Evolution of the endomembrane systems of trypanosomatids

Venkatesh, Divya; Boehm, Cordula; Barlow, Lael D.; Nankissoor, Nerissa N.; O'Reilly, Amanda; Kelly, Steven

Published in:
Journal of Cell Science

DOI:
10.1242/jcs.197640

Publication date:
2017

 Licence:
CC BY

Document Version
Publisher's PDF, also known as Version of record

Link to publication in Discovery Research Portal

Citation for published version (APA):
Venkatesh, D., Boehm, C., Barlow, L. D., Nankissoor, N. N., O'Reilly, A., Kelly, S., Dacks, J. B., & Field, M. C. (2017). Evolution of the endomembrane systems of trypanosomatids: conservation and specialisation. Journal of Cell Science, 130(8), 1421-1434. https://doi.org/10.1242/jcs.197640
Evolution of the endomembrane systems of trypanosomatids – conservation and specialisation

Divya Venkatesh1,2, Cordula Boehm1, Lael D. Barlow3, Nerissa N. Nankissoor3, Amanda O’Reilly2, Steven Kelly4, Joel B. Dacks3 and Mark C. Field1,*

ABSTRACT
Parasite surfaces support multiple functions required for survival within their hosts, and maintenance and functionality of the surface depends on membrane trafficking. To understand the evolutionary history of trypanosomatid trafficking, where multiple lifestyles and mechanisms of host interactions are known, we examined protein families central to defining intracellular compartments and mediating transport, namely Rab, SNAREs and RabGAPs, across all available Euglenozoa genomes. Bodonids possess a large trafficking repertoire, which is mainly retained by the Trypanosoma cruzi group, with extensive losses in other lineages, particularly African trypanosomes and phytomonads. There are no large-scale expansions or contractions from an inferred ancestor, excluding direct associations between parasitism or host range. However, we observe stepwise secondary losses within Rab and SNARE cohorts (but not RabGAPs). Major changes are associated with endosomal and late exocytic pathways, consistent with the diversity in surface proteomes between trypanosomatids and mechanisms of interaction with the host. Along with the conserved core family proteins, several lineage-specific members of the Rab (but not SNARE) family were found. Significantly, testing predictions of SNARE complex composition by proteomics confirms generalised retention of function across eukaryotes.

KEY WORDS: Rab, SNARE, Trafficking, Molecular evolution, TBC domain, Trypanosoma, Proteomics

INTRODUCTION
Membrane trafficking mediates delivery of macromolecules to discrete intracellular compartments from the site of uptake or synthesis. Trafficking is essential to nearly all eukaryotic cells, contributing towards nutrient acquisition, protein processing and turnover, and, in multicellular organisms, it also participates in higher-order tissue organisation. The importance of membrane transport is reflected in the many diseases associated with trafficking, including diabetes, Alzheimer’s disease and cystic fibrosis (Birault et al., 2013; Ollkonen and Ikonen, 2006; Rajendran and Annaert, 2012; Seixas et al., 2013). Development of membrane trafficking was likely a major evolutionary driver enabling the transition from prokaryotic to eukaryotic cells (Field et al., 2011; Dacks et al., 2016). For many pathogens, trafficking has special relevance in maintaining the host–parasite interface, which is at the cell surface, and both the surface and underlying trafficking apparatus are intimately connected with immune evasion, pathogenesis and life cycle progression (Manna et al., 2013).

Membrane transport requires vesicle formation, translocation, tethering, docking and fusion to release cargo (Bonifacino and Glick, 2004), which is achieved through coordinated action by Rab and ARF GTPases, coat complexes, tethers and SNAREs. Members of these paralogous families encode specificity for individual transport events to defined organelles; evolutionary reconstructions suggest stepwise evolution from a simpler ancestral system before the last eukaryotic common ancestor (LECA) (Devos et al., 2004, 2006; Dokudovskaya et al., 2006; Field and Dacks, 2009; Schlacht et al., 2014). A conserved core of Rab, RabGAP and SNARE proteins has been derived by reconstructing eukaryotic evolutionary history and broadly supports this model (Arasaki et al., 2015; Diekmann et al., 2011; Elias et al., 2012; Kienle et al., 2009a; Vedovato et al., 2009; Yoshizawa et al., 2006). Beyond this are examples of lineage-specific features in the endosomal sorting complexes required for transport (ESCRT) system, sortilins and ARF GTPases in metazoans (Field et al., 2007; Gabernet-Castello et al., 2013; Leung et al., 2008). In parasitic organisms, adaptin complexes, important cargo selectors that are otherwise well conserved in most lineages, are rather variable, possibly because of specific adaptation (Woo et al., 2015).

Kinetoplastids are unicellular flagellated protists within the supergroup Excavata, and include free-living species and parasites that cause many important human diseases, together with species afflicting animals and plants. Kinetoplastids exhibit varied lifestyles, host range, distribution and specialisations over long evolutionary periods. Furthermore, kinetoplastid surfaces are highly divergent between lineages, likely reflecting specialisations within membrane transport (Gadelha et al., 2015; Manna et al., 2013). More remarkable is the highly distinct nature of the proteins and glycoconjugates present at the cell surface, with recent data indicating that many surface proteins are restricted to the parasite interface, which is at the cell surface, and both the surface and underlying trafficking apparatus are intimately connected with immune evasion, pathogenesis and life cycle progression (Manna et al., 2013).

SNAREs, mediators of membrane fusion, possess a characteristic domain, and are classified as Q- or R-SNAREs (Bock et al., 2001; Fasshauer et al., 1998). Typically, three Q- (a, b or c) and one R-SNARE form a complex. The complement of Qa and R-SNAREs in...
the LECA has been established with broad taxonomic sampling (Arasaki et al., 2015; Vedovato et al., 2009), but the Qb and Qc families were only assessed by comparative genomics nearly a decade ago, using the limited sampling of genomes then available (Kloepper et al., 2007; Yoshizawa et al., 2006) and warrant re-examination to solidify the definition of the LECA complement. It is key to have a robust estimate of the overall SNARE complement in the LECA as the starting point from which we assume kinetoplastids evolved. Rabs are small GTPases and well established markers of organelar identity (Brighouse et al., 2010; Pereira-Leal and Seabra, 2001). A broad repertoire of ~23 Rabs are predicted in the LECA (Elias et al., 2012). With the exception of metazoa and vascular plants, where great expansion is evident, Rab proteins have evolved mainly through small scale expansions and secondary losses, plus emergence of novel paralogs (Elias et al., 2012; Klopper et al., 2012). Rabs are regulated by GTPase-activating proteins (GAPs), the majority of which possess a Tre-2/Bub2/Cdc16 (TBC) Rab-binding domain. The LECA possessed about ten TBC subtypes and subsequent expansions have included domain swapping (Gabernet-Castello et al., 2013). The greatest number of TBC innovations are in animals and fungi, although novel subclasses are present in a wide range of lineages. Detailed analysis of both Rab and SNARE repertoires in fungi indicates a simple set with minimal variations between different lineages, despite major diversity in morphology and transitions between single and multicellular forms, albeit with evidence for minimisation in some species (Kienle et al., 2009b; Pereira-Leal, 2008).

