

# Evolution of the Multivesicular Body ESCRT Machinery; Retention Across the Eukaryotic Lineage

Ka Fai Leung, Joel B. Dacks and Mark C. Field\*

The Molteno Building, Department of Pathology,  
University of Cambridge, Tennis Court Road, Cambridge,  
CB2 1QP, UK

\*Corresponding author: Mark C. Field, mcf34@cam.ac.uk

**Lysosomal targeting of ubiquitylated endocytic cargo is mediated in part by the endosomal sorting complex required for transport (ESCRT) complexes, a system conserved between animals and fungi (Opisthokonta). Extensive comparative genomic analysis demonstrates that ESCRT factors are well conserved across the eukaryotic lineage and complexes I, II, III and III-associated are almost completely retained, indicating an early evolutionary origin. The conspicuous exception is ESCRT 0, which functions in recognition of ubiquitylated cargo, and is restricted to the Opisthokonta, suggesting that a distinct mechanism likely operates in the vast majority of eukaryotic organisms. Additional analysis suggests that ESCRT III and ESCRT III-associated components evolved through a concerted model. Functional conservation of the ESCRT system is confirmed by direct study in trypanosomes. Despite extreme sequence divergence, epitope-tagged ESCRT factors TbVps23 and TbVps28 localize to the endosomal pathway, placing the trypanosome multivesicular body (MVB) in juxtaposition to the early endosome and lysosome. Knockdown of TbVps23 partially prevents degradation of an ubiquitylated endocytosed transmembrane domain protein. Therefore, despite the absence of an ESCRT 0 complex, the trypanosome ESCRT/MVB system functions similarly to that of opisthokonts. Thus the ESCRT system is an ancient and well-conserved feature of eukaryotic cells but with key differences between diverse lineages.**

**Key words:** endocytosis, ESCRT, multivesicular body, trypanosomes, ubiquitin, vesicle transport

**Received 29 November 2007, revised and accepted for publication 10 July 2008, uncorrected manuscript published online 14 July 2008, published online 26 August 2008**

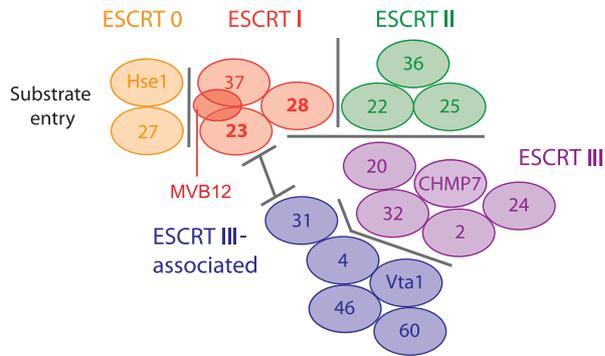
In animals and fungi (the Opisthokonta supergroup), ubiquitylation plays a vital role in downregulation and endocytosis of a considerable proportion of cell surface transmembrane domain (TMD) proteins, including the somatotrophin and epidermal growth factor receptors (1,2). Covalent attachment of ubiquitin to lysine side chains within the cytoplasmic tail of TMD proteins involves engagement of specialized E3 ubiquitin ligases, c-Cbl and Rsp5p, that appear to have specific functions in ubiquitin

modification of cell surface proteins (3,4). Subsequent internalization and recognition of ubiquitylated cargo are achieved, at least in part, by ubiquitin-interacting motifs (UIMs) present in a range of proteins that participate in early endocytic events, including epsin and eps15 (5,6).

Internalized ubiquitylated molecules are delivered to the multivesicular body (MVB), a comparatively late endocytic compartment, where inward budding of membrane creates vesicles entirely encompassed within the endosomal membrane (7). This process requires sequential action of several protein complexes referred to as endosomal sorting complex required for transport or ESCRTs (8) (Figure 1). All ESCRT components are class E vacuolar protein sorting (Vps) genes in *Saccharomyces cerevisiae* and generate an enlarged prevacuolar compartment in knockout cells. The class E compartment accumulates proteins normally destined for the vacuole and results in aberrant secretion of carboxypeptidase Y (9,10). In mammalian cells, the MVB is similarly responsible for sorting of lysosomally targeted proteins and antigen presentation while also contributing a site exploited by many viruses for virion assembly (11). Interestingly, while none of the ESCRT factors appear to be required for viability in yeast, in mammals, several ESCRT components are essential, highlighting potential evolutionary differences between yeast and metazoa (12,13).

Considerable effort has been devoted to generating interactome maps for the ESCRTs, while structural studies have elucidated the detailed conformation of several of the complexes [(14,15), reviewed in 16]. Overall, there is a high degree of subunit and functional conservation between metazoa and yeast. Furthermore, very recent data suggest the presence of, at least, two subunits of the ESCRT III/ESCRT III-associated system in Archaea. The presence of prokaryotic ESCRT precursors points to a possible origin for the system (17).

Current data favor a hierarchical assembly of ESCRT complexes, termed the concentric ring hypothesis, based around an initial targeting event involving recognition of phosphatidylinositol 3-phosphate (PI3P) by the FYVE domain-containing factor Vps27p/Hrs, a component of the ESCRT 0 complex (6,18). Evidence also suggests that ESCRT 0 recognizes ubiquitylated cargo and recruits ESCRT I to the nascent MVB membrane (Figure 1) (6). Subsequent recruitment of ESCRT II and III is associated with formation of inward-budding membranes. While the precise mechanism of membrane deformation remains unknown, multiple copies of ESCRT III factors are believed



**Figure 1: Schematic representation of higher eukaryote ESCRT complex components.** The individual complexes are color coded as indicated, and each lozenge represents a single factor. Multiple copies of the same factor are not represented for simplicity, and the factors are named using the *S. cerevisiae* nomenclature omitting Vps. Major interaction faces between the complexes are indicated with gray lines, and the general arrangement is based on yeast two-hybrid, X-ray crystallography, glutathione S-transferase pull down and additional data from many studies but is simplified for clarity. The ubiquitylated endocytic substrate entry point is indicated, and the two subunits localized in trypanosomes (Vps23 and Vps28) are shown in bold. Note that in mammalian cells, Vps23 is referred to as TSG101, Vps31 as Alix, Hse1 as STAM and Vps27 as Hrs.

to form a lattice on the endosomal membrane [(18) and references therein]. The ESCRT III-associated complex also appears to play a role in the latter steps of this pathway, culminating in the removal of ubiquitin from cargo mediated by the deubiquitylating enzyme Doa4p in yeast (17,19–21) and by ubiquitin-specific processing protease Y (UBPY) and associated molecule with the SH3 domain of STAM (AMSH) in mammals (16). Deubiquitylated cargo is then delivered into the inward-budding vesicular structures. In mammals, ESCRT complexes are disassembled by the activity of the AAA-type adenosine triphosphatase (ATPase) Vps4 (18).

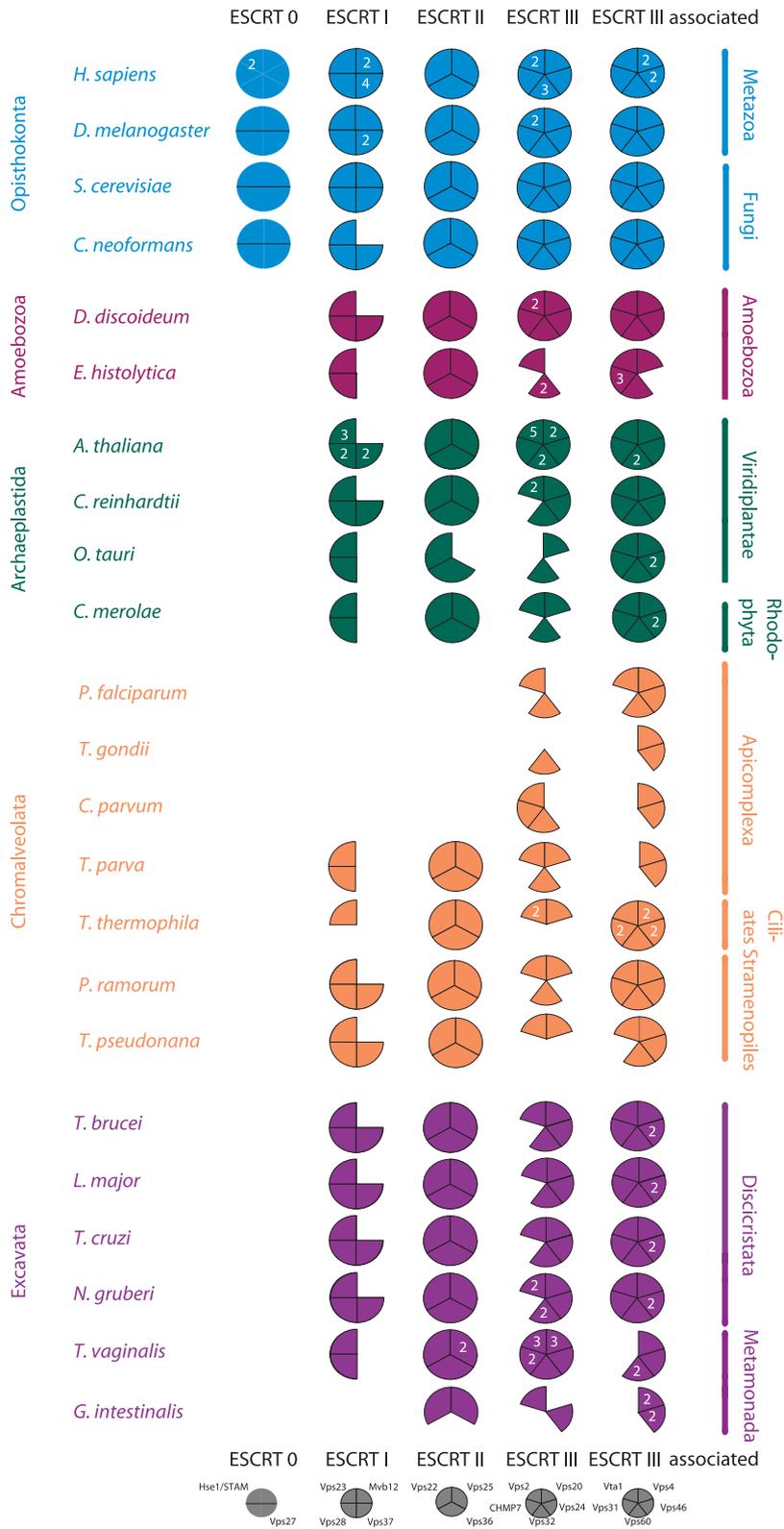
Comparative genomic and phylogenetic studies have determined that the basic features of the intracellular trafficking system were established very early in eukaryote evolution [(22) and references therein]. Much of the molecular machinery was present in a comparatively modern configuration prior to radiation of the major eukaryotic lineages. Specifically, Rab and SNARE proteins, and major subfamilies of these factors, are present across the eukaryotes, indicating probable establishment at least as early as the last common eukaryotic ancestor (LCEA). The basic vesicle coat systems, including coatomers, adaptins and clathrin, are also ancient, as are the tether complexes co-ordinating vesicle fusion (22,23). A more detailed analysis of the endocytic system not only confirms this ancient cellular complexity but also reveals significant evolutionary plasticity (24). Several lineage-specific features are observed, including opisthokont-specific factors such as caveolin, Golgi-localized, gamma-ear-containing,

ADP-ribosylation factor-binding proteins (GGAs), stonins, certain features of the adaptin complexes and lineage-specific expansions of the Rab5 subfamily, suggesting needs for specialized functions in higher eukaryotes (24,25). In contrast, there is ample evidence for secondary losses, including entire adaptin complexes and the endocytic factor Rab4 (26,27). This earlier work also suggested the conservation of ESCRT factors, while a further study has specifically highlighted the distribution of Vps25 (24,28). Similarly, evidence for MVB-like organelles in diverse eukaryotes has been reported (29–32). The best characterized thus far have been among apicomplexans using an ATPase-defective mutant of Vps4, while the apparent absence of several ESCRT components from the lineage was also noted (29).

