of the coiled-coil is a hydrophilic layer termed the '0' layer (ionic layer) containing three glutamines (Q) and one arginine (R). At present it is still unclear what role the '0' layer plays in SNARE function, but mutation of '0' layer residues in the yeast SNARE complex required for exocytosis, causes defects to secretion and cell viability (Ossig et al., 2000). However, a separate study found that secretion still occurred if a mutant SNARE complex containing four glutamines was created (Katz and Brennwald, 2000). Regardless, the observation that nearly all SNAREs contain either a glutamine or arginine residue at the '0' layer has allowed reclassification of SNARE proteins, based on sequence rather than location or inferred function, as R or Q (Fasshauer et al., 1998), with further sub-division of the Q-SNAREs as Qa (or syntaxins), Qb and Qc-SNAREs. The four-helix bundle of the SNARE complex is almost always composed of one helix of each type (Fasshauer et al., 1998).

Comparative genomic studies have established the presence of homologues from all four SNARE families in a broad diversity of eukaryotes (Dacks and Doolittle, 2001; Dacks and Field, 2004; Yoshizawa et al., 2006). Several molecular cell biological studies in eukaryotes from the different supergroups also confirm the universal involvement of SNAREs in cellular trafficking (Sanderfoot et al., 2001; Bogdanovic et al., 2002; Besteiro et al., 2006; Schilde et al., 2006; Ayong et al., 2007). So while not explicitly aimed at ancestral reconstruction, this all strongly points to the duplication giving rise to the Qa, b, c and R-SNARE families having occurred prior to the ancestor of existing eukaryotes (Fig. 2A).

The human genome encodes at least 38 SNARE genes, with each SNARE gene product being localised to a defined set of membranes and hence suggested to be involved in a specific set of fusion reactions (Bock et al., 2001). For example, syntaxin 17 localises to the ER (Steegmaier et al., 2000) and functions in retrieval of escaped ER proteins, syntaxin 4 is localised to the plasma membrane and involved in the fusion of GLUT4-containing vesicles with the plasma membrane (Volchuk et al., 1996). Syntaxin 7 locates to late endosomes and is required for late endosome/lysosome fusion (Mullock et al., 2000). Studies of Qa orthologues in diverse organisms reveal largely similar intracellular localisation (Bogdanovic et al., 2000; Besteiro et al., 2006; Ayong et al., 2007; Kissmehl et al., 2007) and suggest the existence of five major organelle, or transport pathway specific, syntaxin subfamilies (Dacks, 2007). The anterograde endocytic syntaxin subfamily (SynE) contains all of the homologues related to syntaxin 7 and 12, while the plasma membrane syntaxins (SynPM) are all related to syntaxin 4. The remaining subfamilies; Syn5, Syn18 and Syn16 are involved in transport at the *cis*-Golgi, ER and retrograde transport between endosomes and *trans*-Golgi network respectively (Fig. 2A). Phylogenetic studies suggest that the five major syntaxin families likely predate the LCEA (Dacks and Doolittle, 2002, 2004). Further, analyses robustly resolve the ER-associated Syn18 subfamily with the Golgi-associated Syn5 subfamily in evolutionary reconstructions, suggesting that these organelles share an evolutionary connection to the exclusion of the other endomembrane compartments (Dacks, 2007). There is also clear evidence for lineage specific expansion of several SNARE protein subfamilies. Most notably, expansion in metazoa and higher plants of the SNARE proteins involved with late stages of exocytosis may suggest a correlation with the evolution of multicellularity, and hence potential tissue-specific functions (Dacks and Doolittle, 2002; Yoshizawa et al., 2006; Sanderfoot, 2007).

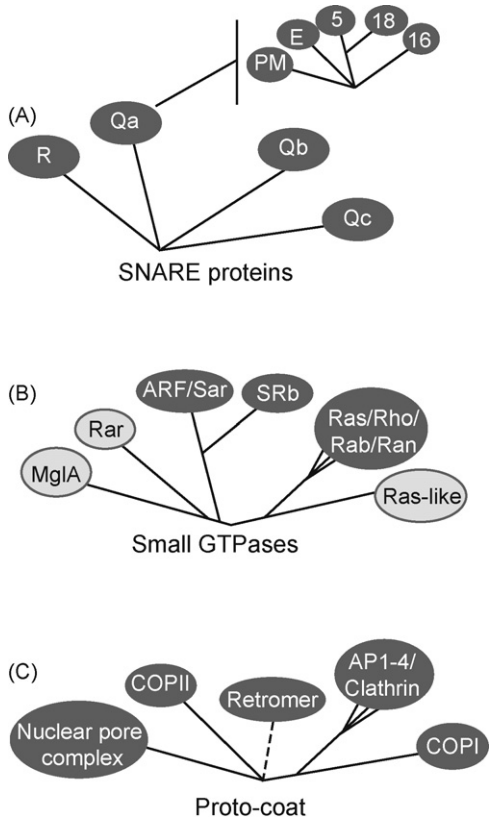

Fig. 2. Evolution of specificity-encoding families or complexes. Ovals denote protein families or complexes; dark grey for eukaryotes and light grey for prokaryotes. Solid black lines denote relationships for which there is good evidence either from homology searching, shared components or molecular phylogenetics. The dashed line denotes a potential evolutionary connection but lacking robust evidence for support. (A) SNAREs. There are four SNARE families that each appears to have evolved before the LCEA. The order in which they evolved is unclear. Inset: The five major subfamilies of the Qa SNAREs, or syntaxins, also appear to have evolved before the LCEA, with evidence suggesting that the Golgi-associated Syn 5 and the ER-associated Syn 18 subfamilies share a closer relationship to one another than to other syntaxins. (B) Small GTPases. The major eukaryotic small GTPases and two prokaryotic classes are shown. Significantly, Sar1 and ARF are clearly close relatives of one another, as are Ras, Ran, Rho and Rab. Surprisingly, the eukaryotic GTPases appear to have two distinct prokaryotic origins. (C) The membrane deformation complexes. Depicted is the relationship between heterogenous protein complexes, and focuses on considering only those subunits for which there is evidence for a common ancestor and structure. The relationship between the three vesicle coats, COP I, COP II and clathrin–adaptin is based on a variety of evidence and is well established. Structural considerations also suggest that some components of the nuclear pore complex scaffold are related to this group. Finally, retromer may also be part of this lineage, although this last association is currently tenuous.

Many Q-SNAREs also possess a short N-terminal peptide (N-peptide) which has a role in binding members of a family of molecules known as Sec1/Munc18 (SM) proteins. First identified in a genetic screen for secretion (Novick et al., 1980), SM proteins are essential for vesicle fusion, but the precise function of the SM proteins remains incompletely defined. SM proteins may inhibit SNARE complex assembly by binding a Qa-SNARE in the closed conformation as, for example, Munc18-1/syntaxin 1 (Misura et al., 2000). But SM proteins also stimulate SNARE complex assembly, for example Sly1/syntaxin 5 and Vps45/syntaxin16 (Toonen and Verhage, 2003). Finally, SM proteins may also proof-read SNARE assembly and prevent formation of inappropriate complexes, such as 2Q:2R (Starai et al., 2008).

Seven SM proteins are encoded in the human genome: Munc18-1, Munc18b (Munc18-2), Munc18c, Sly1, Vps45 and Vps33A and B (Toonen and Verhage, 2003). Homologues are present across the eukaryotes and form four major families Sly1, Vps45, Sec1 (which includes Munc18) and Vps33. Phylogenetic analyses suggest that the four major SM protein subfamilies had already evolved before the LCEA (Arac et al., 2005; Koumandou et al., 2007).