Our study indicates that a significant proportion of the putative LECA repertoire of these protein families is conserved in kinetoplastids, along with some lineage-specific proteins and secondary losses. We infer a large ancestral repertoire of Rab and SNARE proteins (but not TBC Rab-GAPs) associated with endosomal and late exocytic pathways that appears to undergo step-wise secondary loss in several parasitic trypanosomatid lineages (but not so in the Trypanosoma cruzi group). The proteins associated with these pathways also show high levels of variation in numbers and level of conservation within the kinetoplastids, even though no large-scale changes obviously correlate with parasitism, host range or modes of immune evasion.

RESULTS

Confirmation of the LECA complement of Qb and Qc SNAREs

Although recent analyses have addressed most protein families and subfamilies considered here, Qb and Qc SNAREs have not been characterised in a comparable way. In particular, there are conflicting views on the distributions of Qb-SNARE Novel Plant Syntaxin (NPSN) and the Qc-SNARE Syntaxin of Plants 7 (Syp7) (Sanderfoot et al., 2000). As their names imply, both were proposed conflicting views of the distributions of Qb-SNARE Novel Plant characterised in a comparable way. In particular, there are subfamilies considered here, Qb and Qc SNAREs have not been

The kinetoplastid complement of trafficking proteins

We next considered the genomes of 18 Euglenozoa including basal bodonids and trypanosomatids. We identified candidate genes for 382 Rabs, 552 SNAREs and 307 TBC RabGAPs. Where possible these were assigned to subfamilies by phylogenetic analyses with a reference set of previously defined sequences from across eukaryotes. Such data are readily available for Rabs and TBC RabGAPs (Elias et al., 2012; Gabernet-Castello et al., 2013) and were prepared here for SNAREs using SNARE repertoires from at least one representative of each eukaryotic subgroup, other than the Excavata; Homo sapiens and Saccharomyces cerevisiae as Opisthokonta, Dictyostelium discoideum as Amoebozoa, Phytophthora sojae as SAR-CCTH (a heterogenous group comprising stramenopiles, alveolates and Rhizaria along with cryptomonads, centrohelids, telonemids and haplotyphes) and Arabidopsis thaliana for the Archaeplastida. We were able to assign a considerable proportion of the kinetoplastid representatives of all three families (Fig. 2, 3, 4; Fig. S2). Moreover, 86% of LECA SNAREs, 72% of LECA Rabs and 100% of LECA TBC RabGAPs were identified in one or more members of the lineage, suggesting that the core machinery of trafficking is well conserved (Table 1). Most kinetoplastid proteins were assigned by phylogenetic evidence, but three SNARE proteins, VT1-like, Syx6-like and Syx8-like, were assigned only by BLAST and reverse BLAST as phylogenetic support was low.

Comparing coding content and trafficking repertoire

There is loose correspondence between genome and Rab repertoire sizes (Gabernet-Castello et al., 2013). While this may reflect a connection between genome and compartmental complexity, the huge diversity within eukaryotes makes complexity a difficult parameter to define, and analysing a group of organisms where the basic cellular bauplan is well conserved, allows us to re-evaluate this concept (Fig. S1). The coding content of Bodo saltans is considerably larger than the parasites, while the excavates Naegleria gruberi and Euglena gracilis have larger genomes (Ebenzter et al., 2017; Fritz-Laylin et al., 2010). In keeping with genome size, B. saltans and N. gruberi possess greater numbers of SNAREs, Rabs and TBCs compared to kinetoplastids. Despite comparable coding content for the Trypanosoma brucei and cruzi group trypanosomes, the latter have larger Rab, TBC and SNARE repertoires, suggesting further secondary loss in the former group. Even Leishmania spp. generally have slightly larger numbers of these proteins, despite smaller predicted coding content. Overall, this suggests adaptive shaping of trafficking system gene repertoires rather than changes simply reflecting genome size.

Anterograde trafficking genes are conserved across kinetoplastids

The gene complements for SNAREs and Rabs at the kinetoplastid early secretory pathway are highly conserved. In opisthokonts, the SNARE complex of Qa-Syx5, Qb-Bos1, Qc-Bet1 and R-Sec22 mediates endoplasmic reticulum (ER) to Golgi and intra-Golgi transport (Hardwick and Pelham, 1992; Newman et al., 1990). A second Syx5 complex, constituting Qa-Syx5, Qb-Gos1, Qc-Sf1 and R-Ykt6 is thought to exclusively mediate COPII transport within the Golgi in animals and fungi (Ballensiefen et al., 1998). Whereas Sf1 and Bet1 are two alternate Qc SNAREs acting at early anterograde steps (Kloepper et al., 2007), they were unresolved in our phylogenies (Fig. S2B). It is likely that Bet1 fulfils the requirements of the kinetoplastid early anterograde pathway as two Bet1 paralogs are present, similar to in H. sapiens. Significantly
A. thaliana has an expanded set of four of these proteins (designated as Bet11, Bet12, Sft11 and Sft12) and D. discoidium just one. The other SNAREs in these complexes are all singletons in kinetoplastids except for some species-specific losses (Fig. 5A) and duplication of Gos1 in *Trypanosoma congolense*. Syntaxin 17 cycles between the ER and the ER–Golgi intermediate compartment (ERGIC) in mammalian cells (Itakura and Mizushima, 2013) and acts in autophagosome formation and mitochondrial dynamics (Arasaki et al., 2015; Itakura et al., 2012). Syntaxin 17 is generally patchily distributed across eukaryotes and absent from all kinetoplastids, signifying a loss since divergence from the Heterolobosea. Overall, these data suggest possible differentiation in early secretory events, consistent with multiple budding pathways at the trypanosome ER, but also indicating lineage-specific evolution of these steps (Sevova and Bangs, 2009).