Experimental examination of components in evolutionarily distant eukaryotes is important for at least two reasons; functional data are essential (i) for validation of homology where sequence similarity is equivocal and (ii) to provide insight into evolutionarily conserved functions. The African trypanosome, *Trypanosoma brucei*, is a unicellular protozoan parasite and, as a member of the Excavata, evolutionarily distant from metazoan and fungal organisms (33). Because of experimental tractability and medical importance, the trypanosome trafficking system is comparatively well characterized (26,34). Interest also stems from some remarkable features, including the predominance of glycosyl-phosphatidylinositol-anchored proteins and the presence of a specialized flagellar pocket responsible for all membrane traffic to and from the plasma membrane (35). Recently, we described structures morphologically related to MVBs and ubiquitin-dependent internalization and degradation of trypanosome TMD proteins (30,36,37). The ubiquitylation pathway appears biochemically similar to higher eukaryotes but with the distinction that genes for E3 ubiquitin ligases related to Rsp5 or c-Cbl are absent from all non-opisthokont lineages (37).

The absence of identifiable endocytic E3 ligases in trypanosomes, and potentially other factors from these and additional lineages (24,29), prompts a detailed analysis of the ESCRT/MVB machinery across the full eukaryotic range. Additionally, despite clear conservation between mammalian and yeast ESCRTs, both taxa are opisthokonts (33) and hence represent a narrow sampling. Therefore, the complete pattern of eukaryote ESCRT diversity is unexplored.

In this study, a comprehensive comparative genomic analysis of ESCRT components was undertaken in conjunction with direct functional investigation of trypanosome Vps23 and Vps28 orthologues, candidate components of the trypanosome ESCRT complex. Our data indicate a remarkable level of conservation of the ESCRT system but some significant lineage-specific aspects as well.



**Figure 2: Distribution of ESCRT components across the range of eukaryotic taxa.** Data are based on BLAST results together with alignments and phylogenetic analysis and are represented as a Coulson plot (see *Materials and Methods*). Large taxon groupings are color coded, and a key defining the factors represented by each sector is given at the bottom in gray. Filled sector, an identification based on a clear reverse BLAST result and/or additional evidence through analysis of the sequence by CLUSTAL or phylogeny. Open sector, not found. White numbers in some sectors indicate that multiple copies of these factors were found in these genomes and are the products of distinct genes. Accession numbers are provided in Table S1.

## Results

### **An ancient origin for the ESCRT system**

Using comparative genomics, we evaluated the presence of ESCRT complexes 0, I, II, III and the ESCRT III-associated factor genes across a broad range of available eukaryote genomes (Figure 2). We included four opisthokonts, two amoebozoans, four archaeplastidans, seven chromalveolates and finally six excavates (Figure 2); importantly, all sampled genomes are completed, and together, these taxa provide fair coverage of eukaryotic diversity. Each genome was parsed by a combination of homology searches and phylogenetic analyses, enabling classification of identified ESCRT components (Figures 2 and 3, Tables S1 and S2).

From the overall representation of ESCRT coding sequences, the most dramatic finding was that ESCRT I, II, III and ESCRT III-associated complexes are highly conserved across the five supergroups (Figure 2). Therefore, a near-universal system for trafficking of cargo through the MVB system, including internal budding of membranes and recognition of ubiquitylated TMD proteins is probably present throughout the eukaryotic lineage. This taxonomic distribution is most simply interpreted as the result of an origin for the ESCRT complexes predating speciation of the eukaryotic supergroups from the LCEA and implies the presence of an ESCRT system of near-modern configuration in the ancestor. Significantly, the AAA-type ATPase, Vps4, responsible for recycling ESCRT subunits from the MVB was retained by all taxa, indicating a highly conserved mechanism for delivering energy to the system and is consistent with recent evidence for an archaeal origin for Vps4 (17). Interestingly, CHMP7, a component of the ESCRT III complex (63), is found in genomes from plant, amoebozoan and excavate taxa. Although the distribution is sparse, this suggests that CHMP7 is a more general feature of ESCRT machinery than previously assumed and was possibly present as part of the ESCRT machinery in the LCEA. Overall, our data confirm previous work that demonstrate a conserved ESCRT system (8,24,28), but dramatically extend the scope of that conservation.

In addition to a trend of ESCRT complex retention, we observed expansion of genes encoding several subunits. *Homo sapiens* provides the only example of an expanded ESCRT 0 subunit, signal transducing adaptor molecule (STAM) (Hse1 in *S. cerevisiae*). Expansion of ESCRT I subunit copy number, while embracing all four subunits, is restricted to multicellular opisthokonts and plants. It is worth noting that although Mvb12 was found only in humans, yeast and *Drosophila* in this analysis, the high divergence between human and yeast Mvb12, together with its small size, may result in lack of detection by basic local alignment search tool (BLAST) analysis. Thus, we cannot rule out its presence in additional taxa. The sole example of expansion in ESCRT II is in *Trichomonas*

*vaginalis*, an organism with lineage-specific amplification of endocytic machinery factors (38). ESCRT III subunit expansion is more widespread and observed in unicellular and multicellular species. The most frequent ESCRT III-associated multicopy gene is Vps46. However, it is not clear if the LCEA contained multiple Vps46 genes or if this represents an additional example of lineage-specific amplification of endocytic factors (25).

In contrast to the conservation of ESCRT I, II, III and ESCRT III-associated complexes, ESCRT 0 is restricted to the Opisthokonta. The absence of ESCRT 0 from some non-opisthokonts has been suggested previously (24,29), but the present analysis provides exhaustive evidence for this specificity and moreover effectively rules out secondary loss as an explanation. We validated these observations by BLAST searches of 15 additional complete, or near-complete, genomes/expressed sequence tag (EST) databases and also of the entire non-redundant sequence database, restricted to non-opisthokont eukaryotes (see *Materials and Methods*). While VHS domain-containing factors were recovered, there was no evidence for ESCRT 0 subunits from any of these searches.

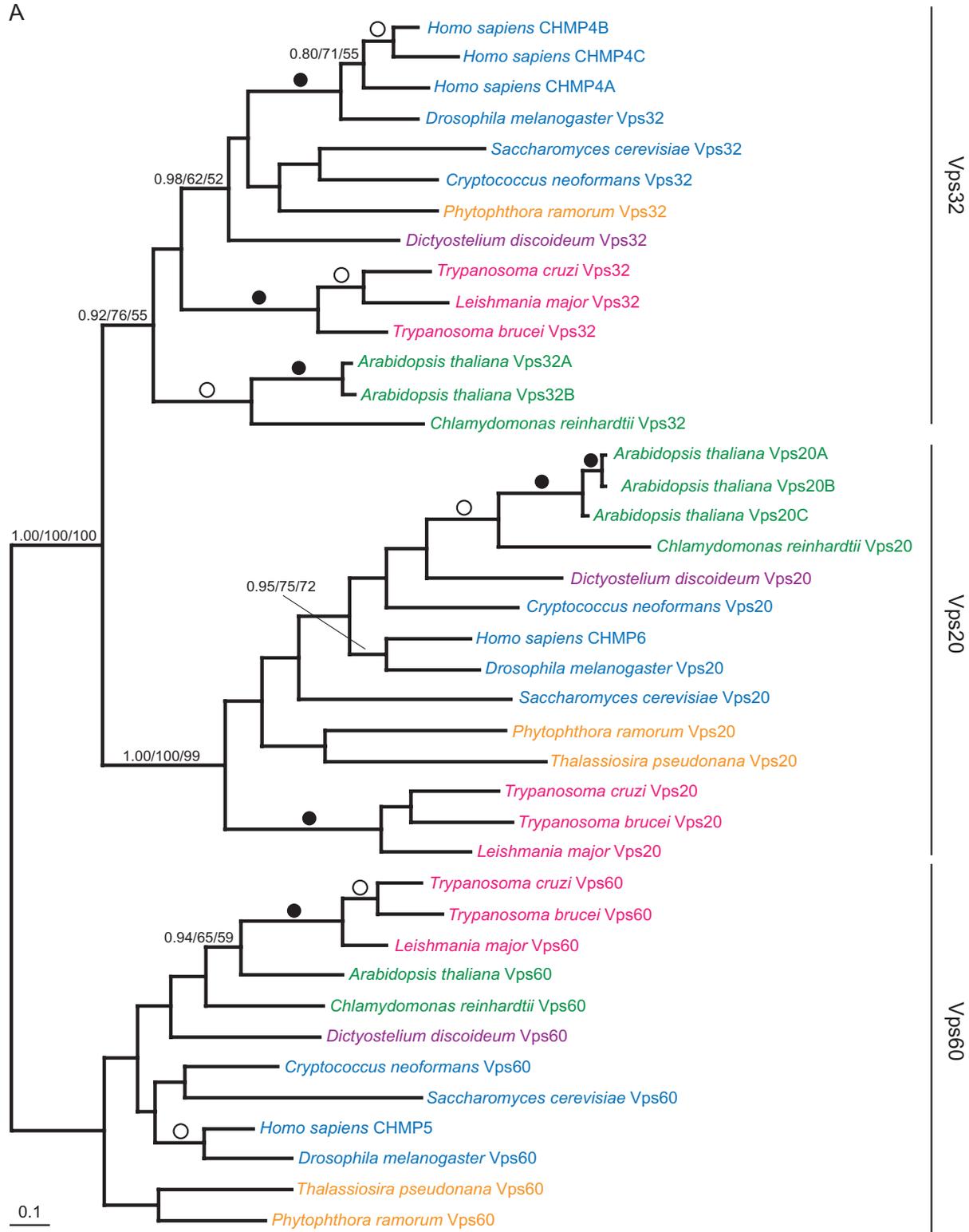
ESCRT 0 provides both the interface for substrate entry and a role in initial targeting of the ESCRT machinery to endosomal membranes. Hence, this finding has a major impact on views of ESCRT function (6,18), implying that the majority of eukaryotes utilize a mechanism distinct from that in the Opisthokonta for initial contact with ubiquitylated cargo and the endosomal membrane. Furthermore, the LCEA could not have used an ESCRT 0-dependent mechanism.

### **Lineage/lifestyle-specific patterns of ESCRT representation**

While substantial conservation of the ESCRT components across eukaryotic evolution is clear (Figure 4A), some secondary loss can be confidently predicted. Absence of individual subunits may be the result of failures in detection. For *Giardia* in particular, we and others find that the extreme evolutionary divergence of this organism frequently results in failure to recover candidate orthologues (24,39,40). This may account for a number of *Giardia* absences, but complete failure to detect ESCRT I-encoding sequences is likely significant and suggests minimized MVB machinery. Evidence for an unusual endocytic system in *Giardia* has been presented (41). CHMP7 homologues are also likely to have been lost on multiple occasions.