SM proteins were initially thought to bind syntaxins in, at least, two distinct modes. Sly1 and Vps45 bound their cognate SNAREs syntaxin 5 and syntaxin 16 in the open conformation (Bracher and Weissenhorn, 2002; Dulubova et al., 2002), while Sec1/Munc18 bound syntaxin 1 in a closed conformation (Misura et al., 2000). This insight provoked the question as to how homologous and structurally related proteins could have evolved distinct binding and regulation modes (Toonen and Verhage, 2003). Recent evidence has resolved this issue. The SNARE N-peptide binds a conserved hydrophobic pocket on the SM protein (N-pocket) (Arac et al., 2005). Munc18-1, Munc18b, Munc18c, Sly1 and Vps45 all retain the N-pocket and their corresponding partners syntaxin 1, 4, 5 and 16, respectively have an N-peptide that can bind this pocket (Hu et al., 2007; Burkhardt et al., 2008). This implies that, despite earlier issues about SM proteins binding SNAREs in a closed versus an open conformation, there may well be a shared common binding mode. However, this binding mode may not be conserved between all SM proteins. For example, human Vps33a lacks an N-pocket and yeast Sec1p/Sso1p and Vps33p/Vam3p lack both the N-pocket and N-peptide (Hu et al., 2007). Nevertheless, these SM proteins still interact with SNAREs. Additional investigation, particularly of homologous proteins from evolutionarily distant species may well help to unravel this issue. Interestingly, Vps33p comprises part of the larger multiprotein HOPS complex (homotypic vacuole fusion and vacuole protein sorting) (Seals et al., 2000), which interacts with the small GTPase Ypt7p and plays an important role in tethering and regulating late endocytic fusion events. Thus, the indirect binding of the syntaxins is likely a derived feature of this SM protein following its co-option into the HOPS tethering complex (Koumandou et al., 2007).

The SNARE hypothesis originally postulated that the combinatorial interaction of V and T SNAREs fully encoded the specificity of a fusion reaction and implicitly also encoded the identity of the target organelle. By this view, the evolution of SNAREs recapitulates the evolution of the trafficking system. Most certainly SNAREs play an important part, but increasing evidence indicates that other molecular players are involved.

3. Rab family small GTPases

The Rab family is the largest member of the Ras superfamily of small GTPases. Their activities appear to be predominantly restricted to control of vesicle transport, *albeit* with a rather great range of function at the molecular level (Zerial and McBride, 2001). There is some evidence that a minority play roles in flagellum biosynthesis (Oro, 2007). Since the description of the first member of the Rab subfamily, Sec4p in *Saccharomyces cerevisiae*, many hundreds of Rabs have been described from multiple genomes. They are a constant presence in eukaryotes, with all organisms possessing at least some Rab proteins. Interest in the Rab family was initially spurred by their unique localisations, for example Rab1 at the ER, Rab5 in early endosomes and Rab7 at the lysosome/vacuole (Zerial and McBride, 2001). In common with other small GTPases, Rab protein structure comprises a compact six-strand beta sheet, together with five alpha helices. A flexible C-terminal hypervariable domain, together with a terminal prenylation signal completes the Rab protein structure. Interestingly, the central core region appears to be highly flexible, in that independent evolution of surface loops between conserved secondary structural elements is a common feature of small GTPases, facilitating an extreme level of flexibility while retaining a conserved core (Blouin et al., 2004).

An early study of Rab evolution suggested that the family was ancient, and that the basic organisation of the Rabs was likely arrived at by the time of the LCEA (Iwabe et al., 1996). A more extensive and systematic analysis (Pereira-Leal and Seabra, 2000) identified multiple sequence elements conserved between all members, but which discriminate Rabs from other Ras-like GTPases. Further, additional subfamily-specific sequences were recognised allowing eight Rab clades to be defined (Pereira-Leal and Seabra, 2001). For the most part, there is some function–phylogeny correlation, i.e. Rabs involved in post-Golgi processes tended to occupy the same clade. This analysis has provided an important framework from which to build and test new models. A more detailed investigation of Rabs, encompassing more diverse eukaryotic genomes and a large sampling of fungal genomes (Pereira-Leal, 2008), suggests that the LCEA possessed at least Rabs 1, 2, 4, 5, 6, 7, 8, 11 and 18, consistent with both syntaxin reconstructions and a suspected high level of complexity in the LCEA. The modern Rab complement in many fungi appears to have also resulted from secondary loss (Pereira-Leal, 2008). Remarkably, in spite of a great range in genome size, the Rab family is maintained at 8–12 members in most fungi species. Over a broader range of genomes, Rab4 is a frequent victim of secondary loss, representing an example of a trafficking pathway that was probably essential to the LCEA, but became dispensable in many extant lineages (Field et al., 2007).

While the situation in fungi appears to be one of complement maintenance or reduction, significant expansion and emergence of novel subfamilies of Rabs has been observed. Specifically in *Entamoeba*, over 90 members are known, and *Trichomonas* has nearly 300 (Lal et al., 2005; Saito-Nakano et al., 2005; Carlton et al., 2007). In each organism, over half of the Rab sequences do not appear to belong to a clade shared with other taxa. The level of sequence divergence within these lineage-specific clades is considerable, and certainly consistent with novel specificity and binding activity. The driver for such massive expansion is unknown, and will require additional genome sequencing to determine if these observations are organism-specific, life-style dependent or due to other factors. More limited lineage-specific expansion of conserved Rab subfamilies is rather frequent and has clearly shaped Rab evolutionary history. In *Arabidopsis thaliana*, Rab sequences can be classified into eight distinct clades. Three of these contain orthologues of Rab11, Rab7, as well as Rabs associated with polarised secretion, and exhibit considerable expansion; the Rab11 clade is particularly large (Rutherford and Moore, 2002; Vernoud et al., 2003). These clades are all associated with endosomal activity or specialised exocytotic events, both of which are clear features of *A. thaliana* biology, and hence there is some suggestion for functional consequences. Further, sequence divergence between members of these clades is considerable, consistent with functional differentiation (Rutherford and Moore, 2002).

Until recently, the Ras superfamily was considered to be essentially a eukaryote-restricted gene family. A rather divergent GTPase family, MglA, is widely distributed amongst eubacteria, and has some similarities to Ras, but divergence is significant enough to suggest distinct function and mechanism. However, recent analysis has uncovered a rather more conserved Ras-like family in both eubacterial and archaeal genomes. These Ras-like proteins retain the five defining G-motifs of the Ras superfamily and are hence are more closely related to Ras than is MglA (Dong et al., 2007). It is not clear at this time how extensive this new Ras-like family is amongst prokaryotes, but the evidence so far is that the representation is quite broad, and likely precludes lateral gene transfer as an origin. Further, phylogenetic reconstruction of the prokaryote and eukaryote GTPases suggests distinct origins for the Ras/Rho/Rab/Ran and ARF/Sar family GTPases (Fig. 2B). This new insight augers well for efforts to extend reconstruction of trafficking pathways and understanding the evolution of specificity beyond the prokaryote–eukaryote boundary.

Rabs exist in two major states, GTP or GDP bound. Cycling between GTP and GDP states provides a molecular switch; the overall protein fold is rather well conserved between the family members, but significantly there is a restricted and rather clear conformational change accompanying hydrolysis of GTP to GDP. This is mainly confined to the switch I and switch II regions. Importantly, these regions contribute the binding sites for multiple factors, and such interactions are therefore nucleotide-state dependent. The slow intrinsic rates of hydrolysis and nucleotide exchange are accelerated by GTPase-activating proteins (GAPs) and guanine exchange factors (GEFs) respectively. Conformational changes and recognition of the Rab switch regions facilitates specific recruitment of GAPs to the GTP-bound form and GEFs to the GDP form. However, given the very large number of factors that recognise specific Rabs, for example Rab5 has ~30 identified binding partners, the contributions of additional sequences to binding sites must play some role in extending specificity (Pfeffer and Aivazian, 2004). Understanding the co-evolution of binding sites for factors that constitute complexes is a major challenge.

While the cellular location of each Rab is highly specific and apparently conserved across large evolutionary distance, the sequence determinants encoding this specificity are unclear. Studies suggested that the Rab hypervariable C-terminal domain provided targeting information (Pfeffer and Aivazian, 2004), but this has been challenged by more detailed work (Steele-Mortimer et al., 1994; Ali et al., 2004). Following synthesis, Rabs are prenylated in the cytosol by the Rab geranylgeranyltransferase (GGTase), and delivered to membranes by Rab escort protein (REP). Interestingly, Rab GGTase and REP are ancient, and likely LCEA components (Rasteiro and Pereira-Leal, 2007). REP, which only facilitates solubilisation and membrane delivery of newly synthesised Rabs, shares a common evolutionary origin with guanine dissociation inhibitor (GDI), a third GTP cycle factor that is able to bind GDP-bound Rabs and extract them from the membrane (Alory and Balch, 2003; Hala et al., 2005). Hence there are some clear functional similarities between REP and GDI; this extends to their broad specificity, as there is only one REP in most genomes and a small number of GDI genes. REP and GDI therefore represent ancient and general aspects of Rab function, but likely contribute little to specificity or the evolution of novel compartments or pathways.