The retrograde pathway from the Golgi to the ER is mediated by a complex of Qa-Syx18, Qb-Sec20, Qc-Sec22 in yeasts (Dilcher et al., 2003). Except Use1, all are conserved across kinetoplastids, bar losses of Syx18 or Sec20 in *Phytomonas EM1* and *HART1* respectively, and duplication of Sec22 in *Trypanosoma theileri* (Fig. 5A). This suggests that canonical retrograde transport systems are present, consistent with the presence of a KDEL receptor and ER-retrieval signals on major ER proteins (Bangs et al., 1996; Schwartz et al., 2013).

In opisthokonts, the late exocytic SNARE complex is a specific ternary complex composed of a plasma membrane syntaxin (denoted SynPM), SNAP25 and Syb1. A SynPM is identified across eukaryotes and absent from all kinetoplastids, signifying a loss since divergence from the Heterolobosea. Overall, these data suggest possible differentiation in early secretory events, consistent with multiple budding pathways at the trypanosome ER, but also indicating lineage-specific evolution of these steps (Sevova and Bangs, 2009).
related lineage-specific duplications. The bodonid *Trypanosoma borrelii* and *cruzi* group possess a second SynPM2 protein. No transmembrane domain (TMD) was detected in SynPM2 of the duplicated SynPM1 genes in *L. braziliensis* and *L. major*, which are otherwise identical, but for the lack of a TMD in one paralog. The TMD is also lost from the single SynPM1 in *Leishmania infantum* and *mexicana* but retained by *L. donovani*. The localisation of *T. brucei* SynPM1 (TbSynPM1) is similar to the TMD-containing SynPM1 from *L. major*, which localises close to the flagellar pocket, while as expected the TMD-lacking SynPM1 in *L. major* is cytosolic (Besteiro et al., 2006). Overall, this pattern indicates a considerable level of species-specific SynPM paralogs, suggesting a common requirement to differentiate this pathway.

Putative Qbc-type SNAREs are present in *B. saltans*, *T. cruzi* and *T. grayi*, but do not cluster with canonical yeast or human Qbc SNAREs, but instead with SAR-CCTH and Excavate sequences (Fig. S2E). The eukaryotic phylogeny of synaptobrevins (Fig. S2D) is also complex, and the presence of brevins in different clusters of diverse taxa indicates that they likely emerged independently in several lineages by loss of the longin domain from a ‘VAMP7’. Moreover, unlike metazoan brevins, which are truncated at the SNARE domain, longin-less VAMP7 orthologues have an extended N-terminal domain in Apicomplexa and Euglenozoa, indicating a separate origin. A single independently evolved VAMP7 with a divergent/undetected longin domain is present in kinetoplastids, and in *T. brucei* this protein TbVAMP7C has an endosomal localisation in juxtaposition to Golgi and lysosomal markers (Fig. 6C) rather than the cell membrane.
Fig. 3. See next page for legend.
Fig. 3. Phylogenetic relationships of kinetoplastid Rabs. The consensus Bayesian topology is shown. Nodes are iconised as colour-coded circles according to MrBayes posterior probabilities and PhyML approximate likelihood ratio test as shown in the key. The figure includes two representative kinetoplastid Rabs from each subfamily cluster along with representative Rabs from across eukaryotes. Assigned kinetoplastid-specific Rabs are in blue, lineage-specific Rabs in red, unassigned orphan Rabs in green, while representative eukaryotic Rabs are black. LECA Rabs for which no kinetoplastid orthologues were found (Rabs 8, 22, 34, 50 and Titan) were not included in the analysis. Rab is used as outgroup.

Rab1, Rab8 and Rab18 are present in the predicted LECA cohort and associated with anterograde pathways (Elias et al., 2012). Interestingly, Rab8 is lost from all kinetoplastids, suggesting differences in post-Golgi trafficking and targeting of material to the flagellum/recycling endosome (Huber et al., 1993; Klöpper et al., 2012). Rab1 and Rab18 have a complex history, and paralogous expansion of Rab1 or Rab18 may have given rise to two lineage-specific Rabs, KSRX1 and UzRX3, which appear across the Euglenids as tandem genes and are also present in the heterolobosids; however, it may be significant that in *T. brucei* these proteins have no obvious role in trafficking (Natesan et al., 2009). While UzRX3 localisation is similar to Rab1 (early secretory pathway), KSRX1 has a more diffuse localisation (Natesan et al., 2009). Interestingly, *B. saltans* has two additional relatives of the Rab1 and Rab18 clade, suggesting further diversification of these pathways.

Post-Golgi pathway Rabs include Rab4, Rab14 and Rab11, which are mostly well-conserved, with the exception that Rab14 is lost by the *T. brucei* group (Fig. 5B). This suggests possible further adaptations of endocytic pathways, an important mechanism for immune evasion in mammalian infective African trypanosomes. Like AP-2, however (Manna et al., 2013), Rab14 is present in the other extracellular trypanosomes of the *T. cruzi* group (Fig. 5B, Fig. S3E). We also identified a Rab11-like protein (Rab11B) that likely arose from duplication of Rab11 at the base of the kinetoplastids, but this is only retained by *B. saltans*, Leishmania spp. and *T. cruzi* group (Fig. 5B; Fig. S3E). Rab11B may provide additional flexibility in surface remodelling, and importantly *T. cruzi* Rab11 localises to the contractile vacuole, regulating trafficking of trans-sialidase to the plasma membrane (Niyogi et al., 2014). Hence, kinetoplastid Rab11 and Rab11B may have distinct functions, as noted for Rab11 paralogs in mammals and Archaeoplastida (Lapierre et al., 2003; Petrzílková and Eliáš, 2014).

### Endocytic, retrograde and lysosomal pathway trafficking

We identified four Qa SNARE proteins (SynE1 and SynE2, Syx16A and SyxB), four Qb SNARE (VTI1-like A, B and NPSNA and NPSNB), four Qc SNAREs (Syx7A, Syx7B, Syx6-like and Syx8-like), and four R-SNAREs (VAMP7A, B, C and D) as orthologues of SNARE proteins involved in endosomal, lysosomal and retrograde trafficking pathways. For several, multiple paralogs exist, particularly in the basal bodonids and the *T. cruzi* group (Fig. 5A; Fig. S3A–D).