The most significant absences are in three apicomplexan taxa: *Plasmodium falciparum*, *Toxoplasma gondii* and *Cryptosporidium parvum*. Confirming and extending earlier work (24,29), we find that all three genomes lack genes encoding ESCRT I and II subunits; several of the subunits from ESCRT III and ESCRT III-associated complexes were also not identified. Given the apparent absence of entire complexes from several related organisms, we consider these

A



**Figure 3: Phylogenetic reconstruction of Snf7 and Vps24 domain ESCRT families.** Bayesian phylogenetic trees for Vps20/32/60 (A) and Vps2/24/46 (B). Numerical values (x/y/z) indicate statistical support for MRBAYES/PHYML/RAxML (posterior probability/bootstrap/bootstrap, respectively) for all nodes with support better than 0.80/50%/50%. Nodes of interest are in bold. Values for highly supported nodes have been replaced by symbols as follows: closed circles denote nodes with 0.99/99%/99%, while open circles denote nodes supported by 0.95/80%/80%. Species names are colored for supergroup and correspond to those in Figure 2. Figure 3 continued on next page.

B

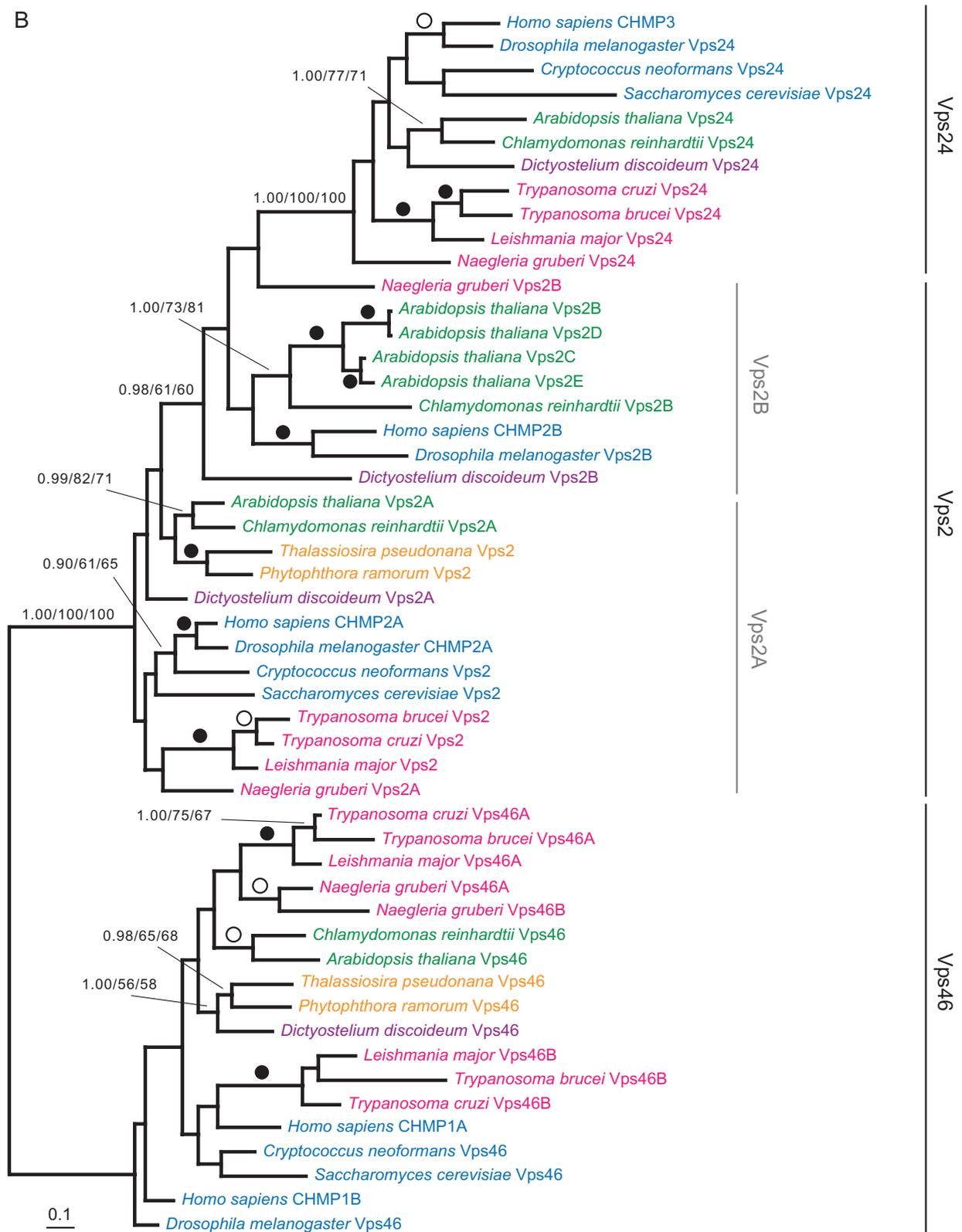
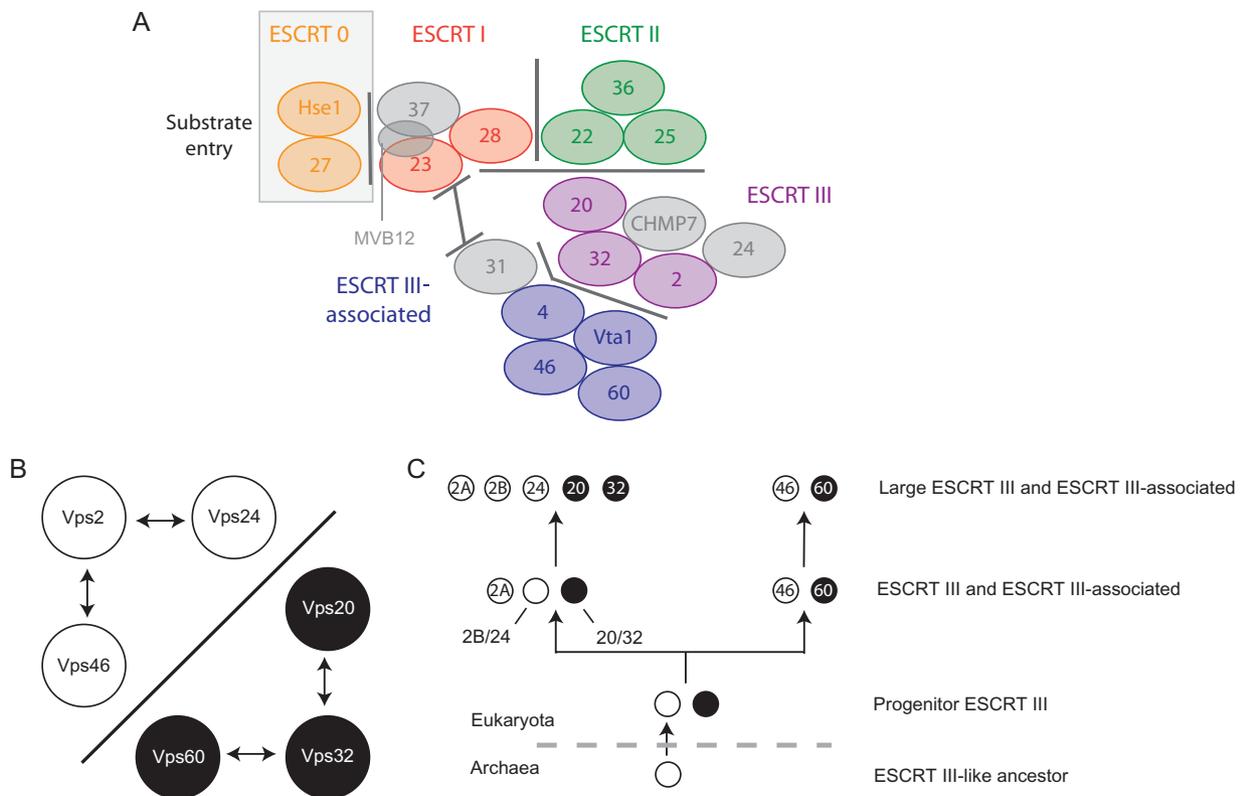


Figure 3: Continued from previous page.



**Figure 4: Model for evolution of ESCRT complexes.** A) Schematic of ESCRT complexes but with subunits absent in several taxa shown in gray and remaining conserved subunits shown using the color scheme of Figure 1. Overall, there is a remarkable level of conservation. The most prominent evolutionary variation is restriction of ESCRT 0 to the Opisthokonta, but critically, it appears that core interactions between the subunits are almost completely conserved, suggesting an evolutionary highly stable configuration. B) Sequence relatedness and symmetry between ESCRT III and ESCRT III-associated subunits. BLAST analysis indicates significant and bidirectionally detectable sequence similarity between Vps2, Vps24 and Vps46 (white) and Vps20, Vps32 and Vps60 (black). Note that in each case, one factor is related to the two others as detected by PSI-BLAST (data not shown), but no interfamily similarity is detected, suggesting independent origins for Vps2/24/46 and Vps20/32/60, consistent with distinct architectures revealed by analysis for coiled-coil domains (Figure S1). C) Model for evolution of ESCRT III and ESCRT III-associated subunits. A progenitor ESCRT complex containing one member each of Vps2/24/46 and Vps20/32/60 (white and black) gave rise by sequential gene duplications to distinct ESCRT III and ESCRT III-associated subunits. The proposed root is supported by analysis of an archaeal homologue of Vps2/24/46. Subsequent duplication of the ESCRT III subunits generated the configuration seen in most eukaryotes.

observations likely to be real and not sampling artifacts. However, the distribution is not easily explained as subunits for ESCRT I and II are recovered from a fourth apicomplexan, *Theileria parva*, and non-apicomplexan chromalveolates. Additional searches, using ESCRT I and II sequences from both *S. cerevisiae* and *T. parva*, against all *Plasmodium* species at NCBI failed to identify additional ESCRT I or II complex candidates within these taxa. By contrast, interrogation of the *Theileria annulata* genome successfully identified orthologues of Vps23p (ESCRT I) and all ESCRT II subunits (data not shown), strengthening this finding.

**Evolution of ESCRT III and ESCRT III-associated subunits**

Position-specific iterated (PSI)-BLAST analyses were performed to probe for deeper sequence relationships

between members of the ESCRT complexes. With two exceptions, no homology was detected between any of the ESCRT factors. However, relationships between a subset of the ESCRT III and ESCRT III-associated subunits were found encompassing two families, each comprised of three subunits; one family shares an Snf7 domain (Vps20, Vps32 and Vps60) and the second a Vps24 domain (Vps2, Vps24 and Vps46) (Figure 4B). These relationships were confirmed by sequence similarity and secondary structure prediction. Both gene families encode proteins of similar molecular weights (~25 kDa) and are predicted to contain extensive coiled-coil regions (Figure S1). Interestingly, the N-terminal glycine of Vps20p is highly conserved in orthologues throughout the supergroups, suggesting that N-myristoylation plays an important role when Vps20 is present. No evidence was found that the Snf7 and Vps24 families are homologous to one another. These data indicate that the Snf7 and Vps24 families are

likely the only ESCRT factors where common ancestry is present within or between the complexes.

Interestingly, within these protein families, extensive gene amplification was detected, for example expansion of Vps20, 32 and 60 in several taxa (Figure 2 and Table S1). Most frequent are expansions in copy numbers of Vps2 (ESCRT III) and Vps46 (ESCRT III-associated) genes, evident in taxa from multiple supergroups. Significantly, these subunits are related at the sequence level (Figure 3B). Vps2p and Vps46p are small proteins and likely present in more than one copy per ESCRT complex (18). However, the third member of this family, Vps24, is not subject to gene expansion but instead appears to have been lost on multiple occasions (Figure 2).