Further, while the molecular identity of the membrane Rab receptor is unknown, the Yip family of membrane proteins from yeast was offered recently as candidates. These proteins recruit Rabs to the membrane by displacement of RabGDI, which normally maintains them in soluble form. However Yip–Rab binding specificity appears rather low and the number of Yips represented in the genome is substantially smaller than the Rab repertoire. For example, there are 14 Yip homologues in *H. sapiens* compared to ~70 Rabs (reviewed in (Pfeffer and Aivazian, 2004)). Still, while clearly unable to provide complete discrimination, it is possible that the Yips broadly define membrane subpopulations, and that subsequent association with other factors targets Rabs to highly specific locations. Candidate Yip orthologues are present in many genomes, and while functionally uncharacterised, may provide a route to understanding the evolution of targeting and localisation. The concept of a membrane microdomain, originally put forward to explain the localisation of Rab5 with Rab4 and Rab11 on contiguous endosomal membranes, is consistent with the idea of a requirement for a self-assembling complex providing the final level of specificity (de Renzis et al., 2002; Rink et al., 2005). This, however, represents a significant challenge for *in silico* pathway reconstruction.

4. Vesicle forming machinery

While much of the focus on identifying molecules encoding organelle identity and specificity has been on components of vesicle fusion, the proteins involved in forming the vesicle coats also clearly play a role. Like the SNARE and Rab machinery, the coats each appear to conform to a broadly generalisable model of action (Gurkan et al., 2006) and yet each are specific and characteristic of particular locations or transport pathways. COP II defines transport from the ER to the Golgi apparatus, while COP I occupies the retrograde route as well as *intra*-Golgi transport. The clathrin/adaptin coats are involved in movement of material between the *trans*-Golgi network (TGN), endocytic organelles and the cell surface (Bonifacino and Glick, 2004).

An additional complex, retromer, is also responsible for movement of material from the endosomes back to the TGN but is rather more exclusive in its cargo, recycling vacuolar hydrolase and mannose-6-phosphate receptors (Seaman, 2005). Like most coat complexes, retromer is composed of a cargo adaptor subcomplex (Vps26, Vps29, Vps35) and an outer membrane-deformation complex, in both mammals and yeast this subcomplex consists of two proteins, sorting nexins 1 and 2 and Vps5, Vps17 respectively (Hierro et al., 2007; Bonifacino and Hurley, 2008). While best described experimentally in these latter two systems, retromer genes have been characterised in eukaryotes from multiple supergroups (Dacks et al., 2003; Nakada-Tsukui et al., 2005; Damen et al., 2006) and identified in many additional eukaryotic genomes indicating a presence early in eukaryote evolution. Vps29 is also directly homologous to prokaryotic metallophosphoesterases, providing a link beyond the LCEA (Collins et al., 2005; Wang et al., 2005).

The remaining vesicle coats have been examined in much greater detail. COP II creates vesicles for transport from the ER to the Golgi body (Barlowe et al., 1994). The ER membrane-anchored protein Sec12 converts the GTPase Sar1 to its GTP-bound form, which then attaches to the site of vesicle assembly at the ER membrane. The cytosolic Sec23–Sec24 complex is then recruited by Sar1. Sec24 acts as a cargo adaptor, while Sec23 acts as a Sar1-GAP. Interestingly, recent evidence suggests that the R-SNARE Sec22 may well be incorporated into ER vesicles via recruitment through Sec23 and Sec24, the first example suggesting a dual role for the COP II GAP (Mancias and Goldberg, 2007). Membrane-deformation is performed by the Sec13/Sec31 complex that assembles into a novel geometry allowing for considerable flexibility in the dimensions of the transport vesicles (Gurkan et al., 2006). COP II machinery has been identified in complete genome sequences from animals to excavates and comparative genomic surveys have confirmed that this vesicle coat system was present in the LCEA (Dacks and Doolittle, 2001; Dacks and Field, 2004).

The COP I, or coatomer, complex is clearly involved in transport from the Golgi body back to the ER and is controversially implicated in anterograde *intra*-Golgi transport (Rabouille and Klumperman, 2005). A small GTPase, ARF, is activated by Sec7-GEFs and recruited to Golgi apparatus membranes via interaction with members of the p24 protein family (Bethune et al., 2006). This assembly creates a 'priming complex' to which the coatomer complex binds. COP I is formally composed of seven proteins. Within the F-Cop subcomplex are two large subunits CopB (Beta) and CopG (Gamma), a CopD medium subunit (Delta) and a small CopZ subunit (Zeta). Three additional factors, CopE (Epsilon), CopA (Alpha), and CopB′ (Beta prime), complete the complex, with the latter two possessing WD repeats with beta-propeller and alpha solenoid structures (Devos et al., 2004). Cargo recognition likely involves several of the coatomer subunits (CopA, CopB′, CopG, CopD) that interact with conserved KKXX motifs in the cytoplasmic tails of the cargo proteins (Bethune et al., 2006). As a small GTPase, ARF1 is a ubiquitous eukaryotic protein (Fig. 2B). Comparative genomic surveys have identified COPI components from all eukaryotic supergroups (Dacks and Doolittle, 2001; Dacks and Field, 2004). The role of COP I has been experimentally verified in many organisms and it is clear that COPI is an ancient coat system (Faulstich et al., 1996; Maier et al., 2001; Donohoe et al., 2007; Smith et al., 2007).

Beyond the Golgi complex, cargo is transported to a variety of locations. ARF again performs a central role (Bonifacino and Glick, 2004). Cargo proteins with the appropriate conserved amino acid sorting signals are bound by one of the adaptin (AP) complexes. These hetero-tetrameric complexes are composed of two large subunits, a medium and a small subunit (Robinson, 2004). Proteins from the TGN are packaged into AP1 carriers for transport to the endosomes, or into vesicles with the AP4 cargo adaptor, destined for the plasma membrane (Simmen et al., 2002; Robinson, 2004). The AP3 adaptor complex (in mammalian cells) is localised primarily to endosomes and plays a role in the delivery of lysosomal proteins (Theos et al., 2005). Adaptins, specifically AP2, also act at the cell surface, selecting cargo into clathrin-coated pits for entry into the endocytic system. While the vesicle coat for AP1, AP2 and AP3 vesicles is clathrin in mammalian cells, the coat for AP4 vesicles is not as clear (Robinson, 2004). Based on comparative genomic (Dacks and Field, 2004; Field et al., 2007), phylogenetic (Dacks et al., 2008) and cell biological evidence (Morgan et al., 2001; Lefkir et al., 2003; Happel et al., 2004; Elde et al., 2005), it is clear that the various adaptins and clathrin were present in the last common ancestor.

Although not a separate coat complex in its own right, an set of important cargo adaptors are the GGAs (Golgi-localising, Gamma-ear containing, ARF-binding proteins) (Robinson, 2004). These molecules have homology to the Gamma adaptin 1 subunit and act in transport between the TGN and endosomes. Unlike the four major adaptin complexes, the taxonomic distribution of the GGAs is limited to the opisthokonts, implying that these molecules are a recent, lineage-specific, innovation (Boehm and Bonifacino, 2001; Field et al., 2007). While the history of the adaptins is dominated by early complexity and some lineage-specific innovation, secondary loss has also shaped this system. AP4 is the most commonly lost complex, with evidence for independent loss in fungi, some animals, *Leishmania* and *Giardia*. Further examples of AP2 and AP3 loss are also present (Boehm and Bonifacino, 2001; Field et al., 2007).

The best evidence for homology between the vesicle coats (Fig. 2C) comes from higher-order structural analysis. Several components of COP I, COP II and the clathrin/AP coats share a common but characteristic set of protein folds restricted to beta propellers and alpha solenoids, suggesting that they may have a common origin (Devos et al., 2004). A link with the retromer complex is unclear, although Vps35 is predicted to be an alpha solenoid protein (Hierro et al., 2007). Retromer also shares with the other three coat systems a basic underlying mechanism of distinct cargo adaptor and membrane deformation complexes. Regardless of the relationship with retromer, additional evidence for homology of the three major coats is found in shared machinery and sequence relationships. While COP II involves Sar1 as the GTPase responsible for nucleating vesicle formation, both COP II and the adaptins use ARF (Bonifacino and Glick, 2004). Inclusion of the same GTPase is evidence for shared ancestry of the latter complexes, while Sar1 is the closest relative to ARF amongst the Ras GTPase superfamily (Dong et al., 2007) (Fig. 2B). The F-COP subcomplex and the adaptins are also homologous with both being composed of two large subunits, a medium and a small subunit (Duden et al., 1991). Homology searching and phylogenetic analysis has demonstrated that each subunit is the result of a series of gene duplications giving rise first to the COPI component and then to the four adaptin paralogues (Schledzewski

et al., 1999; Boehm and Bonifacino, 2001). Further, the two large subunits are homologous, as are the medium and small subunit and with one exception, all of these adaptin subunit duplications likely predate the LCEA. The one exception, however, provides potential insight into the processes that give rise to the organelles and the specificity-encoding machinery.