Except for *B. saltans* and *T. vivax*, SynE1 is present in all kinetoplastids and possibly duplicated in the parasitic bodonid *T. borrelli*. SynE2 is also present in all kinetoplastids and duplicated in the *T. cruzi* group. Two paralogs of Syx16, Syx16A and Syx16B, are conserved across kinetoplastids. Thus, both SynE and Syn16 underwent duplication at least once early in the lineage evolution. As in the SynPM proteins, all Syx16A proteins lack a TMD, while all Syx16B except *T. borrelli* retain the TMD. *L. braziliensis* also shows lineage-specific duplication of Syx16A, a surprisingly conserved isoform, which suggests that it must retain a function, and is perhaps a regulatory protein. Paralogous Syx16 sequences from other organisms all retain their TMD, suggesting loss of the TMD is kinetoplastid-restricted. Notably, the Sec1/Munc-like protein (SM) Vps45, which regulates Syn16, is also duplicated at the base of the kinetoplastids (Koumandou et al., 2007), and possibly in the last kinetoplastid common ancestor (LKCA), suggesting independent regulation of multiple Syn16 complexes.

The Qb (two VT11 and two NPSN proteins), Qc (two Syp7 and two Syx6-like proteins) and R-SNAREs (four VAMP7 proteins) are expanded in kinetoplastids. Specifically, Qb-NPSNA-2, Qc-Syp7B, R-VAMP7D together with Qa-SynPM2 appear to have arisen at the base of the kinetoplastids but are retained only in the bodonid and *T. cruzi* groups (Fig. 5A). These SNAREs are implicated in plasma membrane trafficking in plants, suggesting they may mediate a similar pathway in the kinetoplastida. In addition, the Syp7B clade also shows evidence for bodonid and *T. cruzi* group-specific expansions. Most SNARE losses are scattered with no obvious pattern, but there are clear instances of co-evolutionary loss of components of predicted SNARE complexes. SNAREs Qa-SynPM2, Qbc-SNAP25 and R-VAMP7D, which could form a putative exocytic complex are all lost in African trypanosomes as well as the *Leishmania* and Phytomonad clades. Qb-VT11-like A2, Qb-NPSNA2, and Qc-Syx6-like 2, predicted to form complexes with VAMP7, are coincidently lost (see Fig. 5A). Except for SNAP-25 (which shows a broader pattern of loss), these are all SNAREs derived from kinetoplastid-specific expansions, and likely indicate loss of kinetoplastid-specific post-Golgi pathways.

Of the endocytic Rabs (Rab5, Rab20 or Rab24, Rab21, Rab22 and Rab50), kinetoplastids retain Rab5, Rab21 and Rab24, albeit with Rab24 bodonid restricted (Fig. 5). Early endocytic Rab5 experienced a single basal duplication in kinetoplastids and both paralogs are stably retained (Fig. S3E) and have acquired distinct functions (Pal et al., 2002). This contrasts with the fungal Ypt5, which is less well retained, with several instances of species or clade-specific expansions and losses (Pereira-Leal, 2008). Rab21 is generally stable with no major expansion in eukaryotes, but there are at least two paralogs at the base of the kinetoplastids (Fig. S3E); Rab21 mediates intermediate endosomal trafficking in *T. brucei*, a conserved function with higher eukaryotes (Ali et al., 2014). An expanded Rab21 subfamily is only retained by bodonids and the *T. cruzi* group, suggesting additional endosomal pathways. In *Phytomonas* spp., which have lost Rab21, there is clear minimisation of endocytic Rabs. Rab2 and Rab6, which mainly mediate retrograde trafficking at the Golgi complex, are also stably retained. Two copies of Rab2 are present in the bodonids and *T. grayi*, but cluster separately from the canonical Rab2.

Rab7, which mediates late endocytic trafficking, is represented by a single paralog in trypanosomatids; however, a Rab7-like protein is found in *B. saltans* and *T. borrelli* (Fig. 5). Rab28 also has an endosomal function in *T. brucei* (Lumb et al., 2011) and except for Phytomonads is present across all kinetoplastids. Rab23 and the Rab-like intraflagellar transport protein 27 (IFT27), which are involved in the biogenesis of cilia/flagella, are both fully retained as expected. Rab32 is involved in the biogenesis of lysosome-related organelles (LROs) and autophagosome formation in mammalian cells (Hirota and Tanaka, 2009), as well as mitochondrial dynamics (Bui et al., 2010; Ortiz-Sandoval et al., 2014). In *T. cruzi*, Rab32 is involved in the biogenesis and maintenance of acidicalcosomes (Niyogi et al., 2015), and presents complex evolution in kinetoplastids, being present only in the bodonids and *T. cruzi* group, where one copy of the canonical Rab32 appears to have
undergone duplication giving rise to Rab32-like proteins (Fig. 5B). Many trypanosomes may thus have reduced or alternative pathways for autophagosome biogenesis and/or mitochondrial dynamics, but by contrast, the emergence of Rab32-like proteins likely suggests the development of novel or more complex LROs, particularly in B. saltans, which has up to five Rab32-like proteins. TBCs associated with endosomal Rabs include TBC-B, which acts on Rab2, Rab7, Rab11 and Rab21 and TBC-D, which is a GAP for Rab1, Rab7, Rab11 and Rab21 and TBC-D, which is a GAP for Rab1, Rab7.
Table 1. Extent of conservation of LECA trafficking family proteins in kinetoplastids

| Family | LECA repertoire | Total in LECA | Absent from kinetoplastids | LECA proteins retained in kinetoplastids | non-LECA proteins |
|--------|----------------|--------------|----------------------------|----------------------------------------|------------------|
| SNAREs |                |              |                            |                                        |                  |
| Qa     | Syx18, Syx5, Syx16, SynE, SynPM, Syx17 | 20           | Syx17                      |                                        | Qa1, Syx16       |
|         |                |              |                            |                                        | (2)              |
| Qb     | Sec20, Bos1, Gos1, VT11, NPSN         | None         |                            |                                        | VT11-like (3),   |
|         |                |              |                            |                                        | NPSN (3)         |
| Qc     | Use1, Bet1, Syx6, Syx8, Syp7          | Use1         |                            |                                        | Bet1 (2), Syp7   |
|         |                |              |                            |                                        | (2), Syx6-like (2)|                  |
| Qbc    | SNAP-25        | Very restricted |                            |                                        |                  |
| R      | Sec22, Ykt6, VAMP7, R_reg              | None         |                            |                                        | VAMP7 (4)        |
| Rab    | 1, 2, 4, 5, 6, 7, 8, 11, 14, 18, 20/24, 21, 22, 23, 28, 32A/B, 34, 50, RTW, IFT27, Titan | 22           | 16/22 (72%)               | Rab11-like, Rab32-like, Rab21 (2), |                  |
| TBC    | B, D, E, F, I, L, M, N, Q, RootA     | 10           | None                       | Rab11-like, Rab32-like, Rab21 (2),     |                  |
|        |                |              |                            | KSRX1, UzRX3, KSRX4                   |                  |
|        |                |              |                            |                                        | ExA (2)          |

Putative LECA subtypes of each protein family are presented along with those that were not identified in kinetoplastids. In the penultimate column, the proportion of these LECA proteins identified in kinetoplastids are listed. In the final column, proteins predicted to be in the last kinetoplastid common ancestor (LKCA) but not in the LECA are shown. These include proteins predicted to be derived from expansion of LECA proteins. Numbers in parentheses indicate the number of copies present.