Phylogenetic analysis was used to investigate evolution of the various paralogous ESCRT components (see *Materials and Methods* and *Supporting Information* for full details). Analysis of Snf7 sequences containing Vps20, Vps32 and Vps60 from all taxa in Figure 2 provided moderate resolution, separating Vps60 from Vps20 and Vps32 and also verifying the assignment of these sequences as Vps60 orthologues (Figure S2). Further analysis following removal of long-branch/highly divergent sequences provided robust support for monophyly of Vps60, Vps32 and Vps20 with representatives of all five supergroups present in each clade (Figure 3A) and confirms that the duplications giving rise to these three protein families predate their ancestor.

Similarly, analysis of Vps60, Vps32, Vps20 and CHMP7 sequences was also performed. CHMP7 consists of two domains, each with homology to SNF7, and so the N-terminal and C-terminal domains were analyzed individually. Preliminary analysis suggested that, despite CHMP7 homology, the *T. vaginalis* genes may have a distinct evolutionary history (data not shown) and were therefore excluded from further analysis. The remaining N-terminal CHMP7 domains clustered with good support (Figure S3). Additional analysis with only CHMP7 C-terminal domains and SNF7 recovered a cluster of the C-terminal CHMP7 domains. However, support for C-terminal domain monophyly was obtained for Bayesian analysis only, with other methods showing no support overall (data not shown). This is highly suggestive of a single origin for CHMP7 predating the LCEA, but the sparse identification of CHMP7 and equivocal phylogenetic relationships means that conclusive reconstruction of CHMP7 evolution cannot be obtained with presently available data or methods.

Vps24 domain-containing protein sequences were subjected to phylogenetic analysis, allowing classification of orthologues of all three families (Figure S4, Table S2, data not shown). This analysis recovered monophyly for Vps46 and Vps24 but indicated possible paraphyly of the Vps2 family (Figure S4), suggesting that the duplications giving rise to the Vps24 domain-containing proteins predates LCEA. A smaller data set with divergent sequences

removed confirmed this result (Figure 3B). The root leading to Vps46 was placed within the Vps2 clade in the best topology, and a clade of Vps2B sequences plus Vps24 was moderately well supported. This implies that Vps24 is actually a subset of a larger Vps2 family, indicating the presence of an additional gene duplication in the Vps2/24/46 family of ESCRT III factors compared with the analogous Snf7 family. A further analysis, including several archaeal homologues of Vps2, placed the root for the Vps2 domain family between clades of ESCRT III-associated and ESCRT III homologues (Figure S5).

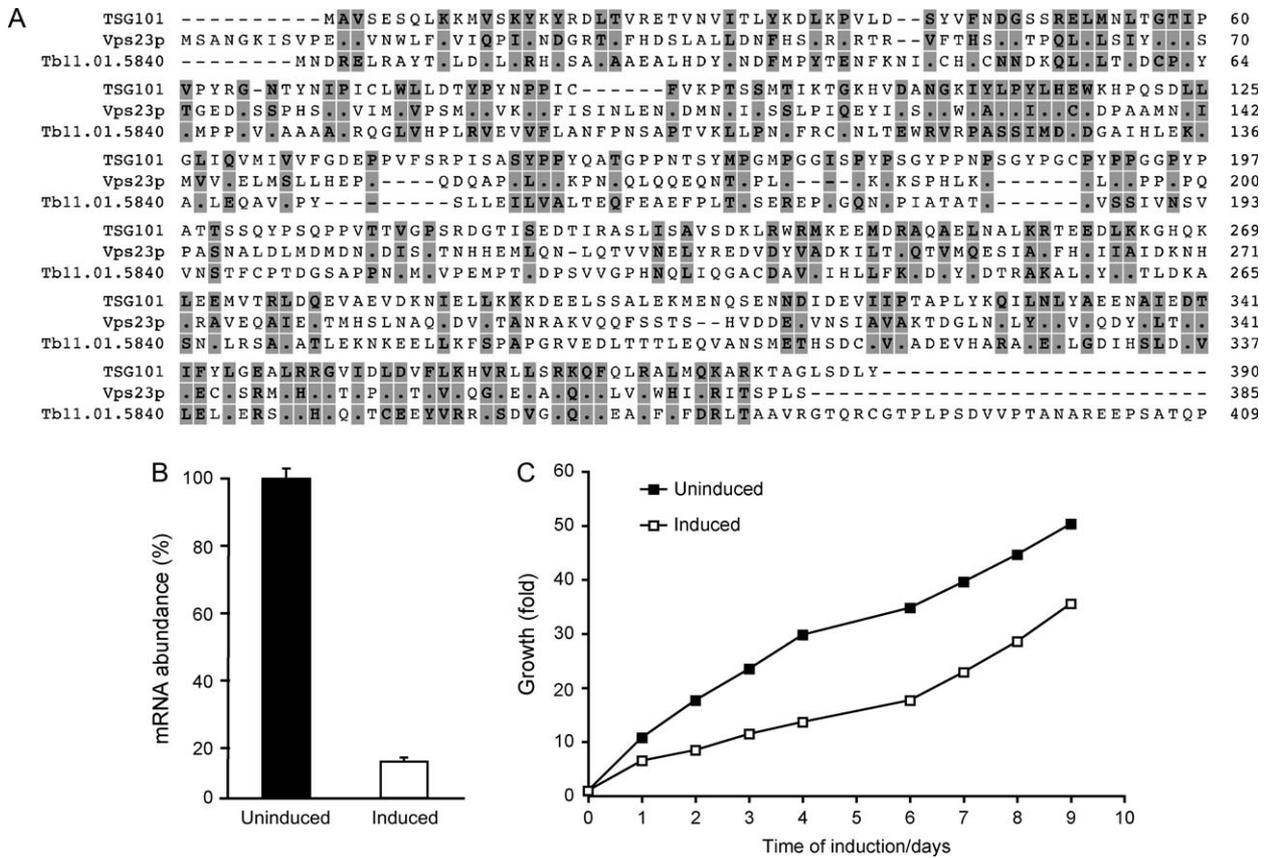
#### **Analysis of ESCRT function in a divergent system**

We considered sequence divergence, absence of the ESCRT 0 complex from non-opisthokont species and the additional taxon-specific specializations (24) to warrant direct experimental validation of predicted ESCRT genes in at least one organism. Furthermore, the differential phenotypes of ESCRT gene ablation/knockdown on endocytic trafficking in yeast and metazoa raises the possibility of significant functional divergence when considering organisms from additional supergroups. We selected the African trypanosome on account of evolutionary position and the presence of an ubiquitylation pathway within the endocytic system (35–37). Components of ESCRT I were targeted as, in metazoa and fungi, these factors provide dramatic phenotypes (42).

The trypanosome orthologues of both Vps23 and Vps28 were examined (Tb11.01.5840, Figure 5A, and Tb11.01.2510, Figure S6, respectively). Alignment of Tb11.01.5840 to *S. cerevisiae* Vps23p and *H. sapiens* TSG101 illustrates the level of divergence between the excavate and the opisthokont sequences; while the opisthokont sequences are 23% identical, Tb11.01.5840 is only 14% identical to *H. sapiens* TSG101 and 13% identical to *S. cerevisiae* Vps23. Furthermore, Tb11.01.5840 lacks a predicted ubiquitin E2 variant (UEV) domain; this may be because of extreme sequence divergence or true absence. Regardless, we consider the assignment of functional similarity equivocal based on sequence alone, compounded further by the presence of a second candidate trypanosome Vps23 orthologue, Tb11.01.4703, with marginally lower identity to Vps23 and TSG101 (Figure S7). Analysis of Tb11.01.5840 gene product function strongly indicated an orthologous relationship to Vps23 and thus is designated as TbVps23. Nonetheless, this does not rule out Tb11.01.4703 as an additional Vps23 orthologue.

#### **TbVps23 is required for proliferation and localizes to the endosomal system**

We initially ascertained if TbVps23 expression was required for cell proliferation. Knockdown of TbVps23 messenger RNA by RNA interference (RNAi) was specific and achieved >80% ablation of message, as determined by quantitative real-time polymerase chain reaction (qRT-PCR) (Figure 5B). A rapid and significant impact on the growth of bloodstream form (BSF) trypanosomes was



**Figure 5: The trypanosome orthologue of Vps23 is required for robust growth.** A) Clustal alignment of *H. sapiens* TSG101, *S. cerevisiae* Vps23p and Tb11.01.5840. Residues identical in two or more sequences are shaded gray, while residues identical to the top sequence are shown as '.'. Dashes indicate gaps introduced to optimize the alignment, and an ~100 amino acid C-terminal extension from Tb11.01.5840 is omitted. B) Knockdown of TbVps23 results in specific loss of Tb11.01.5840 messenger RNA (mRNA). qRT-PCR data are shown for induced (open symbols) or uninduced (closed symbols) cells after 48-h induction. Abundance of Tb11.01.5840 mRNA is normalized to  $\beta$ -tubulin, and knockdown is approximately 80%. Error bars derive from duplicate experiments and show standard deviation. C) Growth for cells under knockdown for TbVps23. Cells induced (open symbols) or uninduced (closed symbols) were grown over a period of 9 days and cell numbers monitored. A cumulative growth curve expressed as fold cell number is shown. Note that, at day 6, the RNAi has clearly broken, and the cells are growing at a normal rate. Similar data have been obtained for multiple independent experiments.

obtained within 24 h (Figure 5C), indicating that TbVps23 is required for normal cellular function. The culture escaped from the RNAi system after ~6 days, and a normal growth rate was restored, a common occurrence with RNAi (43,44). No evidence for a specific cell cycle block, accumulation of vesicles or enlargement of the flagellar pocket was observed (data not shown), indicating that TbVps23 does not participate directly in early endocytic events or cell cycle control (45).

TbVps23 was localized by ectopic expression of a C-terminal hemagglutinin (HA)-tagged version of the protein. Immunofluorescence demonstrated diffuse staining across the cytoplasm, with multiple puncta in interphase cells located between the nucleus and the kinetoplast, the region of the cell containing essentially all the endosomal apparatus in trypanosomes (Figure 6A) (34). Selective gating of the fluorescence signal indicated that these puncta contained

the most concentrated TbVps23 signal, and in subsequent images, selective gating was used to more clearly facilitate colocalization with endosomal markers. Co-staining with affinity-purified antibodies against trypanosome Rab5A, Rab11, clathrin, EpsinR and the lysosomal marker p67 [(46–48); Gabernet-Castello and M. C. F., manuscript in preparation] demonstrated that TbVps23 located to a compartment juxtaposed to the Rab5A early endosome and the lysosome (Figure 6B). TbVps23 also showed some colocalization with EpsinR and clathrin. However, the TbVps23 compartment is clearly distinct from the Rab11 recycling endosome, consistent with a location for TbVps23 between the early and the late regions of the endosomal system, similar to the location of TSG101 in yeast and mammalian cells (42).