5. Paralogy as a driving force in organelle complexity

The molecular machinery described above shares some common features. There are organelle or transport pathway specific versions of each complex and/or protein family. These different versions are the products of gene duplications that in many cases predate the LCEA. And, surprisingly, the LCEA itself appears to have possessed a very full complement of both molecular machinery and membrane-trafficking organelles. Brought together, this evidence suggests a process by which the membrane-trafficking system has risen in complexity to the inferred state in the LCEA and also in modern eukaryotes (Fig. 3). The manner in which organelle specificity is encoded has been hotly debated and various authors (Cavalier-Smith, 2002; Dacks and Doolittle, 2002; Jekely, 2003; Arac et al., 2005; Yoshizawa et al., 2006; Dacks, 2007; Sanderfoot, 2007) have suggested one protein family or another as the key determinants, whereby the evolution of these proteins by gene duplication would have been concurrent with, or causal to, the evolution of the endomembrane organelles (Fig. 3B). However, it has now become clear that there are multiple determinants of organelle identity rather than a single protein family using a mailing address and house number type of system (see Cai et al., 2007a). It follows that gene duplications in each of these protein families, and most significantly, co-evolution of their interacting factors would have enabled evolutionary differentiation from a single endomembraneous organelle into an initial basic functional division and then subsequent diversification into specialised endomembrane organelles (Fig. 3). This model may extend to organelles other than within the membrane-trafficking system; specifically a topological and evolutionary connection between the ER and peroxisomes has been proposed (Hoepfner et al., 2005; Schluter et al., 2006). Even more convincingly, the only other occurrence of the unique combination of alpha solenoid and beta propeller observed in the vesicle coat proteins is within the nuclear pore complex (NPC), a structure that is also involved in membrane deformation (Devos et al., 2004). Supporting this hypothesis is the dual function of Sec13, which forms part of both the NPC and COP II coat.

A recent study lends further credence to this model. Phylogenetic analyses of the syntaxins involved in anterograde endocytic transport suggest that the LCEA possessed only a single subfamily member (Dacks et al., 2008) whereas animals, plants and fungi each retain multiple paralogues that differentially act at the recycling and degradative organelles (Collins et al., 2002; Pelham, 2002; Rojo et al., 2003). Those same organelles are also serviced by distinct Rab proteins (Rab5 and Rab7). These clearly evolved before the LCEA (Pereira-Leal and Seabra, 2001). However, in several eukaryotic lineages, multiple paralogues of Rab5 exist in various systems that have subtly distinct functions. These Rab5 genes are themselves lineage-specific, i.e. the product of gene duplications that occurred after the LCEA (Dacks et al., 2008). Finally, the gene duplications that gave rise to the F-COP and adaptin subunits all appear to pre-date the LCEA, with the exception of the Beta (B) subunit. Here the duplications giving rise to COPB, AP3B and AP4B predate the LCEA, but several lineages apparently possess only a single Beta subunit functioning in both the AP1 and AP2 complexes (Boehm and Bonifacino, 2001; Dacks et al., 2008). Phylogenetic analysis indicates this to be the ancestral state and organisms possessing separate AP1 and AP2 subunits have derived these via independent

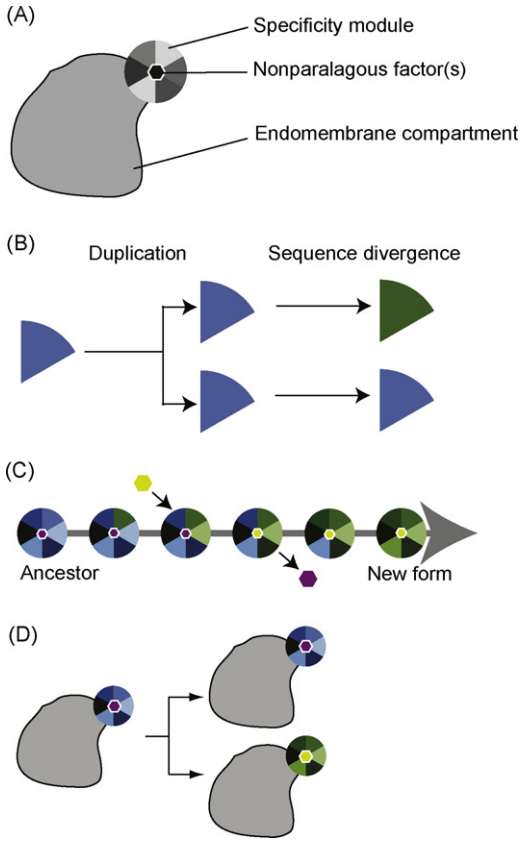

Fig. 3. Organelle evolution driven by gene duplication of the identity-encoding machinery. (A) An initial endomembraneous compartment is shown, with an as-yet undifferentiated set of identity-encoding machinery shown. The segmented circle indicates a group of subunits that are part of extensive paralagous families (Rabs, SNAREs etcetera), while the central hexagon is a non-paralagous factor (e.g. tethering complexes). (B) Gene duplication and sequence divergence of individual components of the identity-encoding machinery would produce new members of these protein families that could potentially associate with new organelles. (C) The various protein factors within the identity-encoding machinery would undergo gene duplication and co-evolution, as part of a gradual process and with replacement of different components not occurring in a synchronous manner. The process would eventually create a novel identity-encoding machine that would control trafficking for a novel transport step. This new identity would be reinforced by the inability to interact with accessory factors (yellow hexagon) and acquisition of novel factors (purple hexagon). (D) Gradual subunit replacement of an identity-encoding machine would produce new endomembranous organelles, with several iterations giving rise to the observed complexity of organelles in the membrane-trafficking system, as well as possibly other non-endosymbiotically derived compartments. The concept of the identity-encoding machinery is virtual—a single complex does not appear to encode specificity, as discussed in the text.

gene duplications after they diverged from the LCEA (Dacks et al., 2008). These data are all consistent with a model where the endocytic compartment began differentiating into discrete recycling and degradative organelles, with specific associated Rabs, by the time of the LCEA, but the gene duplications giving rise to the remaining components of the specificity-encoding machinery was not yet complete. These then occurred in parallel and in multiple lineages. Such a scenario is precisely what would be expected for any system in the midst of differentiation at the point of radiation following evolution of the LCEA, and assuming a model of paralogy-driven organellogenesis (Dacks et al., 2008).

6. Tethers

While these data can explain the origin of a large proportion of the specificity-encoding machinery, they cannot explain the origins

of all components. The tethering complexes are a heterogeneous assembly of multi-subunit proteins mediating initial recognition and attachment of an incoming vesicle to a target membrane (Sztul and Lupashin, 2006). The tether complexes vary in composition and structure, and each complex mediates not only transport to specific organelles, but also in different orientations, i.e. anteriograde versus retrograde transport. DSL1 is involved in anteriograde ER to Golgi body transport, while TRAPPI mediates transport between the same organelles, but in the opposite direction. COG and TRAPPII mediate retrograde Golgi traffic, with TRAPPII also being implicated in endosome to TGN vesicle transport. GARP mediates a similar step as well, which may reflect additional differentiation of endosomal trafficking pathways and possible endosomal populations. The CORVET complex mediates TGN to endosome transport. HOPS is involved in endosome to lysosome and homotypic lysosomal fusion; while exocyst is involved in transport from the TGN to the plasma membrane and secretion (Cai et al., 2007a). Comparative genomics confirms that the majority of these complexes are represented in sufficiently diverse eukaryotes to suggest an origin prior to the LCEA (Koumandou et al., 2007). Some subunits, and even entire complexes cannot be reliably identified in certain taxa; it is unclear if this is due to extreme sequence divergence or secondary loss in those organisms. The origins of the CORVET complex are not known at this time.

On the face of it, paralogy is at play in this system as well. Intra-complex homology was observed between members of several of the tethering complexes (Koumandou et al., 2007). For example COG1, 2, 4, 5, 7, and 8 all appear to be related, as do HOPS subunits Vps11/Vps18/Vps39 and Vps41. The most intriguing example is in the TRAPPI complex where three of the seven members are all Bet3 family members and three are Bet5 members, suggesting two sets of gene duplications of two factors explaining the derivation of the complex (Koumandou et al., 2007). The origin may even be traced back to a prokaryotic Bet3 homologue (Podar et al., 2008).