Rab11 and RabL5 (Fukuda, 2011; Gabernet-Castello et al., 2013). This latter TBC is implicated in recycling of VAMP7 vesicles and endosome to Golgi transport. TBC-B is retained in all kinetoplastids as a single copy but lost in the phytomonads. TBC-D is stably retained across all eukaryotes except for the opportunistic parasite Trypanosoma. In the final column, proteins predicted to be in the last kinetoplastid common ancestor (LKCA) but not in the LECA are shown. These include proteins predicted to be derived from expansion of LECA proteins. Numbers in parentheses indicate the number of copies present.

**Conservation of VAMP7C interactions**

To examine VAMP7C interactions, we used immunosolation and mass spectrometry. In opisthokonts, VAMP7 forms two complexes, the first with SynE, Syx6 and VT11B, to mediate lysosomal transport, and the second with Syx6, Syx16 and VT11A for endosomal trafficking (Jahn and Scheller, 2006). In T. brucei, we consistently identified VT11-like A and B, Syx6-like, Syx6-like, Syx16B and SynE (Table 2), a complex composition that is also consistent with their localisation (Fig. 6E). Opisthokont VAMP7 is also implicated in plasma membrane trafficking, with Qa-SynPM and Qbc SNAP-25 or SNAP-23 (Jahn and Scheller, 2006) and in plants in complex with Qa-SynPM, QB-NPSN and Qc-Syp7 (Suwastika et al., 2008; Zheng et al., 2002). We reproducibly recovered TbNPSNA as a candidate VAMP7C interactor, albeit with poor emPAI support, but not TbSyp7, suggesting that a different T. brucei VAMP7 (TbVAMP7A or B) is required or that the SNARE complex composition for plasma membrane targeting is divergent. Therefore, endosomal interactions between VAMP7 and partners are apparently conserved with the opisthokont SNARE complexes.

**DISCUSSION**

The kinetoplastids encompass many parasitic species with a particularly broad range of niches and life cycles. Surface components are critical to their success, which is partly reflected within the trafficking system, and our analysis here provides a comprehensive view of these processes. The repertoire of Rabs, RabGAPs and SNAREs at the LCKA resembles that predicted for LECA, and the free-living B. saltans possesses the largest repertoire, with several kinetoplastid-specific members that were likely to originate in the LCKA. Significantly, we also find that the frequency of paralogous pairs is high in kinetoplastids, which may indicate a whole genome duplication event at the base of the kinetoplastid lineage. Rather than a precipitous decrease as
Kinetoplastids transitioned from free-living to parasitic forms; there is gradual loss of trafficking genes, and likely corresponding pathway simplification (Jackson et al., 2015). Whereas some LECA proteins (e.g. SNAP-25, Rab24, Rab32 and TBC-N) are lost, a large proportion of losses are of lineage-specific duplications originating in the last common excavate ancestor or in the LCKA. Hence, radical intracellular remodelling did not accompany the transition to parasitism. The extensive bodonid repertoire is mainly due to retention of the LECA/LKCA gene complement, and the most prominent expansions appear to be Syp7 and Rab32.

Leishmania and trypanosomes have both lost the endosomal Rab32.
Fig. 6. Sub-cellular localisation of *T. brucei* R-SNAREs. (A–D) Myc-tagged TbVAMP7A (A), HA-tagged TbVAMP7B (B), TbVAMP7C (C) and TbYkt6 (D) in procyclic *T. brucei* cells are shown. Localisation of each SNARE (red) is shown relative to markers (green) for the endoplasmic reticulum (TbBiP) (top), the lysosome (p67) (middle) and the Golgi complex (TbGRASP) (bottom). The nucleus and kinetoplast are stained with DAPI and pseudocoloured in blue. Tags were visualised with rat anti-Myc antibody or rat anti-HA antibody as appropriate. (E) Localisation of candidate TbVAMP7C interactors. HA-tagged VAMP7C (red), is shown relative to putative interactors tagged with Myc (top five rows) and V5 (lower two rows), in green. The nucleus and kinetoplast were stained with DAPI. VAMP7C was stained with rat anti-HA, SNARE interactors were stained with mouse anti-Myc antibody and non-SNARE interactors were stained with mouse anti-V5 antibody. Scale bar: 2 µm.
Table 2. Proteins identified as potential VAMP7C interactors

| Accession   | ID       | # of peptides | emPAI | Family      |
|-------------|----------|---------------|-------|-------------|
| Tb427.10.790| Handle-VAMP7C | 10              | 5.74  | SNARE       |
| Tb09.211.3920| Qa-Syx16B | 10              | 1.98  | SNARE       |
| Tb427.03.5570| Qa-SynE  | 7               | 1.27  | SNARE       |
| Tb427.10.1030| Qc2a-SyP7B | 6              | 1.25  | SNARE       |
| Tb427.08.3470| Qb2-VTI1-like A | 6           | 1.22  | SNARE       |
| Tb427.08.1120| Qb3-VTI1-like B | 2           | 0.87  | SNARE       |
| Tb427.10.2340| Qc3-Syx8-like | 2             | 0.67  | SNARE       |
| Tb427.10.6780| VPS45    | 10              | 0.67  | SM protein  |
| Tb427.10.9950| Qa-Syx18 | 4               | 0.32  | SNARE       |
| Tb427.10.14200| Qa-Syx5 | 3                | 0.28  | SNARE       |

Data are arranged by emPAI rank order. Bold indicates an interaction validated by co-immunofluorescence (Fig. 6). All experiments and identifications were confirmed in a total of four biological replicates.