We also examined the location of a second ESCRT I component, TbVps28 (Tb11.01.2510, Figure S6), to confirm

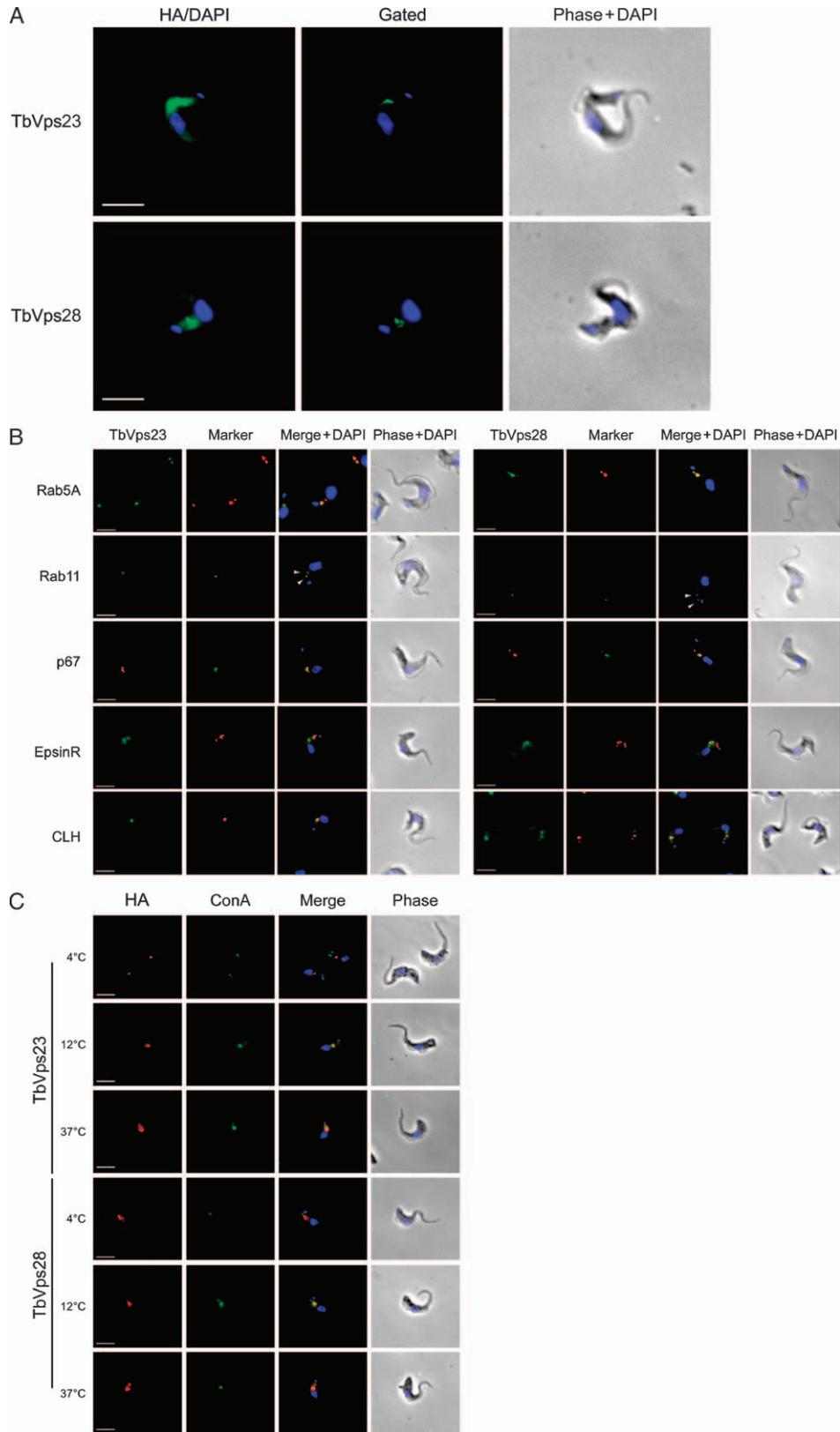


Figure 6: Legend on next page.

the assignment of TbVps23. TbVps28 is rather better conserved with opisthokont homologues than Vps23 and therefore could be more confidently assigned based on sequence. Again, an HA-epitope-tagged form of the protein was expressed in trypanosomes, and a location similar to TbVps23 was obtained (Figure 6A). The antigen localized to the posterior region of the cell between the nucleus and the kinetoplast, was juxtaposed to Rab5A and p67, and distinct from Rab11 (Figure 6B). As with TbVps23, TbVps28 showed some colocalization with EpsinR and clathrin. Furthermore, both TbVps23 and TbVps28 were juxtaposed with ConA under conditions where endocytosis is blocked at the early endosome by reduced temperature [(47), Figure 6C]. Finally, by creation of a double transformant, we coexpressed HA-TbVps23 and FLAG-tagged TbVps28 to facilitate colocalization; fluorescence signals for both proteins were coincident (Figure 7). These colocalization data essentially eliminate concerns with mistargeting of epitope-tagged protein as the probability of equivalent mistargeting of both constructs is highly remote. Hence, both TbVps23 and TbVps28 are localized to the same endosomal structures residing between the nucleus and the kinetoplast and intimately associated with both the early endosomes and the lysosome. We previously reported the presence of MVB-like structures in trypanosomes, which includes electron-dense vesicular profiles enclosed within a limiting membrane between the nucleus and the kinetoplast (30). That position is similar to the MVB in animals and fungi, and consistent with the location of TbVps23 and TbVps28, demonstrating conservation of positioning of MVBs in trypanosomes and opisthokonts.

#### **TbVps23 and TbVps28 are cytosolic factors**

We further investigated the subcellular localization of TbVps23 and TbVps28 by subcellular fractionation (Figure 8). While the TMD fusion protein BiPNTm was predominantly found in the membrane fraction, both TbVps23 and TbVps28 were almost exclusively found in the soluble fraction, consistent with the large levels of cytosolic staining observed by immunofluorescence (Figure 6A). Furthermore, this is consistent with the current model, whereby ESCRT I subunits are mainly cytosolic but recruited onto the endosomal membrane of sorting MVBs

either by ESCRT 0 or, presumably, by some other mechanism in trypanosomes.

#### **TbVps23 is required for turnover of ubiquitylated endocytic cargo proteins**

In a final test for conservation of function, we directly analyzed the role of trypanosome ESCRTs in endocytic trafficking. We previously demonstrated that the cytosolic domain of invariant surface glycoprotein (ISG)65 contains conserved lysine residues, which are essential for endosomal targeting and rapid turnover. Furthermore, using a chimeric reporter construct, we demonstrated covalent attachment of ubiquitin in a context-specific manner to proteins entering the endocytic pathway (36,37). These data suggest a high degree of similarity between endosomal trafficking of trypanosome surface proteins and MVB/lysosomal targeting in higher eukaryotes. We anticipated that knockdown of TbVps23 would result in stabilization of the protein.

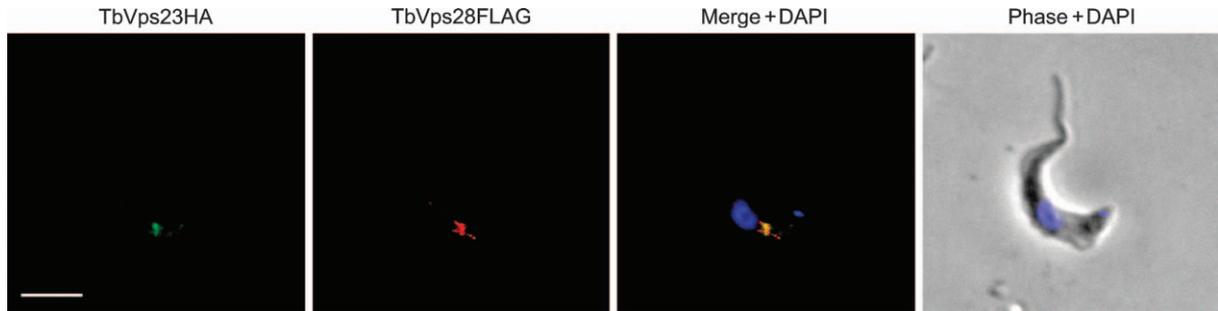
Indeed, we observed a decrease in turnover of endogenous ISG65 over a protracted period in the presence of TbVps23 knockdown in bloodstream form (BSF) cells compared with control cells. At 4 h after addition of cycloheximide, nearly 40% of ISG65 was degraded in the uninduced cells, while less than 20% was degraded when TbVps23 was knocked down (Figure 9). This stabilization was incomplete and probably because of residual expression of TbVps23 (Figure 5) or possibly alternate pathways. However, these data confirm a very similar function for TbVps23, Vps23p and TSG101 and most importantly provide validation of functional equivalence between highly divergent ESCRT complex orthologues retaining less than 15% sequence identity.

## **Discussion**

We describe an evolutionary analysis of the factors comprising the eukaryotic ESCRT complex, greatly extending several earlier partial studies [(24,28,29) and reviewed in 16]. The major feature is high conservation of ESCRT components, and with the exception of three apicomplexa, there is substantial retention of ESCRT I, II, III and ESCRT III-associated complexes by all taxa. Our findings also

---

**Figure 6: TbVps23 and TbVps28 localize to the endocytic system of trypanosomes.** A) Location of TbVps23 and TbVps28 HA-tagged proteins. Both proteins were visualized with mouse anti-HA antibody and a FITC-conjugated secondary antibody (green) and cells counterstained with DAPI for DNA (blue). TbVps23 and TbVps28 localize to the posterior of the cell, which corresponds to the location of the trypanosome endocytic apparatus. Both punctate and diffuse cytosolic staining were observed; the 'gated' panels show more harshly processed data to emphasize only the punctata. Similar gating is used for (B) and (C). B) Colocalization with a panel of endosomal markers to Rab5A, Rab11, p67, EpsinR and clathrin heavy chain (CLH). Cells are counterstained with DAPI for DNA (blue). There is clear juxtaposition for TbVps23 and TbVps28 with Rab5A (early endosome), p67 (lysosome), clathrin and EpsinR (sorting endosomes) but not for Rab11 (recycling endosome), suggesting a location close to the early endosome/lysosome axis. Positions of Rab11 and TbVps23 or TbVps28 are indicated by arrowheads in some panels for clarity. C) Demonstration that TbVps23 and TbVps28 are on endosomal compartments. Cells expressing TbVps23HA or TbVps28HA were allowed to take up FITC-conjugated ConA (green) at the indicated temperatures, fixed and co-stained for HA (red). There is clear colocalization at 12°C, where ConA is preferentially blocked at the Rab5 early endosome and some at 37°C when ConA has reached the lysosome (47). At 4°C, ConA is not internalized and remains in the flagellar pocket. Cells were co-stained for DNA with DAPI (blue). Scale bars in all panels are 2 µm.



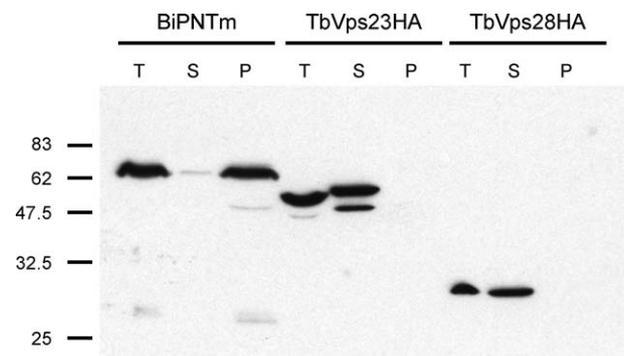
**Figure 7: TbVps23 and TbVps28 colocalize.** To confirm that TbVps23 and TbVps28 are present on the same structures, a FLAG epitope-tagged version of TbVps28 was coexpressed in the cell line already expressing HA-tagged TbVps23. When these cells were examined with anti-FLAG (red) and anti-HA (green) antibodies, clear coincidence of both epitope-tagged proteins was obtained. Cells were co-stained for DNA with DAPI, and the scale bar is 2  $\mu\text{m}$ .

suggest that CHMP7 may contribute to ESCRT function in more taxa than previously assumed. This bioinformatic evidence is extended by functional validation of selected subunits in *T. brucei*, a taxon evolutionarily distant from the organisms where ESCRTs have been characterized to date. These data indicate functional conservation from opisthokonts to excavates for at least two ESCRT I subunits and altogether suggest conservation and general applicability of all but one of the major features of the system. Bolstering this conclusion are reports of organelles morphologically similar to MVBs from taxa of the major eukaryotic supergroups (29–32).