However, unlike the other examples discussed so far, there is only equivocal evidence suggesting that the tethers are derived from a single ancestral complex. While homology searching algorithms do find weak evidence for relationships between some subunits of the GARP, DSL1, COG and exocyst complexes, the values are bordering on insignificance, and the regions of homology are confined to small coiled-coil forming stretches (Whyte and Munro, 2001; Koumandou et al., 2007). No evidence for homology is found for HOPS or the TRAPP complexes either with each other or with the additional tether complexes. The heterogeneous composition of the tethers also points towards an independent origin for the various complexes. This could potentially act to reinforce the differentiation of distinct organelles and solidify the burgeoning organelle identity (Fig. 3).

7. The late endosomal system

Perhaps the set of membrane-trafficking machinery that best exemplifies this balance of paralogy-driven complexity and cementing of different organelle identity is the ESCRT system. In yeast, metazoa, and most other eukaryotes, endocytosis is a complex process involving multiple pathways. Some of these, for example the caveolin pathway, are comparatively recent lineage-restricted innovations, while others, like clathrin-dependent mechanisms, are shared across nearly all the eukaryota and are ancient features (Field et al., 2007). However, in all cases endocytic cargo molecules must be recognised, sorted, and sent onto their next destination. Covalent attachment of ubiquitin chains to many receptor molecules has been extensively demonstrated in animals and fungi (Hurley and Emr, 2006), and this modification serves to direct host molecules into a pathway that ultimately delivers the protein to the lysosome/vacuole for degradation. Attachment of the ubiquitin chain to a substrate is performed by an E3 ligase at the plasma membrane—these ligases are the Rsp5p and c-Cbl gene products in opisthokonts.

These initial stages of the pathway serve to deliver ubiquitylated cargo to the multivesicular body (MVB), a late endosomal structure characterised by the presence of small vesicles located within a separate bounding membrane; creation of this structure requires inward budding of the surface membrane. This budding process, recognition of ubiquitylated cargo and ultimately deubiquitylation and delivery of the cargo molecule to the lysosome is the responsibility of the ESCRT proteins (endosomal sorting complex required for transport) (Williams and Urbe, 2007). The ESCRT system is a supercomplex of over twenty proteins, consisting of five subcomplexes, 0, I, II, III and III-associated. These assemble from a soluble cytosolic form in a sequential manner to deliver ubiquitylated cargo to the MVB membrane (primarily the action of ESCRT 0, I and II), to deform the membrane (primarily ESCRT III and III-associated) and to ultimately disassemble and recycle the components (ESCRT III-associated). The net result is delivery of ubiquitylated molecules into the membrane of the internal vesicles within the MVB. Knock-out of these factors in yeast and metazoan cells frequently results in severe blockade to lysosomal sorting and function (Babst et al., 2000).

A role for ubiquitylation in endocytosis appears to be ancient. Direct demonstration of a requirement for ubiquitylation in internalization of type I *trans*-membrane domain surface molecules has recently been obtained for *T. brucei*, an excavate (Chung et al., 2008), and comparative genomics indicates that the ESCRT complexes I, II, III and III-associated are likewise widely distributed across the eukaryotes, thus strongly implying an ancient origin predating the LCEA (Field et al., 2007; Leung et al., 2008). This, together with the presence of structures morphologically resembling MVBs in diverse taxa argues for conservation of the basic ubiquitylation, MVB and ESCRT systems (Tse et al., 2004; Yang et al., 2004; Walker et al., 2006; Allen et al., 2007). It is not clear if a stimulus is required to initiate ubiquitylation in non-opisthokont taxa, or how numerous the substrates may be in any given cell type. Our understanding of the cell biology of the ESCRT system in most taxa is extremely limited.

In contrast to the overall conservation of the ESCRT system, and the presence of a functional ubiquitylation endocytic pathway, several aspects appear to be rather less well-conserved across the eukaryotes. Most significantly, the ESCRT 0 heterodimer appears to be specific to the Opisthokonta. This is functionally important as this complex is responsible for both recognition of ubiquitylated cargo at the MVB, and also recruitment of ESCRT I factors, suggesting that the initial docking and recognition mechanisms are distinct between opisthokonts and all other eukaryotes (Field et al., 2007; Leung et al., 2008). It is possible that ESCRT I, in fact, acts as the cargo receptor in these systems, consistent with a partial defect to sorting as observed for disruption of interactions between Vps27p and Vps23p in yeast (Bilodeau et al., 2003). However, the presence of a novel complex, or of extreme divergence precluding recognition cannot be excluded.

This likely mechanistic distinction is also echoed in the evolutionary history of the epsin and E3 ubiquitin ligase family. Proteins with an epsin N-terminal homology domain (ENTH) occur in most taxa and are clearly an ancient feature. However, in opisthokonts ENTH proteins are found as both epsin and epsin-related, or epsinR, forms. The latter differ from epsin in lacking a UIM and indeed appear to be the widely distributed and ancient form. This evolutionary distribution is reflected in a second epsin-like family, eps15; again a UIM-containing form is found only in opisthokonts,

while the UIM-lacking eps15R is broadly distributed (Field et al., 2007). Even more extreme is the absence of clear orthologues of the Rsp5 and c-Cbl E3 ligases from non-Opisthokonta, and therefore the ligases responsible for surface protein ubiquitylation in the majority of taxa remain unidentified (Chung et al., 2008). The complexity of the ubiquitin ligase family makes the identification of the non-opisthokont E3 ligases responsible for endocytic function a significant challenge. Frequent secondary losses are also seen amongst the ESCRTs, especially in complexes I and II (Field et al., 2007; Leung et al., 2008).

Rather more robust retention of ESCRT III and III-associated subunits is observed, and hints that these factors may encode a minimal MVB trafficking system. Similar to the tethers, the components of the ESCRT I and II complexes exhibit no obvious homology to one another or to other ESCRT machinery. However, the ESCRT III and III-associated complexes are composed of subunits derived from paralogous protein family expansion. The ESCRT III proteins Vps20, 32 and the ESCRT III-associated protein Vps60 are all related and contain a SNF7 domain. Further, CHMP7 is a protein containing two tandem SNF7 domains, which presumably are the result of a tandem duplication event, and which interacts with the ESCRT III complex. Similarly Vps2A, Vps2B, and Vps24 are homologous ESCRT III components, while their relative Vps46 acts in the ESCRT III-associated complex. All of these factors appear to be widely distributed in eukaryotes and derived via gene duplications that established the families prior to the LCEA (Leung et al., 2008). Importantly an archaeal homologue of Vps2 has been identified (Obita et al., 2007) and can be used to root the tree, as one would expect, between the ESCRT III and ESCRT III-associated complex subunits, and which also provides a link beyond the prokaryote–eukaryote divide. A final ESCRT III-associated subunit, Vps4 also has an archael homologue (Obita et al., 2007). Given that the Snf7 and Vps2 domain proteins appear capable of assembling lattice-like structures on cell membranes (Hanson et al., 2008), and that Vps4, an ATPase, can introduce energy into the system likely facilitating conformational change, such a minimal configuration could represent an ancestral mechanism for membrane deformation. In common with the Rab proteins, the ESCRT system provides an example of potential prokaryotic origins for systems until recently considered to be solely the domain of eukaryotic organisms. The ESCRT system therefore encapsulates within a single system many of the forces that have likely shaped the membrane-trafficking system in general: prokaryotic building blocks hinting at origins of the system; component amplification driven by gene duplication; acquisition of additional factors to enhance and solidify function; loss of features after the LCEA in some lineages and lineage-specific innovation in others.

8. Conclusions—a model for autogenous organelle evolution

Vesicle formation and vesicle fusion were originally thought to be quite separate cellular events, and competing theories were held about how specificity is encoded. Increasingly, however, it has become apparent that the two processes are intimately tied together; SNAREs interact with vesicle coats (Mancias and Goldberg, 2007), as do tethers (Cai et al., 2007b, ENTH domain cargo adaptors bind to SNAREs (Miller et al., 2007; Chidambaram et al., 2008). SM proteins can be part of tethering complexes that then interact with Rab GTPases (Seals et al., 2000). These connections provide an additional, non-genetic, level of information on the organisation of the cell. On the one hand they can be seen as an obstacle to evolutionary reconstruction of cellular history. On the other hand, it is precisely the similarities in the modes of evolution of the various specificity-determining factors, i.e. coats, Rabs and SNAREs and knowledge of their interactions that provided the step forward beyond focusing on the history of single factors to the concept of co-evolution of trafficking machinery giving rise to endomembrane complexity.