Materials and Methods

Sequence data collection
To identify SNARE sequences, a validated dataset of 26 predicted SNARE sequences obtained for T. brucei (Murungi et al., 2014) and 27 sequences from L. major (Besteiro et al., 2006) were used to query predicted proteomes for Trypanosoma brucei brucei 927, T. b. gambiense, T. congolense, T. vivax, T. cruzi, T. grayi, Leishmania major, L. mexicana, L. braziliensis, L. infantum, Bodo saltans, Phytomonas serpens, Phytomonas EM1, Phytomonas HART1 (http://www.genedb.org/ or TriTrypDB) and the heterolobosid Naegleria gruberi (Fritz-Laylin et al., 2010) (http://genome.jgi.doe.gov/Naegr1) by BLAST. tBLASTp was used to search transcriptome data from T. theileri, T. carassii, Trypanoplasma borrelii and Euglena gracilis (S.K. and M.C.F., unpublished data). All sequences returned with e-values <10^{-3} were retained. This dataset was parsed to remove redundancies using >99% sequence identity as criteria for exclusion. ClustalW (Thompson et al., 2002) was used to align each dataset and generate neighbour-joining (NJ) trees (Saitou and Nei, 1987). Sequences which were excluded or weakly clustering were validated for a SNARE domain via Interpro (Hunter et al., 2009, http://www.ebi.ac.uk/interpro) and SNARE-DB (Kloepper et al., 2007, http://bioinformatics.mpibpc.mpg.de/snare/index.jsp). Sequences incorporating a SNARE-domain (SNARE, t/v-SNARE, sec20, syntaxin, longin and synaptobrevin) were retained. Sequences in which no domain was detected but were between 70 and 500 residues were also retained.

A pfam Ras domain (PF00071) HMMSCAN (Eddy, 1998) (http://hmmer.org) search was conducted against 50 eukaryotic proteomes, including those listed above to identify RabS. Sequences with e-values <10^{-3} were retained and used to generate a NJ tree. Kinetoplastid Rab candidates were identified and tentatively assigned on the basis of association with defined Rab sequences. One sequence from each cluster was used as a query to define the cluster by reciprocal best hit (rbh) BLAST and another round of rbhBLAST performed using sequences from the first round as queries. rhb requires that potential positive hits retrieve the original query when used to search. The tree was annotated accordingly to verify initial assignments. Additionally, assignments were made using Rabifier (Diekmann et al., 2011). All collected kinetoplastid sequences were classed as either a tentative Rab subfamily member or as stray. Identical procedures were undertaken with the TBC domain (pfam PF00566) and HMMSCAN to recover sequences.

For analyses addressing the evolution of NPSN and SyP7, a broader eukaryotic sampling and separate phylogenetic methods were used. Homology searches were performed using Homo sapiens, Saccharomyces cerevisiae and Arabidopsis thaliana sequences; protein sequences from a broad sampling of eukaryotes were retained. Positive BLAST hits for each protein of interest were aligned using MUSCLE v3.8.31 (Edgar, 2004). Where very similar sequences were present from the same or closely related organisms, those that aligned least well were removed to limit overrepresentation of some taxonomic groups in the Hidden Markov Model (HMM). For each potential orthologue identified by HMMER (e-value 0.05), reverse BLAST searches were performed as before. Homology search results were summarised using Coulson Plot Generator (Field et al., 2013).

Phylogenetic reconstruction
SNARE, Rab and RabGAP datasets were aligned using MAFFT (Katoh et al., 2005) using the E-INS-i strategy and manually edited in Jalview (Waterhouse et al., 2009). The alignment was used to generate maximum-likelihood trees in PhyML v3.0 (Guindon et al., 2010) using the LG model with the following parameters: number of substitution rates, 4; starting tree, BioNJ, tree topology search; NNI moves and statistics, aLRT SH-like and bootstrap (100 or 1000 replicates as indicated in the figures). Bayesian inference was implemented in MrBayes v3.2 on the CIPRES server (Miller et al., 2010; Ronquist and Huelsenbeck, 2003), generally with 8×10^6 Markov chain Monte Carlo (MCMC) generations where convergence was achieved, as measured by a splits frequency below 0.01. Substitution models employed for inferring trees were selected using ProtTest v3 (Abascal et al., 2005) for PhyML and the mixed model for MrBayes. One representative from each well supported clade (>0.9, 90, MrBayes, PhyML support respectively) along with a panel of known eukaryotic SNAREs, RabS and TBC Rab-GAPs were analysed to determine orthology and define subfamilies. A SNARE panel was created from SNARE sequences from Homo sapiens, Saccharomyces cerevisiae, Arabidopsis thaliana, Phytophthora sojae and Entamoeba histolytica obtained from SNARE-DB. For Qb and Qc, LECA complement datasets were assembled from the searches. Alignments generated for HMMs or phylogenetic analysis were constructed using MUSCLE v3.8.31 (Edgar, 2004) and edited using Mesquite (Maddison and Maddison, 2015). Maximum likelihood phylogenies were constructed using RAXML (Stamatakis, 2014) with the Protein GAMMA model for rate heterogeneity and the LG4X substitution matrix. Values from 100 bootstrap were mapped onto the best of 20 topologies constructed from the original alignment. Bayesian phylogenetic analysis was performed using MrBayes, and run for 8–10 million MCMC generations with the Mixed substitution model, and the final average standard deviation of split frequencies decreased to less than 0.008. Several trees were constructed iteratively for each dataset, with removal of manually selected sequences to improve resolution.

Trypanosoma cell culture
Procyclic culture form (PCF) T. brucei Lister 427 were grown as previously described (Brn et al., 1979). Expression of plasmid constructs was maintained using antibiotic selection at the following concentrations: G418 or hygromycinB at 25 μg/mL, blastidicin at 10 μg/mL and puromycin at 2 μg/mL.

Expression constructs
Putative trypanosome SNARES TbVAMP7C (Tb427.10.790), TbVAMP7B (Tb427.2.5120), TbVAMP7B (Tb427.5.3560), TbYkt6 (Tb927.9.14080) were amplified from T. b. brucei 427 genomic DNA using Hercules DNA polymerase (Aligent Technologies).

For hemagglutinin (HA)-tag fusion, the PCR products containing a coding sequence for C-terminal HA-epitope were cloned into the PCF expression vector pLew79 (Wirtz et al., 1999) using AvrII and BamHI (TbVAMP7 isoforms) using the following primers: 5′-TTGTGTCCTAGAGTGCTTTATATCTGCGCTCTTT-3′ pLewVAMP7A_F and ACTCAAGGATCCCTAAGCTGAATCCTGAAACATGTTGATCTCTTTGCTCTAGGTTTATCG-3′ pLewVAMP7A_R; 5′-TTGTGTCCTAGAGTGCTTTATATCTGCGCTCTTT-3′ pLewVAMP7B_F and 5′-ACTCAAGGATCCCTAAGCTGAATCCTGAAACATGTTGATCTCTTTGCTCTAGGTTTATCG-3′ pLewVAMP7B_R; 5′-TTGTGTCCTAGAGTGCTTTATATCTGCGCTCTTT-3′ pLewVAMP7C_F and 5′-ACTCAAGGATCCCTAAGCTGAATCCTGAAACATGTTGATCTCTTTGCTCTAGGTTTATCG-3′ pLewVAMP7C_R;
The PCR product of TbVkt6 was cloned into pHDI034 containing an N-terminal HA-epitope using HindIII and AmpliSeq using the following primers: 5′-GGCCGAAGTGTTTACTCTCCGGCACA-3′ and 5′-CCGTCTTATGTCATACGGTCGGCAACA-3′ pHDI034Ykt6F and 5′-CCGTCTTATGTCATACGGTCGGCAACA-3′ pHDI034Ykt6R.