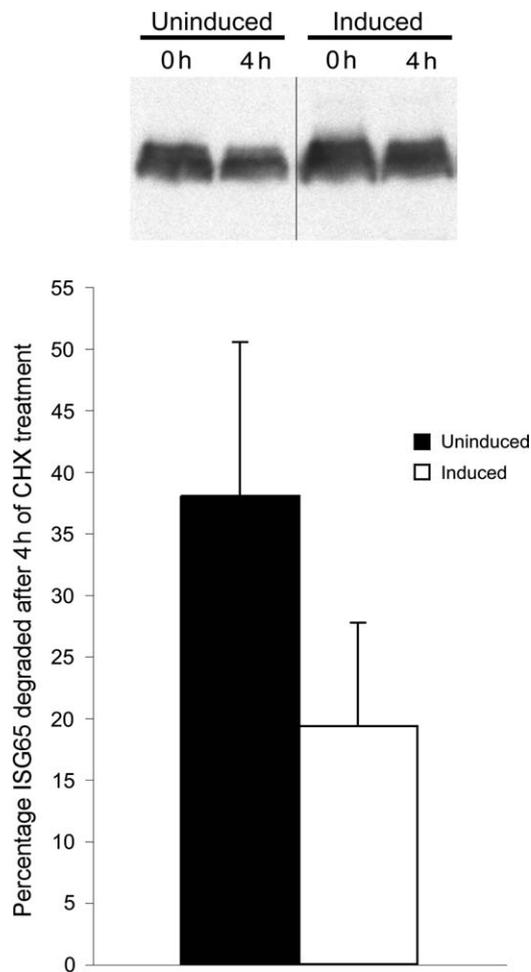
Significantly, however, we also found that the heterodimeric ESCRT 0 complex is restricted to members of the Opisthokonta. Vps27p, Hrs in mammals, is a key component for recognition of ubiquitylated cargo, binding clathrin and targeting ESCRT components to the endosomal membrane by recognition of PI3P through its FYVE domain (reviewed in 6). The role of this complex in seeding assembly of the ESCRT system is inherent to both the conveyor belt and the concentric circle models for ESCRT recruitment (18) but is clearly not generalizable across the eukaryotes. The nature of the cargo recognition mechanism and membrane targeting in non-opisthokonts and in the eukaryotic ancestor is presently completely unknown. Direct recognition through ESCRT I components is possible, specifically by the UEV domain of Vps23 (49), consistent with disruption of interactions between Vps27p and Vps23p in yeast that produces only a partial defect in cargo sorting. This suggests that recruitment of ESCRT I by ESCRT 0 is non-essential to initial cargo recognition (50). Alternatively, additional factors or an analogous complex to ESCRT 0 may be present. For example, there are multiple open reading frames (ORFs) that encode potential, and uncharacterized, FYVE domain-containing proteins in most eukaryotic genomes. In trypanosomes, PI3P is selectively present on endosome membranes and PI3-kinase activity is necessary for normal endocytosis, suggesting that endosomal membranes are selectively tagged by PI3P in such systems (51). There is no direct evidence to discrimi-

nate between these possibilities, but the present data do highlight potential plasticity within the ESCRT system. Furthermore, absence of ESCRT 0 suggests a distinct mechanism for interaction with the clathrin coat system in non-opisthokonts, and hence, regulation of delivery of material to the MVB must occur by a separate mechanism; for example, conservation across the eukaryotes of a clathrin-binding PPXY motif close to the N-terminus of Vps25 (28) indicates the presence of additional and conserved mechanisms for clathrin interaction. Significantly, trypanosome Vps23/TSG101 does not appear to retain a UEV, again suggesting lineage-specific aspects to ESCRT assembly and cargo recognition. Direct identification of such factors from divergent organisms is required to advance our understanding.

The absence of ESCRT I and II in *Plasmodium*, *Toxoplasma* and *Cryptosporidium* is also significant for both the evolution of these systems and their trafficking potential.



**Figure 8: TbVps23 and TbVps28 are predominantly cytosolic.** BSF cells overexpressing TbVps23HA, TbVps28HA or the TMD protein BiPNTm were subjected to subcellular fractionation. The apparent molecular weights of BiPNTm, TbVps23HA and TbVps28 are approximately 62, 52 and 28 kDa, respectively. BiPNTm was mostly found in the membrane fraction, while TbVps23HA and TbVps28 were almost exclusively in the cytosol. T denotes total lysate, S for supernatant and P for pellet fraction. Numbers indicate the molecular weight in kiloDalton.



**Figure 9: TbVps23 is required for degradation of ubiquitylated endocytic cargo.** Turnover of ISG65 in TbVps23/TSG101 knockdown cells. Following RNAi induction, cells were treated with cycloheximide (CHX) to initiate chase, and samples were taken at time 0 and 4 h after addition of cycloheximide. Cell lysates were then subjected to SDS-PAGE. Top panel: ISG65 was detected by Western immunoblotting using an anti-ISG65-specific antibody, and levels were estimated by densitometry and normalized to BiP expression detected by reprobings. Bottom panel: ISG65 levels were normalized to time 0 (at 100%). Graph shows the percentage of ISG65 degraded in the presence or absence of TbVps23 knockdown and represents the mean of three completely independent experiments, with the standard error indicated.

Recent phylogenomic analyses (52) place *Plasmodium* and *Theileria* as sister taxa, with *Toxoplasma* and finally *Cryptosporidium* emerging as basal lineages. This suggests that the ESCRT system has been under relaxed selection in the Apicomplexa with potentially three distinct losses of ESCRT complexes I and II. Given that all the Apicomplexa analyzed have complex life cycles, are digenetic and have an intracellular parasitic lifestyle, a clear biological rationale for the multiple losses of ESCRT I and II from *P. falciparum*, *T. gondii* and *C. parvum*, but not from

*Theileria* species, is not apparent. *T. gondii* appears to have progressed further along the line of minimization of the ESCRT system (29) and retains few ESCRT III or ESCRT III-associated subunits. Expression of an inactive form of *Plasmodium* Vps4 in *T. gondii* results in appearance of MVB-like structures located close to early endosomes, suggesting that despite the loss of most ESCRT factors, *T. gondii* retains some MVB functionality. These data argue for a similar loss of selective pressure for ESCRT I and II on multiple occasions in the Apicomplexa. Moreover, this may indicate that ESCRTs III or ESCRT III-associated components can provide minimum functionality for endosomal targeting. The role of ubiquitylation in surface protein turnover, however, is unexplored in these species, and at present, a biological rationale for these multiple and similar losses is not obvious. Interestingly, these losses are similar to the absence of several tether complex subunits from these taxa (23).

By contrast, ESCRT III or ESCRT III-associated components are retained well, including in the Apicomplexa. Vps4 can be recruited to the MVB membrane by multiple factors and does appear to have several roles in coordinating the latter stages of ESCRT function. Hence, selection for retention of ESCRT III and ESCRT III-associated factors is potentially stronger than for ESCRT I and II and also consistent with minimal functionality being provided by these complexes. This hypothesis is further supported by recent evidence for a Vps4-like AAA-ATPase and Vps2-like family homologue in Archaea (17) and together with the data in this study suggests that the ancient portions of the ESCRT system may indeed be the ESCRT III or ESCRT III-associated complexes. Given that the Snf7 and Vps24 domain proteins appear capable of assembling lattice-like structures on cell membranes (17,18), and that Vps4 can introduce energy into the system facilitating conformational change, such a minimal configuration could represent an ancestral mechanism for membrane deformation. Experimental evidence is obviously required to determine if these archaeal ESCRT precursors indeed possess homologous roles specifically in the creation of intracellular vesicles, as observed in *Ignicoccus* (53).

We also uncovered evidence for restricted multicopy genes for several ESCRT factors. Several expansions in multicellular taxa suggest a possible requirement for ESCRT diversity in different tissues and cell types. Expansion of SNARE homologues associated with secretion has been associated with multicellularity in the past (54). ESCRT subunits are also expanded in *T. vaginalis*, which has undergone expansion of its endocytic membrane trafficking complement (38).

Phylogenetic analysis allowed us to consider evolution of the ESCRT system beyond the LCEA. The presence of structurally related subunits in ESCRT III or ESCRT III-associated complexes, multiple related Snf7 and Vps24

family members and archaeal Vps24 homologues suggests a model for evolution of these factors by duplication and functional specialization (Figure 4C). The most parsimonious model is that the initial ESCRT complex was composed of a single Snf7 protein and single Vps24-related ancestral factor, with at least the latter derived from an archaeal source. Gene duplications of both the Snf7 and the Vps24 domain proteins at this stage allowed functional diversification into ESCRT III and ESCRT III-associated complexes. An additional duplication of the ESCRT III Snf7 factors gave rise to Vps32 and Vps20, while two duplications would have given rise first to Vps2A and the ancestor of 2B/24 and then to Vps2B and Vps24. Both the functional division between the ESCRT III and the ESCRT III-associated complexes and the placement of the root in the Vps2 phylogeny with the archaeal homologues support the topology proposed by our model (Figure 4C).

We extended bioinformatic insight into ESCRT system conservation by functional analysis of putative ESCRT I complex homologues in *T. brucei*. This allowed increased confidence in *in silico* predictions and examination of aspects of the ESCRT system that have been retained and are present in evolutionarily distant eukaryotes. Despite low sequence similarity for TbVps23, this factor retains a similar location within the endosomal system of trypanosomes, as in animal and fungal cells. TbVps23 also retains a similar function, blocking degradation of ubiquitylated surface proteins when subjected to RNAi-mediated knockdown. Moreover, the basic architecture and morphology of the trypanosome MVB are very similar to higher eukaryotes, both at the ultrastructural level where clear MVB-like structures can be detected (30) and in terms of relationships with the early endosomes and the terminal vacuole or lysosome.

Conservation of the endocytic system extends beyond basic characterization of molecular markers for pathways (34). Specifically, the overall features of the endocytic ubiquitylation, and possibly NEDDylation, pathway are conserved from metazoa to trypanosomes (37). However, as well as absence of an ESCRT 0 complex, we also note the absence of detectable homologues of the E3 ligases c-Cbl and Rsp5 and canonical UIM-containing epsin and eps15 molecules from non-opisthokont lineages (24,37). Overall, this suggests a distinct system operates in many, if not most, eukaryotic lineages for conjugation and recognition of surface molecules by ubiquitin.

Two separate mechanisms have been invoked to explain the evolutionary emergence of components of the eukaryotic membrane trafficking system. Components such as the SNAREs, Rabs, Sec1/Munc18 (SM) proteins and vesicle coats are proposed to encode organelle identity in some manner (55). These protein families are the products of gene duplications whereby, with one notable exception (25), the major organelle-associated families had already appeared prior to the LCEA (reviewed in 22) and most

consistent with concerted expansion, giving rise to the diverse trafficking organelles in modern cells (22,25). By contrast, evidence from systems such as the tethering complexes suggests that independent innovation of trafficking machinery also occurred at a stage prior to the LCEA (23). Clear patterns of increased complexity by gene duplication exist for ESCRT III and ESCRT III-associated complexes, but no other homology was detectable for the additional complex subunits. Thus, ESCRT evolution appears to encompass both mechanisms.

Three major vesicle coats, COPI, COPII and clathrin, are likely to be derived from a single ancestral membrane deformation complex through paralogous expansion (22). The ESCRT machinery likely evolved independently, because there is no obvious homology with these vesicle coats, and deform membranes by a distinct mechanism and in the opposite topology, all suggesting that LCEA possessed at least two vesicle forming systems with distinct origins, but which subsequently merged into a single endomembrane system.