The story of how specificity is encoded in trafficking pathways and how identity is encoded for compartments is still incomplete. The story of how the trafficking system evolved is incomplete as well, and not surprisingly. As our understanding of the former is improved, we will have more and more targets to examine by molecular evolutionary means and more information to feed into our evolutionary reconstructions. The connection of trafficking organelles to other compartments such as peroxisomes and the nucleus are equally promising for understanding both cellular organisation and evolution. And as evolutionary analysis becomes more sensitive and robust, we may be further able to unravel our cellular history and the origins of the eukaryotic cell.

Charles Darwin almost certainly never contemplated that his ideas would extend to the molecular details of cellular function and the membrane-trafficking system. The ER and the Golgi apparatus were first described some 15 years after his passing and he could not know that genes encoding protein factors underlying trafficking between endomembranous organelles or that mathematical theories derived decades after the publication of his work would be used to untangle the history of the eukaryotic cell. And so it is a reminder that Darwin's remarkable insights contributed a powerful and incredibly well-supported hypothesis, one that itself continues to mature and evolve, that the concept of evolution by natural selection may be rigorously applied to organisms and structures well beyond its original scope. This is a remarkable testament to the power of clear thinking and careful objective observation.

Acknowledgements

This work was supported by project and program grants from the Wellcome Trust (to MCF), by the Parke-David Foundation (to J.B.D.) and the Medical Research Council (to A.P.). We are also highly grateful to Michael Rout for discussions.

References

Ali BR, Wasmeier C, Lamoreux L, Strom M, Seabra MC. Multiple regions contribute to membrane targeting of Rab GTPases. J Cell Sci 2004;117:6401–12.
Allen CL, Liao D, Chung WL, Field MC. Dileucine signal-dependent and AP-1-independent targeting of a lysosomal glycoprotein in *Trypanosoma brucei*. Mol Biochem Parasitol 2007;156:175–90.
Alory C, Balch WE. Molecular evolution of the Rab-escort-protein/guanine-nucleotide-dissociation-inhibitor superfamily. Mol Biol Cell 2003;14:3857–67.
Antonin W, Fasshauer D, Becker S, Jahn R, Schneider TR. Crystal structure of the endosomal SNARE complex reveals common structural principles of all SNAREs. Nat Struct Biol 2002;9:107–11.
Arac D, Dulubova I, Pei J, Huryeva I, Grishin NV, Rizo J. Three-dimensional structure of the rSly1 N-terminal domain reveals a conformational change induced by binding to syntaxin 5. J Mol Biol 2005;346:589–601.
Ayong L, Pagnotti G, Tobon AB, Chakrabarti D. Identification of *Plasmodium falciparum* family of SNAREs. Mol Biochem Parasitol 2007;152:113–22.
Babst M, Odorizzi G, Estepa EJ, Emr SD. Mammalian tumor susceptibility gene 101 (TSG101) and the yeast homologue, Vps23p, both function in late endosomal trafficking. Traffic 2000;1:248–58.
Barlowe C, Orci L, Yeung T, Hosobuchi M, Hamamoto S, Salama N, et al. COPII: a membrane coat formed by Sec proteins that drive vesicle budding from the endoplasmic reticulum. Cell 1994;77:895–907.
Besteiro S, Coombs GH, Mottram JC. The SNARE protein family of *Leishmania major*. BMC Genom 2006;7:250.
Bethune J, Wieland F, Moelleken J. COPI-mediated transport. J Membr Biol 2006;211:65–79.
Bilodeau PS, Winistorfer SC, Kearney WR, Robertson AD, Piper RC. Vps27-Hse1 and ESCRT-I complexes cooperate to increase efficiency of sorting ubiquitinated proteins at the endosome. J Cell Biol 2003;163:237–43.
Blouin C, Butt D, Roger AJ. Rapid evolution in conformational space: a study of loop regions in a ubiquitous GTP binding domain. Protein Sci 2004;13:608–16.
Bock JB, Matern HT, Peden AA, Scheller RH. A genomic perspective on membrane compartment organization. Nature 2001;409:839–41.