Putative SNARE interactors TbSyx16B (Tb427.12.110.300.5570), TbSyx1E (Tb427.03.5570), TbVkt6-like A (Tb427.8.3470), TbVkt6-like B (Tb427.8.1120), TbSyx6-like 1 (Tb427.10.1830) and TbSyx6-like 2 (Tb427.10.2340) were also similarly amplified. For 6× Myc tagging, the product of each gene was cloned into pRPΔDOP (with thanks to Lucy Glover, Institut Pasteur, Paris, France) containing 6× cMyc-epitope using HindIII and XbaI.

Putative SNARE interactors using the following primers: 5′-GGCCGAAGTGTTTACTCTCCGGCACA-3′ and 5′-CCGTCTTATGTCATACGGTCGGCAACA-3′ pHDI034Ykt6F and 5′-CCGTCTTATGTCATACGGTCGGCAACA-3′ pHDI034Ykt6R.

Competing interests
The authors declare no competing or financial interests.

Author contributions
D.V., L.D.B., J.B.D. and M.C.F. conceived and designed the study. D.V., C.B., L.D.B. and N.N.N. performed the experiments and analyses, and generated figures and tables. S.K. provided data for T. theilerii, T. borenii and T. carassii. A.O.R. performed initial sequence collection. D.V., L.D.B., C.B., N.N.N., J.B.D. and M.C.F. wrote the manuscript. J.B.D., D.V. and M.C.F. edited the manuscript.

Funding
This work was supported by the Wellcome Trust (grant 082813 to M.C.F.) and the Gates Cambridge Trust (to D.V.), a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery grant (to J.B.D.), a NSERC PGSD (to L.D.B.), and an Alberta Innovates Health Solutions summer studentship (to N.N.N.). M.C.F. is a Wellcome Trust Investigator. Deposited in PMC for immediate release.

Supplementary information
Supplementary information available online at http://jcs.biologists.org/lookup/doi/10.1242/jcs.197640.supplemental

References
Abascal, F., Zardoya, R. and Posada, D. (2005). ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21, 2104-2105.
Adua, V. O., Gadela, C. and Field, M. C. (2013). Proteomic analysis of clathrin interactions in trypanosomes reveals dynamic evolution of endocytosis. Traffic 14, 440-457.
Ali, M., Leung, K. F. and Field, M. C. (2014). The ancient small GTPase Rab21 functions in intermediate endocytic steps in trypanosomes. Eukaryot. Cell 13, 304-319.
Arasaki, K., Shimizu, H., Mogari, H., Nishida, N., Hirota, N., Furuno, A., Kudo, Y., Baba, M., Baba, N., Cheng, J. et al. (2015). A role for the ancient SNARE Ufe1p and Sec20p. Dev. Cell 35, 49-59.
Balslevsen, W., Ossipov, D. and Schmitt, H. (1998). Recycling of the yeast v-SNARE Sec22p involves COPI-proteins and the ER transmembrane proteins Ufe1p and Sec20p. J. Cell Sci. 111, 1507-1520.
Bangs, J. D., Brouch, E. M., Ransom, D. M. and Roggy, J. L. (1996). A soluble secretory reporter system in Trypanosoma brucei. Studies on endoplasmic reticulum targeting. J. Biol. Chem. 271, 18387-18393.
Besteiro, S., Coombs, G. H. and Mottram, J. C. (2006). The SNARE protein family of Leishmania major. BMC Genomics 7, 250.
Hardwick, K. G. and Pelham, H. R. (1992). SEDS encodes a 39-kD integral membrane protein required for vesicular transport between the ER and the Golgi complex. J. Cell Biol. 119, 513-521.

Hirota, Y. and Tanaka, Y. (2009). A small GTPase, human Rab32, is required for the formation of autophagic vacuoles under basal conditions. Cell. Mol. Life Sci. 66, 2913-2932.

Huber, L. A., Pimplikar, S., Parton, R. G., Virth, H., Zerial, M. and Simons, K. (1993). Rab8, a small GTPase involved in vesicular traffic between the TGN and the basolateral plasma membrane. J. Cell Biol. 123, 35-45.

Hunte, S., Apweiler, R., Atwood, T. K., Bairoch, A., Bateman, A., Binns, D., Bork, P., Das, U., Daugherty, L., Duquehne, L. et al. (2009). InterPro: the integrative protein signature database. Nucleic Acids Res. 37, D211-D215.

Itakura, E. and Mizushima, N. (2013). Syntaxin 17: the autophagosomal SNARE. Autophagy 9, 917-919.

Itakura, E., Kishi-Itakura, C. and Mizushima, N. (2012). The hairpin-type tail-anchored SNARE syntaxin 17 targets to autophagosomes for fusion with endosomes/liposomes. Cell 151, 1256-1269.

Jackson, A. P. (2016). Gene family phylogeny and the evolution of parasite cell surfaces. Mol. Biochem. Parasitol. 209, 64-75.

Johansen, D. T., Otto, T. D. (2004). Rab11 resides in a vesicular compartment distinct from Rab11a in intestinal cells and other epithelial cells. Exp. Cell Res. 298, 322-331.

Leung, K. F., Dacks, J. B. and Field, M. C. (2008). Evolution of the multivesicular body ESCRT machinery; retention across the eukaryotic lineage. Traffic 9, 1607-1616.

Lumb, J. H., Leung, K. F., DuBois, K. N. and Field, M. C. (2011). Rab28 function in trypanosomes: interactions with retromer and ESCRT pathways. J. Cell Sci. 124, 3771-3783.

Maddison, W. P. and Maddison, D. R. (2015). Mesquite: a modular system for evolutionary systematics. Version 2.75, 2011, http://mesquiteproject.org.