This model of ESCRT complement and function appears to be broadly generalizable across the eukaryotes based on bioinformatic and functional studies. However, identification of the ESCRT 0 complex as an animal- and fungal-specific feature, along with the clear functional and component differences in *T. brucei*, stand as strong reminders that, while the system is largely conserved, important differences also exist between lineages.

## Materials and Methods

### Databases

Taxon sampling representing five of the six eukaryotic supergroups was used (33); a representative Rhizarian has not been fully sequenced. In the main analysis, only complete genomes were included so that failure to retrieve a candidate could be more reliably ascribed to absence or extreme divergence. At least 2 representative taxa were included from each supergroup (total of 21) to facilitate detection of secondary losses versus absence within an entire group and to minimize detection failure because of species-specific divergence. Trypanosomatid, *Dictyostelium discoideum* and *Entamoeba histolytica* data are at the Sanger Institute (<http://www.genedb.org/>), yeast data at Munich Information Center for Protein Sequences (MIPS) (<http://mips.gsf.de/genre/proj/yeast/index.jsp>) or the yeast genome database (<http://www.yeastgenome.org/>). Data for *Chlamydomonas reinhardtii*, *Naegleria gruberi*, *Ostreococcus tauri*, *Phytophthora ramorum* and *Thalassiosira pseudonana* were from Joint Genome Institute (JGI) ([http://genome.jgi-psf.org/euk\\_cur1.html](http://genome.jgi-psf.org/euk_cur1.html)). *T. gondii* data are at ToxoDB (<http://www.toxodb.org/>), *C. parvum* data from CryptoDB (<http://www.cryptodb.org/cryptodb/>), *T. parva* and *Tetrahymena thermophila* data from The Institute for Genomic Research (TIGR) (<http://www.tigr.org/tdb/e2k1/tpa1/>, <http://www.tigr.org/tdb/e2k1/ttg/>), *Giardia intestinalis* data at GiardiaDB (<http://www.giardiadb.org/giardiadb/>), *Cryptococcus neoformans* data from the Broad Institute ([http://www.broad.mit.edu/annotation/genome/cryptococcus\\_neoformans/Blast.html](http://www.broad.mit.edu/annotation/genome/cryptococcus_neoformans/Blast.html)) and *Cyanidioschyzon merolae* data at the *C. merolae* BLAST server (<http://merolae.biol.s.u-tokyo.ac.jp/blast/blast.html>). Additional organismal genome data were obtained through the NCBI BLAST interface (<http://www.ncbi.nlm.nih.gov/BLAST/>). In addition, predicted proteomes for most species were downloaded by ftp from the respective database for local analysis.

To further validate the prediction that ESCRT 0 factors are opisthokont specific, the following additional near-complete genomes or EST data sets were searched for Vps27/HRS and Hse1/STAM; *Hartmannella vermiformis*, *Euglena gracilis*, *Malawimonas jakobiformis*, *Trimastix pyriformis* and *Reclinomonas americana* are at NCBI (<http://www.ncbi.nlm.nih.gov/BLAST/>), *Emiliania huxleyi*, *Micromonas pusilla*, *Physcomitrella patens*, *Populus trichocarpa* and *Volvox carteri* at JGI ([http://genome.jgi-psf.org/euk\\_cur1.html](http://genome.jgi-psf.org/euk_cur1.html)), *Eimeria tenella* at GeneDB (<http://www.genedb.org/>) and *Acanthamoeba castellanii* was from the Human Genome Sequencing Center (<http://www.hgsc.bcm.tmc.edu/bcm/blast/microbialblast.cgi?organism=>). Although some VHS domain proteins and FYVE domain proteins were obtained by these searches, none had the complete domain architecture of ESCRT 0 subunits or completed reverse BLAST routines. Finally, we also searched the entire non-redundant sequence database at NCBI for these factors, restricting the analysis to 'Eukaryota, NOT Opisthokonta'.

### Taxonomic homology survey

Models of the ESCRT complexes and their components are based on published descriptions [(15) *inter alia*]. Initial queries typically used *H. sapiens*- or *S. cerevisiae*-predicted proteins and BLAST (56) using the BLOSUM45 amino acid matrix with manual cutoff. Where initial searches failed to recover a candidate, if a clear orthologue was available, query sequences from a closely related taxon to the target genome were used or a Smith and Waterman (57) search was conducted locally using CLC WORKBENCH V 3.0.1 with Cube hardware acceleration (<http://www.clcbio.com/>) against the predicted proteome. All recovered sequences were subjected to reverse BLAST using BLOSUM45 or Smith and Waterman for confirmation of orthology. Sequences were also parsed through the NCBI conserved domain database (CDD) using default parameters and analyzed by alignment for the presence of significant sequence similarities. In addition, candidates with weak BLAST scores were subjected to phylogeny to confirm orthology. A candidate orthologue was considered retrieved if reverse BLAST recovered the original query within the top five hits. Failure to complete these tests resulted in an assignment of 'not found'. To address the possibility that distant sequence relationships were present between the ESCRT subunits, all the *S. cerevisiae* sequences were analyzed using three iterations of PSI-BLAST against the *S. cerevisiae* genome. All results were inspected. Additional alignments were performed to verify relationships.

### Alignments and phylogenetic reconstructions

Alignments were initially created using CLUSTAL (58) and adjusted manually. Only unambiguous homologous regions were retained for phylogenetic analysis. Alignments are available upon request from the authors. A data set of Snf7 homologues containing all homologues from all taxa sampled was initially analyzed to detect long-branch sequences and then further refined to produce the topology shown in Figure 3A. A similar pair of data sets were created and analyzed to produce the data in Figure 3B. Additional data sets were created to enable the specific classification of sequences from particular taxa, to root the Vps2 domain family tree with archaeal homologues and to assess the evolution of the CHMP7. Details of data sets are in Table S2.

For each analysis, PROT-TEST (59) was used to estimate the appropriate model of sequence evolution. In all cases, this resulted in the application of a WAG model plus a four-category gamma correction for among-site rate variation. Phylogenetic analysis was performed by three separate methods. To obtain the best Bayesian tree topology and posterior probability values, the program MRBAYES version 3.1.2 was used (60). Analyses were run for  $10^6$  MCMC generations, with removal of all trees prior to a plateau, established by graphical estimation. All calculations were checked for convergence and had a splits frequency <0.1. Maximum-likelihood (ML) analysis was also performed using the programs PHYLML (61) and RAxML (62) with 100 bootstrap replicates. The initial and final (Figure 3) SNF7 and Vps2 data sets were additionally analyzed by Parsimony and Neighbor-joining algorithms using the programs PROTPARS and NEIGHBOR from the PHYLIP package (64) with 100 bootstrap replicates.

Assessment of long-branch artifact and identification of rapidly evolving taxa were performed by analysis of an initial data set for each of the SNF7 (Figure S2) and Vps2 (Figure S4) protein families by all five analytical methods. The resulting topologies were mostly consistent, but the support values in the parsimony and neighbor-joining analyses were significantly lower, particularly for monophyly of Vps60 (Table S3), suggesting that long-branch artifact might be interfering with phylogenetic signal. Consistent with this, several sequences appeared to be rapidly evolving as assessed by incongruous regions in the alignments and in divergent branch lengths in the resulting tree topologies. Once removed, the support values for the key nodes rose in general (Figures 3, Figures S8 and S9, Table S3) and were similar across all methods suggesting alleviation of long-branch artifacts.

Nodes with better than 0.95 and 80% bootstrap support are considered robust, and nodes with better than 0.80 posterior probability and 50% bootstrap support are shown.

### Culturing of bloodstream forms of T. brucei

Bloodstream form Molteno Institute trypanosomal antigen type (MITat) 1.2, derived from Lister strain 427 and expressing VSG 221, were cultured in Hirumi medium iteration (HMI-9) complete medium (HMI-9 supplemented with 10% heat-inactivated FBS, 100 U/mL penicillin, 100 U/mL streptomycin and 2 mM L-glutamine) (65) at 37°C with 5% CO<sub>2</sub> in a humid atmosphere in non-adherent culture flasks with vented caps. Cells were maintained at densities between 10<sup>5</sup> and 5 × 10<sup>6</sup> cells/mL. Single marker bloodstream (SMB) line was used for expression of tetracycline-inducible constructs (66). Ectopic expression of plasmid constructs was maintained using antibiotic selection at the following concentrations: G418 and hygromycin B at 2.5 µg/mL and puromycin at 0.1 µg/mL.

### Plasmid constructs for overexpression and RNA interference

Primers for amplification of an RNAi target fragment were designed using RNAit (67). RNAi target fragments were PCR amplified using *Taq* DNA polymerase with TbVps23RNAi-F (5'-ACTTTTGTCCACAGACGG-3') and TbVps23RNAi-R (5'-CCAACAACCTCCAACACATCG-3') and cloned into the tetracycline-inducible RNAi expression vector p2T7<sup>TAblue</sup>, linearized with Eam1105I. Trypanosome orthologues of Vps23 and Vps28 were PCR amplified from trypanosome 427 genomic DNA using *Vent* DNA polymerase (New England Biolabs). PCR products were cloned into the BSF expression vector pXS5, containing sequence for a C-terminal HA-epitope tag, using *Hind*III and *Apal* using the following primers: TbVps23-F 5'-CTAGAAGCTTATGAATGATAGGG AGTTGCGG and TbVps23HA-R 5'-CTATGGGCCCGACATAATGGCATGAAGCTG for TbVps23HA; TbVps28-F 5'-CATAAGCTTATGGAAGTGGCATTCCACATTTC and TbVps28HA-R5'-CTATGGGCCCGCCTTTGGCCGTCGCGC for TbVps28HA. TbVps28FLAG was PCR amplified with TbVps28-F and Vps28FLAG-R 5'-CATGAATTCCTAC TTGTCGTCATCGTCTTTGTAGTCCGCTTTGGCCGTCGCG-3'. The pH1034-GRASP-RFP vector and TbVps28FLAG PCR product were digested with *Bam*HI and *Eco*RI, respectively, and then blunt-ended with Klenow fragment for 3 h at 37°C. Klenow-treated vector and PCR product were purified and digested with *Hind*III. Following DNA purification, the plasmid vector was treated with shrimp alkaline phosphatase for 24 h at 37°C. All constructs were verified by standard sequencing methods (Geneservice Ltd) prior to introduction into trypanosomes, and expression was further verified by western blotting where appropriate.

### Transfection of BSF T. brucei

3 × 10<sup>7</sup> cells (0.9–1.0 × 10<sup>6</sup> cells/mL) per transfection were harvested at 800 × g for 10 min and washed with cytomix (2 mM EGTA, pH 7.6; 120 mM KCl; 0.15 mM CaCl<sub>2</sub>; 10 mM K<sub>2</sub>HPO<sub>4</sub>/KH<sub>2</sub>PO<sub>4</sub>, pH 7.6; 5 mM MgCl<sub>2</sub>; 0.5% glucose; 100 µg/mL BSA; 1 mM hypoxanthine and 25 mM HEPES, pH 7.6). Cells were resuspended in 400 µL of cytomix and transferred to 2-mm gap electrocuvettes containing 10–25 µg of linearized DNA plasmid. Electroporation was performed using a Bio-Rad Gene Pulser II (1.4 kV and 25 µF). Cells were transferred to a flask containing HMI-9 complete medium and allowed to recover at 37°C for 6 h. Antibiotic(s) was added for selection, and

cells were subdivided into a 24-well plate. Positive clones were selected 5- to 6-day post-transfection.