Boehm M, Bonifacino JS. Adaptins: the final recount. Mol Biol Cell 2001;12:2907–20.
Bogdanovic A, Bennett N, Kieffer S, Louwagie M, Morio T, Garin J, et al. Syntaxin 7, syntaxin 8, Vti1 and VAMP7 (vesicle-associated membrane protein 7) form an active SNARE complex for early macropinocytic compartment fusion in Dictyostelium discoideum. Biochem J 2002;368:29–39.
Bogdanovic A, Bruckert F, Morio T, Satre M. A syntaxin 7 homologue is present in *Dictyostelium discoideum* endosomes and controls their homotypic fusion. J Biol Chem 2000;275:36691–7.
Bonifacino JS, Glick BS. The mechanisms of vesicle budding and fusion. Cell 2004;116:153–66.
Bonifacino JS, Hurley JH. Retromer. Curr Opin Cell Biol 2008;20:427–36.
Bracher A, Weissenhorn W. Structural basis for the Golgi membrane recruitment of Sly1p by Sed5p. EMBO J 2002;21:6114–24.
Burkhardt P, Hattendorf DA, Weis WI, Fasshauer D. Munc18a controls SNARE assembly through its interaction with the syntaxin N-peptide. EMBO J 2008;27:923–33.
Cai H, Reinisch K, Ferro-Novick S. Coats, tethers, Rabs, and SNAREs work together to mediate the intracellular destination of a transport vesicle. Dev Cell 2007a;12:671–82.
Cai H, Yu S, Menon S, Cai Y, Lazarova D, Fu C, et al. TRAPPI tethers COPII vesicles by binding the coat subunit Sec23. Nature 2007b;445:941–4.
Carlton JM, Hirt RP, Silva JC, Delcher AL, Schatz M, Zhao Q, et al. Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. Science 2007;315:207–12.
Cavalier-Smith T. The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. Int J Syst Evol Microbiol 2002;52:297–354.
Chidambaram S, Zimmermann J, von Mollard GF. ENTH domain proteins are cargo adaptors for multiple SNARE proteins at the TGN endosome. J Cell Sci 2008;121:329–38.
Chung WL, Leung KF, Carrington M, Field MC. Ubiquitylation is required for degradation of transmembrane surface proteins in trypanosomes. Traffic 2008;9:1681–97.
Collins BM, Skinner CF, Watson PJ, Seaman MN, Owen DJ. Vps29 has a phosphoesterase fold that acts as a protein interaction scaffold for retromer assembly. Nat Struct Mol Biol 2005;12:594–602.
Collins RF, Schreiber AD, Grinstein S, Trimble WS. Syntaxins 13 and 7 function at distinct steps during phagocytosis. J Immunol 2002;169:3250–6.
Dacks JB. Evolution of the SNARE protein superfamily. In: Banfield D, Fillippini F, editors, Banfield D and Fillippini F's. SNARE Proteins: Landes Bioscience; 2007. http://www.eurekah.com/chapter/3344.
Dacks JB, Davis LA, Sjogren AM, Andersson JO, Roger AJ, Doolittle WF. Evidence for Golgi bodies in proposed 'Golgi-lacking' lineages. Proc Biol Sci 2003;270(Suppl 2):S168–71.
Dacks JB, Doolittle WF. Reconstructing/deconstructing the earliest eukaryotes: how comparative genomics can help. Cell 2001;107:419–25.
Dacks JB, Doolittle WF. Novel syntaxin gene sequences from *Giardia*. Trypanosoma and algae: implications for the ancient evolution of the eukaryotic endomembrane system. J Cell Sci 2002;115:1635–42.
Dacks JB, Doolittle WF. Molecular and phylogenetic characterization of syntaxin genes from parasitic protozoa. Mol Biochem Parasitol 2004;136:123–36.
Dacks JB, Field MC. Eukaryotic cell evolution from a genomic perspective: the endomembrane system. In: Hirt RP, Horner DS, Hirt RP, Horner DSs, editors. Organelles, genomes and eukaryote phylogeny: an evolutionary synthesis in the age of genomics. London: CRC Press; 2004. p. 309–34.
Dacks JB, Field MC. Evolution of the eukaryotic membrane-trafficking system: origin, tempo and mode. J Cell Sci 2007;120:2977–85.
Dacks JB, Poon PP, Field MC. Phylogeny of endocytic components yields insight into the process of nonendosymbiotic organelle evolution. Proc Natl Acad Sci U S A 2008;105:588–93.
Damen E, Krieger E, Nielsen JE, Eygensteyn J, van Leeuwen JE. The human Vps29 retromer component is a metallo-phosphoesterase for a cation-independent mannose 6-phosphate receptor substrate peptide. Biochem J 2006;398:399–409.
Darwin C. On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. London: John F. Murray; 1861.
de Renzis S, Sonnichsen B, Zerial M. Divalent Rab effectors regulate the sub-compartmental organization and sorting of early endosomes. Nat Cell Biol 2002;4:124–33.
Devos D, Dokudovskaya S, Alber F, Williams R, Chait BT, Sali A, et al. Components of coated vesicles and nuclear pore complexes share a common molecular architecture. PLoS Biol 2004;2:e380.
Dong JH, Wen JF, Tian HF. Homologs of eukaryotic Ras superfamily proteins in prokaryotes and their novel phylogenetic correlation with their eukaryotic analogs. Gene 2007;396:116–24.
Donohoe BS, Kang BH, Staehelin LA. Identification and characterization of COPIa- and COPIb-type vesicle classes associated with plant and algal Golgi. Proc Natl Acad Sci U S A 2007;104:163–8.
Duden R, Griffiths G, Frank R, Argos P, Kreis TE, Beta-COP. a 110 kd protein associated with non-clathrin-coated vesicles and the Golgi complex, shows homology to beta-adaptin. Cell 1991;64:649–65.
Dulubova I, Yamaguchi T, Gao Y, Min SW, Huryeva I, Sudhof TC, et al. How Tlg2p/syntaxin 16 'snares' Vps45. EMBO J 2002;21:3620–31.
Elde NC, Morgan G, Winey M, Sperling L, Turkewitz AP. Elucidation of clathrin-mediated endocytosis in *Tetrahymena* reveals an evolutionarily convergent recruitment of dynamin. PLoS Genet 2005;1:e52.
Fasshauer D, Sutton RB, Brunger AT, Jahn R. Conserved structural features of the synaptic fusion complex: SNARE proteins reclassified as Q- and R-SNAREs. Proc Natl Acad Sci U S A 1998;95:15781–6.
Faulstich D, Auerbach S, Orci L, Ravazzola M, Wegchingel S, Lottspeich F, et al. Architecture of coatomer: molecular characterization of delta-COP and protein interactions within the complex. J Cell Biol 1996;135:53–61.
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.
Gurkan C, Stagg SM, Lapointe P, Balch WE. The COPII cage: unifying principles of vesicle coat assembly. Nat Rev Mol Cell Biol 2006;7:727–38.
Hala M, Elias M, Zarsky V. A specific feature of the angiosperm Rab escort protein (REP) and evolution of the REP/GDI superfamily. J Mol Biol 2005;348:1299–313.
Hanson PI, Roth R, Lin Y, Heuser JE. Plasma membrane deformation by circular arrays of ESCRT-III protein filaments. J Cell Biol 2008;180:389–402.
Happel N, Honing S, Neuhaus JM, Paris N, Robinson DG, Holstein SE. Arabidopsis mu A-adaptin interacts with the tyrosine motif of the vacuolar sorting receptor VSR-PS1. Plant J 2004;37:678–93.
Hartman H, Fedorov A. The origin of the eukaryotic cell: a genomic investigation. Proc Natl Acad Sci U S A 2002;99:1420–5.
Hierro A, Rojas AL, Rojas R, Murthy N, Effantin G, Kajava AV, et al. Functional architecture of the retromer cargo-recognition complex. Nature 2007;449:1063–7.
Hoepfner D, Schildknegt D, Braakman I, Philippsen P, Tabak HF. Contribution of the endoplasmic reticulum to peroxisome formation. Cell 2005;122:85–95.
Hu SH, Latham CF, Gee CL, James DE, Martin JL. Structure of the Munc18c/Syntaxin4 N-peptide complex defines universal features of the N-peptide binding mode of Sec1/Munc18 proteins. Proc Natl Acad Sci U S A 2007;104:8773–8.
Hurley JH, Emr SD. The ESCRT complexes: structure and mechanism of a membrane-trafficking network. Annu Rev Biophys Biomol Struct 2006;35:277–98.
Iwabe N, Kuma K, Miyata T. Evolution of gene families and relationship with organismal evolution: rapid divergence of tissue-specific genes in the early evolution of chordates. Mol Biol Evol 1996;13:483–93.
Jekely G. Small GTPases and the evolution of the eukaryotic cell. Bioessays 2003;25:1129–38.
Katz L, Brennwald P. Testing the 3Q:1R "rule": mutational analysis of the ionic "zero" layer in the yeast exocytic SNARE complex reveals no requirement for arginine. Mol Biol Cell 2000;11:3849–58.
Kissmehl R, Schilde C, Wassmer T, Danzer C, Nuehse K, Lutter K, et al. Molecular identification of 26 syntaxin genes and their assignment to the different trafficking pathways in *Paramecium*. Traffic 2007;8:523–42.
Koumandou VL, Dacks JB, Coulson RM, Field MC. Control systems for membrane fusion in the ancestral eukaryote; evolution of tethering complexes and SM proteins. BMC Evol Biol 2007;7:29.
Lal K, Field MC, Carlton JM, Warwicker J, Hirt RP. Identification of a very large Rab GTPase family in the parasitic protozoan *Trichomonas vaginalis*. Mol Biochem Parasitol 2005;143:226–35.
Lane CE, Archibald JM. The eukaryotic tree of life: endosymbiosis takes its TOL. Trends Ecol Evol 2008;23:268–75.
Lefkir Y, de Chassey B, Dubois A, Bogdanovic A, Brady RJ, Destaing O, et al. The AP-1 clathrin-adaptor is required for lysosomal enzymes sorting and biogenesis of the contractile vacuole complex in *Dictyostelium* cells. Mol Biol Cell 2003;14:1835–51.
Leung KF, Dacks JB, Field MC. Evolution of the multivesicular body ESCRT machinery; retention across the eukaryotic lineage. Traffic 2008;9:1698–716.
Maier AG, Webb H, Ding M, Bremser M, Carrington M, Clayton C. The coatomer of *Trypanosoma brucei*. Mol Biochem Parasitol 2001;115:55–61.
Mancias JD, Goldberg J. The transport signal on Sec22 for packaging into COPII-coated vesicles is a conformational epitope. Mol Cell 2007;26:403–14.
Mazzarello P. A unifying concept: the history of cell theory. Nat Cell Biol 1999;1:E13–5.
Miller SE, Collins BM, McCoy AJ, Robinson MS, Owen DJ. A SNARE-adaptor interaction is a new mode of cargo recognition in clathrin-coated vesicles. Nature 2007;450:570–4.
Misura KM, Scheller RH, Weis WI. Three-dimensional structure of the neuronal-Sec1-syntaxin 1a complex. Nature 2000;404:355–62.
Morgan GW, Allen CL, Jeffries TR, Hollinshead M, Field MC. Developmental and morphological regulation of clathrin-mediated endocytosis in *Trypanosoma brucei*. J Cell Sci 2001;114:2605–15.
Mullock BM, Smith CW, Ihrke G, Bright NA, Lindsay M, Parkinson EJ, et al. Syntaxin 7 is localized to late endosome compartments, associates with Vamp 8, and is required for late endosome-lysosome fusion. Mol Biol Cell 2000;11:3137–53.
Nakada-Tsukui K, Saito-Nakano Y, Ali V, Nozaki T. A retromerlike complex is a novel Rab7 effector that is involved in the transport of the virulence factor cysteine protease in the enteric protozoan parasite *Entamoeba histolytica*. Mol Biol Cell 2005;16:5294–303.
Novick P, Field C, Schekman R. Identification of 23 complementation groups required for post-translational events in the yeast secretory pathway. Cell 1980;21:205–15.
Obita T, Saksena S, Ghazi-Tabatabai S, Gill DJ, Perisic O, Emr SD, et al. Structural basis for selective recognition of ESCRT-III by the AAA ATPase Vps4. Nature 2007;449:735–9.
Oro AE. The primary cilia, a 'Rab-id' transit system for hedgehog signaling. Curr Opin Cell Biol 2007;19:691–6.