Manna, P. T., Kelly, S. and Field, M. C. (2013). Adaptin evolution in kinetoplastids and emergence of the variant surface glycoprotein coat in African trypanosomatids. Mol. Phylogenet. Evol. 67, 123-128.

Miller, M. A., Pfeffer, W. and Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In Proceedings of the Gateway Computing Environments Workshop (GCE), 14 Nov. 2010, New Orleans, LA pp 1-8.

Murungi, E., Barlow, L. D., Venkatesh, D., Adung’a, V. O., Dacks, J. B., Field, M. C. and Christoffels, A. (2014). A comparative analysis of trypanosomatid SNAREs. Parasitol. Int. 63, 341-348.

Natesan, S. K. A., Peacock, I. C., Leung, K. F., Matthews, K. R., Gibson, W. and Field, M. C. (2009). The trypanosome rab-related proteins RabX1 and RabX2 play no role in intracellular trafficking but may be involved in fly infectivity. PLoS ONE 4, e7217.

Nishida, A. P., Shim, J. and Ferro-Novick, S. (1990). BET1, BOS1, and SEC22 are members of a group interacting yeast genes required for transport from the endoplasmic reticulum to the Golgi complex. Mol. Cell. Biol. 10, 3405-3414.

Niyogi, S., Mucci, J., Campetella, O. and Docampo, R. (2014). Rab11 regulates trafficking of trans-sialidase to the plasma membrane through the contractile vacuole complex of Trypanosoma cruzi. PLoS Pathog. 10, e1004224.

Niyogi, S., Jimenez, V., Girard-Dias, W., de Souza, W., Miranda, K. and Docampo, R. (2015). Rab32 is essential for maintaining functional acidocatalysomes, and for growth and infectivity of Trypanosoma cruzi. J. Cell Sci. 128, 2363-2373.

Overholzer, M., Morgen, S., Kunz, S. and Seebec, T. (2006). A vector series for rapid PCR-mediated C-terminal in situ tagging of Trypanosoma brucei genes. Mol. Biochem. Parasitol. 145, 117-120.
Olkkonen, V. M. and Ikonen, E. (2006). When intracellular logistics fails—genetic defects in membrane trafficking. J. Cell Sci. 119, 5031-5045.

Ortiz-Sandoval, C. G., Hughes, S. C., Dacks, J. B. and Simmen, T. (2014). Interaction with the effector dynamin-related protein 1 (Drp1) is an ancient function of Rab32 subfamily proteins. Cell. Logist. 4, e096399.

Pal, A., Hall, B. S., Nesbeth, D. N., Field, H. I. and Field, M. C. (2002). Differential endocytic functions of Trypanosoma brucei Rab5 isoforms reveal a glycosylphosphatidylinositol-specific endosomal pathway. J. Biol. Chem. 277, 9529-9539.

Pereira-Leal, J. B. (2008). The Ypt/Rab family and the evolution of trafficking in fungi. Traffic 9, 27-38.

Pereira-Leal, J. B. and Seabra, M. C. (2001). Evolution of the Rab family of small GTP-binding proteins. J. Mol. Biol. 313, 889-901.

Petřklová, R. and Eliaš, M. (2014). Contrasting patterns in the evolution of the Rab GTPase family in Archaeplastida. Acta Soc. Bot. Pol. 83, 303-315.

Rajendran, L. and Annaert, W. (2007). A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in Trypanosoma brucei. Mol. Biochem. Parasitol. 99, 89-101.

Rendl, J., Helkova, R. and Eliaš, M. (2014). Contrasting patterns in the evolution of the Ypt/Rab family and the evolution of trafficking in Archaeplastida. J. Mol. Biol. 313, 889-901.

Sanderfoot, A. (2007). Increases in the number of SNARE genes parallels the rise of multicellularity among the green plants. Plant Physiol. 144, 6-17.

Sanderfoot, A. A., Assaad, F. F. and Raikhel, N. V. (2000). The Arabidopsis genome. An abundance of soluble N-ethylmaleimide-sensitive factor adaptor protein receptors. Plant Physiol. 124, 1558-1569.

Schlacht, A., Herman, E. K., Klute, M. J., Field, M. C. and Dacks, J. B. (2014). Missing pieces of an ancient puzzle: evolution of the eukaryotic membrane-trafficking system. Cold Spring Harb. Perspect. Biol. 6, a016048.

Schwartz, K. J., Peck, R. F. and Bangs, J. D. (2013). Intracellular trafficking and glycobiology of TbsPDI2, a stage-specific protein disulfide isomerase in Trypanosoma brucei. Eukaryot. Cell 12, 132-141.

Seixas, E., Barros, M., Seabra, M. C. and Barral, D. C. (2013). Rab and Arf proteins in genetic diseases. Traffic 14, 871-885.

Sevova, E. S. and Bangs, J. D. (2009). Streamlined architecture and glycosylphosphatidylinositol-dependent trafficking in the early secretory pathway of African trypanosomes. Mol. Biol. Cell 20, 4739-4750.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312-1313.

Suwastika, I. N., Uemura, T., Shina, T., Sato, M. H. and Takeyasu, K. (2008). SYP71, a plant-specific Qc-SNARE protein, reveals dual localization to the plasma membrane and the endoplasmic reticulum in Arabidopsis. J. Cell Sci. 121, 1104-1118.

Thompson, J. D., Gibson, T. J. and Higgins, D. G. (2002). Multiple sequence alignment using ClustalW and ClustalX. Curr. Protoc. Bioinforma. Chapter 2, Unit 2.3.

Vedovato, M., Rossi, V., Dacks, J. B. and Filippini, F. (2009). Comparative analysis of plant genomes allows the definition of the “PhytoHs”: a novel non-SNARE longin domain protein family. BMC Genomics 10, 510.

Wang, Y. and Seemann, J. (2011). Golgi biogenesis. Cold Spring Harb. Perspect. Biol. 3, 1-12.

Waterhouse, A. M., Procter, J. B., Martin, D. A. M., Clamp, M. and Barton, G. J. (2009). Jalview Version 2-A multiple sequence alignment editor and analysis workbench. Bioinformatics 25, 1189-1191.

Wirtz, E., Leal, S., Ochatt, C. and Cross, G. A. M. (1999). A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in Trypanosoma brucei. Mol. Biochem. Parasitol. 99, 89-101.

Zheng, H., Bednarek, S. Y., Sanderfoot, A. A., Alonso, J., Ecker, J. R. and Raikhel, N. V. (2002). NPSN11 is a cell plate-associated SNARE protein that interacts with the syntaxin KNOLLE. Plant Physiol. 129, 530-539.