Ossig R, Schmitt HD, de Groot B, Riedel D, Keranen S, Ronne H, et al. Exocytosis requires asymmetry in the central layer of the SNARE complex. EMBO J 2000;19:6000–10.
Pelham HR. Insights from yeast endosomes. Curr Opin Cell Biol 2002;14:454–62.
Pereira-Leal JB. The Ypt/Rab family and the evolution of trafficking in fungi. Traffic 2008;9:27–38.
Pereira-Leal JB, Seabra MC. The mammalian Rab family of small GTPases: definition of family and subfamily sequence motifs suggests a mechanism for functional specificity in the Ras superfamily. J Mol Biol 2000;301:1077–87.
Pereira-Leal JB, Seabra MC. Evolution of the Rab family of small GTP-binding proteins. J Mol Biol 2001;313:889–901.
Pfeffer S, Aivazian D. Targeting Rab GTPases to distinct membrane compartments. Nat Rev Mol Cell Biol 2004;5:886–96.
Podar M, Wall MA, Makarova KS, Koonin EV. The prokaryotic V4R domain is the likely ancestor of a key component of the eukaryotic vesicle transport system. Biol Direct 2008;3:2.
Rabouille C, Klumperman J. Opinion: The maturing role of COPI vesicles in intra-Golgi transport. Nat Rev Mol Cell Biol 2005;6:812–7.
Rasteiro R, Pereira-Leal JB. Multiple domain insertions and losses in the evolution of the Rab prenylation complex. BMC Evol Biol 2007;7:140.
Rink J, Ghigo E, Kalaidzidis Y, Zerial M. Rab conversion as a mechanism of progression from early to late endosomes. Cell 2005;122:735–49.
Robinson MS. Adaptable adaptors for coated vesicles. Trends Cell Biol 2004;14:167–74.
Roger AJ. Reconstructing early events in eukaryotic evolution. Am Nat 1999;154:S146–63.
Rojo E, Zouhar J, Kovaleva V, Hong S, Raikhel NV. The AtC-VPS protein complex Is localized to the tonoplast and the prevacuolar compartment in *Arabidopsis*. Mol Biol Cell 2003;14:361–9.
Rutherford S, Moore I. The *Arabidopsis* Rab GTPase family: another enigma variation. Curr Opin Plant Biol 2002;5:518–28.
Saito-Nakano Y, Loftus BJ, Hall N, Nozaki T. The diversity of rab GTPases in *Entamoeba histolytica*. Exp Parasitol 2005;110:244–52.
Sanderfoot A. Increases in the number of SNARE genes parallels the rise of multicellularity among the green plants. Plant Physiol 2007;144:6–17.
Sanderfoot AA, Kovaleva V, Bassham DC, Raikhel NV. Interactions between syntaxins identify at least five SNARE complexes within the Golgi/prevacuolar system of the *Arabidopsis* cell. Mol Biol Cell 2001;12:3733–43.
Schilde C, Wassmer T, Mansfeld J, Plattner H, Kissmehi R. A multigene family encoding R-SNAREs in the ciliate *Paramecium tetraurelia*. Traffic 2006;7:440–55.
Schledzewski K, Brinkmann H, Mendel RR. Phylogenetic analysis of components of the eukaryotic vesicle transport system reveals a common origin of adaptor protein complexes 1, 2 and 3 and the F subcomplex of the coatomer COPI. J Mol Evol 1999;48:770–8.
Schluter A, Fourcade S, Ripp R, Mandel JL, Poch O, Pujol A. The evolutionary origin of peroxisomes: an ER-peroxisome connection. Mol Biol Evol 2006;23:838–45.
Seals DF, Eitzen G, Margolis N, Wickner WT, Price A. A Ypt/Rab effector complex containing the Sec1 homolog Vps33p is required for homotypic vacuole fusion. Proc Natl Acad Sci U S A 2000;97:9402–7.
Seaman MN. Recycle your receptors with retromer. Trends Cell Biol 2005;15:68–75.
Simmen T, Honing S, Icking A, Tikkanen R, Hunziker W. AP-4 binds basolateral signals and participates in basolateral sorting in epithelial MDCK cells. Nat Cell Biol 2002;4:154–9.
Smith SS, Pfluger SL, Hjort E, McArthur AG, Hager KM. Molecular evolution of the vesicle coat component betaCOP in *Toxoplasma gondii*. Mol Phylogenet Evol 2007;44:1284–94.
Sollner T, Whiteheart SW, Brunner M, Erdjument-Bromage H, Geromanos S, Tempst P, et al. SNAP receptors implicated in vesicle targeting and fusion. Nature 1993;362:318–24.
Stanier RY. The origins of photosynthesis in Eukaryotes. In: Carlile MJ, Skehel JJ, Carlile MJ, Skehel JJs, editors. Evolution in the Microbial World. Cambridge, U.K.: Cambridge University Press; 1974. p. 219–40.
Starai VJ, Hickey CM, Wickner W. HOPS proofreads the trans-SNARE complex for yeast vacuole fusion. Mol Biol Cell 2008;19:2500–8.
Steegmaier M, Oorschot V, Klumperman J, Scheller RH. Syntaxin 17 is abundant in steroidogenic cells and implicated in smooth endoplasmic reticulum membrane dynamics. Mol Biol Cell 2000;11:2719–31.
Steele-Mortimer O, Clague MJ, Huber LA, Chavrier P, Gruenberg J, Gorvel JP. The N-terminal domain of a rab protein is involved in membrane-membrane recognition and/or fusion. EMBO J 1994;13:34–41.
Sutton RB, Fasshauer D, Jahn R, Brunger AT. Crystal structure of a SNARE complex involved in synaptic exocytosis at 2, 4 A resolution. Nature 1998;395: 347–53.
Sztul E, Lupashin V. Role of tethering factors in secretory membrane traffic. Am J Physiol Cell Physiol 2006;290:C11–26.
Theos AC, Tenza D, Martina JA, Hurbain I, Peden AA, Sviderskaya EV, et al. Functions of adaptor protein (AP)-3 and AP-1 in tyrosinase sorting from endosomes to melanosomes. Mol Biol Cell 2005;16:5356–72.
Toonen RF, Verhage M. Vesicle trafficking: pleasure and pain from SM genes. Trends Cell Biol 2003;13:177–86.
Tse YC, Mo B, Hillmer S, Zhao M, Lo SW, Robinson DG, et al. Identification of multivesicular bodies as prevacuolar compartments in *Nicotiana tabacum* BY-2 cells. Plant Cell 2004;16:672–93.
van der Giezen M. Mitochondrial amazement. EMBO Rep 2006;7:478.
Vernoud V, Horton AC, Yang Z, Nielsen E. Analysis of the small GTPase gene superfamily of *Arabidopsis*. Plant Physiol 2003;131:1191–208.
Volchuk A, Wang Q, Ewart HS, Liu Z, He L, Bennett MK, et al. Syntaxin 4 in 3T3-L1 adipocytes: regulation by insulin and participation in insulin-dependent glucose transport. Mol Biol Cell 1996;7:1075–82.
Walker G, Dacks JB, Martin Embley T. Ultrastructural description of *Breviata anathema*, n. gen., n. sp., the organism previously studied as "*Mastigamoeba invertens*". J Eukaryot Microbiol 2006;53:65–78.
Wang D, Guo M, Liang Z, Fan J, Zhu Z, Zang J, et al. Crystal structure of human vacuolar protein sorting protein 29 reveals a phosphodiesterase/nuclease-like fold and two protein-protein interaction sites. J Biol Chem 2005;280: 22962–7.
Whyte JR, Munro S. The Sec34/35 Golgi transport complex is related to the exocyst, defining a family of complexes involved in multiple steps of membrane traffic. Dev Cell 2001;1:527–37.
Williams RL, Urbe S. The emerging shape of the ESCRT machinery. Nat Rev Mol Cell Biol 2007;8:355–68.
Yang M, Coppens I, Wormsley S, Baevova P, Hoppe HC, Joiner KA. The *Plasmodium falciparum* Vps4 homolog mediates multivesicular body formation. J Cell Sci 2004;117:3831–8.
Yoshizawa AC, Kawashima S, Okuda S, Fujita M, Itoh M, Moriya Y, et al. Extracting sequence motifs and the phylogenetic features of SNARE-dependent membrane traffic. Traffic 2006;7:1104–18.
Zerial M, McBride H. Rab proteins as membrane organizers. Nat Rev Mol Cell Biol 2001;2:107–17.
Zwilling D, Cypionka A, Pohl WH, Fasshauer D, Walla PJ, Wahl MC, et al. Early endosomal SNAREs form a structurally conserved SNARE complex and fuse liposomes with multiple topologies. EMBO J 2007;26:9–18.