### RNA interference and growth analysis

Following transfection with the relevant RNAi construct, clones were selected for hygromycin antibiotic resistance. RNAi was induced with 1  $\mu\text{g}/\text{mL}$  tetracycline, which was added fresh daily. Cells were quantified using a Z2 Coulter Counter (Coulter Electronics) and maintained between  $10^5$  and  $5 \times 10^6$  cells/mL.

### Quantitative real-time polymerase chain reaction

$1 \times 10^8$  cells were harvested at  $3500 \times g$  for 10 min at  $4^\circ\text{C}$  and washed with ice-cold PBS and quick frozen in dry ice for 1 min. RNA was purified using the RNeasy mini kit (Qiagen) according to the manufacturer's instructions. RNA concentration was quantified using a ND-1000 spectrophotometer and NANODROP software (Nanodrop Technologies). qRT-PCR was performed using iQ-SYBRGreen Supermix on a MiniOpticon Real-Time PCR Detection System (Bio-Rad) and was quantified using OPTICON3 software (Bio-Rad). The following primers were used for qRT-PCR: TbVps23-RTF2 5'-AGCAATTTGAAGCTCGGTTT and TbVps23-RTR2 5'-ATGGCTCTTCTCGTG-CATT and was performed exactly as described (37).

### Immunofluorescence analysis

Cells were harvested at  $800 \times g$  for 10 min at  $4^\circ\text{C}$  and washed with ice-cold Voorheis's-modified phosphate-buffered saline (vPBS; PBS supplemented with 10 mM glucose and 46 mM sucrose, pH 7.6) (68). Cells were then fixed in 3% paraformaldehyde in vPBS for 10 min at  $4^\circ\text{C}$ . Fixed cells were washed with ice-cold vPBS and applied to poly-lysine microscope slides (VWR International) and sectioned with an ImmEdge Pen (Vector Laboratories) for 20 min. For permeabilization, cells were incubated with 0.1% Triton-X-100 in PBS for 10 min at room temperature and washed three times for 5 min with PBS. Samples were blocked in 20% FBS in PBS for 1 h at room temperature. Fixed cells were incubated with primary antibodies for 1 h at room temperature, followed by three washes for 5 min in PBS. Secondary antibodies were then applied for 1 h at room temperature and washed as above. Samples were dried and coverslips were mounted using Vectashield mounting medium supplemented with 4',6-diamidino-2-phenylindole (DAPI) (Vector Laboratories, Inc.). Coverslips were sealed with nail varnish (Max Factor Inc.).

Antibodies were used at the following dilutions: mouse and rabbit anti-HA epitope immunoglobulin G (IgG) (both from Santa Cruz Biotechnology Inc.) at 1:1000, rabbit anti-ISG65 (from M. Carrington, Cambridge, UK) at 1:1000, rabbit anti-Rab5A at 1:200, rabbit anti-Rab11 at 1:400, mouse anti-p67 (from J. Bangs, Madison, WI, USA) at 1:1000, rabbit anti-clathrin heavy chain at 1:250 and rabbit anti-TbEpsinR at 1:500. Secondary antibodies were used at the following dilutions: anti-mouse Oregon Green (Molecular Probes) at 1:1000 and anti-rabbit Cy3 (Sigma) at 1:1000. The cells were examined on a Nikon Eclipse E600 epifluorescence microscope fitted with optically matched filter blocks and a Hamamatsu ORCA charge-coupled device camera. Digital images were captured using METAMORPH software (Universal Imaging Corp.) on a Windows XP computer (Microsoft Inc.), and the raw images were processed using ADOBE PHOTOSHOP 7.0 (Adobe Systems Inc.).

### Concanavalin A uptake assay

$1.5 \times 10^6$  cells were harvested at  $800 \times g$  for 10 min at  $4^\circ\text{C}$  and washed with serum-free HMI-9 supplemented with 1% BSA. Cells were subsequently incubated in serum-free HMI-9 with 1% BSA for 20 min at  $4^\circ\text{C}$ ,  $12^\circ\text{C}$  or  $37^\circ\text{C}$  in Eppendorf tubes, after which they were labeled with 50  $\mu\text{g}/\text{mL}$  fluorescein isothiocyanate (FITC)-conjugated concanavalin A (ConA) (Molecular Probes) and incubated for a further 30 min at the above temperatures to allow for uptake. FITC-ConA was diluted out by the addition of vPBS, and cells were harvested by centrifugation at  $800 \times g$  for 10 min at  $4^\circ\text{C}$ . Labeled cells were fixed and co-stained as described above.

### Subcellular fractionation studies

$1 \times 10^8$  cells were washed twice in ice-cold PBS and subsequently resuspended in 100  $\mu\text{L}$  of hypotonic lysis buffer (10 mM Tris-HCl, pH

7.5). Cells were incubated on ice for 5 min and then centrifuged at  $4^\circ\text{C}$  for 10 min at  $20\,000 \times g$ . The supernatant was transferred to a fresh tube, mixed with an equal volume of  $2\times$  sodium dodecyl sulfate (SDS) sample buffer and incubated at  $95^\circ\text{C}$  for 10 min. The pellet fraction was washed with 100  $\mu\text{L}$  of hypotonic lysis buffer and then resuspended in 100  $\mu\text{L}$  of ice-cold sample lysis buffer (50 mM Tris-HCl, pH 7.5; 150 mM NaCl and 1% Nonidet P-40) and incubated on ice for 25 min to lyse membranes. An equal volume of  $2\times$  SDS sample buffer was added and incubated at  $95^\circ\text{C}$  for 10 min. The supernatant and membrane fractions were then subjected to SDS-PAGE and Western immunoblotting.

### Protein turnover

SMB cells transfected with RNAi constructs were grown in the presence or absence of tetracycline for 48 h. Protein synthesis was blocked by the addition of cycloheximide (50  $\mu\text{g}/\text{mL}$ ), and  $1 \times 10^7$  cells were harvested at various time-points by centrifugation at  $800 \times g$  for 10 min at  $4^\circ\text{C}$ . Cells were washed in ice-cold PBS, then resuspended in  $1\times$  SDS sample buffer and incubated at  $95^\circ\text{C}$  for 10 min. Samples were subjected to electrophoresis on 12.5% SDS-PAGE gels, transferred to polyvinylidene difluoride membranes and proteins were detected by western immunoblotting. Antibodies were used at the following dilutions: rabbit anti-ISG65 at 1:5000, rabbit anti-BiP (J. Bangs) at 1:10 000, anti-rabbit IgG peroxidase conjugated (A0545; Sigma) at 1:10 000 and mouse anti-HA epitope IgG (Santa Cruz Biotechnology Inc.) at 1:10 000. Densitometry quantification was achieved using IMAGEJ software (National Institutes of Health).

### Acknowledgments

This study was supported by the Wellcome Trust (project and program grants to M. C. F. and a traveling fellowship to J. B. D.). We are indebted to CAMPOD for provision of equipment funds (to M. C. F.) and the Parke-Davis fund for salary support (to J. B. D.). We also acknowledge the CamGrid computational resource on which some of these analyses were performed. We thank Mark Carrington (Cambridge, UK) for antibody to ISG65, James Bangs (Madison, WI, USA) for anti-p67 antibody and Carme Gabernet-Castello (Cambridge, UK) for anti-TbEpsinR antibody. Alignments are available upon request from the authors.

### Supporting Information

Additional Supporting Information may be found in the online version of this article:

**Table S1:** Distribution of ESCRT components across a range of eukaryotic taxa. Data are based on BLAST results together with alignments and phylogenetic analysis (see *Materials and Methods*). Typically, the *S. cerevisiae* or *H. sapiens* sequences were used as queries; large taxon groupings are color coded. Individual accession numbers are given. Note that because of lability of some genome data, these accession numbers may not be available through the web interface but can be found in the compiled fasta files of the unfiltered predicted peptides available as downloads from individual genome Web sites. File is offered in Excel (.xls) format.

**Table S2:** Details of phylogenetic analyses. A summary of data set and analytical details for each of the data sets analyzed. The question column refers to the objective of the analysis, for example Vps24 family without long-branch taxa (LBT) = Vps24all-LBT, SNF7 data set with LBT excluded but *Cyanidioschyzon merolae* reincluded to assess the classification of its SNF7 homologues, etc = *C. merolae* specific.

**Table S3:** Support values for key nodes under various methods of phylogenetic analysis. This table summarizes the support values for the major families of SNF7 and Vps2 proteins in both the full and the restricted data sets. Note that in the full data sets, the neighbor-joining and parsimony analyses

show significantly lower support values than the Bayesian and maximum-likelihood (ML) methods, consistent with long-branch artifact interfering with the phylogenetic analyses. For each data set, the values are listed by families and by methods. MB, MrBAYES, PHY, PHYML, RAx, RAxML, NJ, neighbor-joining, Pars, parsimony.

**Figure S1: Coiled-coil architecture of Vps20/32/60 and Vps2/24/46.** Raw output from coils ([http://www.ch.embnet.org/software/COILS\\_form.html](http://www.ch.embnet.org/software/COILS_form.html)) is shown. Note that the arrangement of the predicted coiled-coil regions is distinct between the two families.

**Figure S2: Phylogenetic analysis of all identified SNF7 family homologues from all taxa sampled.** The tree topologies and support values from Bayesian, PHYML, RAxML, neighbor-joining and parsimony analyses of the data set SNF7.R1 are shown here.

**Figure S3: Phylogenetic analysis of SNF7 and CHMP7 homologues.** Bayesian phylogenetic tree for Vps60/32/20 and the two domains of CHMP7 homologues. This analysis shows the single origin of CHMP7 N-terminal domains. In this analysis, a selection of eukaryotes was analyzed based on the alignment from Figure 3A and including taxa with putative CHMP7 homologues. Numerical values indicate statistical support for MrBAYES (posterior probability) for all nodes with support better than 0.95, while specific nodes are also shown that differentiate the subfamilies with statistical support for MrBAYES/PHYML/RAxML (posterior probability/bootstrap/bootstrap, respectively). Taxon colors represent supergroups and correspond to those in Figure 2. Sequences are identified by the two letter abbreviation of their genus and species, followed by their gene name (Vps60, 32, 20 or SNF7). CHMP7 homologues are designated by C7 and then CT or NT for C-termini and N-termini, respectively.

**Figure S4: Phylogenetic analysis of all identified Vps2 family homologues from all taxa sampled.** The tree topologies and support values from Bayesian, PHYML, RAxML, neighbor-joining and parsimony analyses of the data set Vps2.R1 are shown here.

**Figure S5: Phylogenetic reconstruction of Vps24 ESCRT factor family with archaeal orthologues.** Bayesian phylogenetic tree for Vps2/24/46 families. Numerical values indicate statistical support for MrBAYES (posterior probability) for all nodes with support better than 0.95, while specific nodes are also shown that differentiate the subfamilies with statistical support for MrBAYES/PHYML/RAxML (posterior probability/bootstrap/bootstrap, respectively). Taxon colors represent supergroups and correspond to those in Figure 2.

**Figure S6: Alignment of Tb11.01.2510 and Vps28 from *S. cerevisiae* and *H. sapiens*.** A Clustal alignment of the two *H. sapiens* VPS28 isoforms, *S. cerevisiae* Vps28 and Tb11.01.2510 (TbVps28) is shown using default parameters. Fully conserved residues are marked '\*', strongly conserved are indicated by ':' and weakly conserved by '.'. Dashes indicate gaps introduced to optimize the alignment. A percent identity matrix is also shown; overall conservation of Vps28 sequences is considerably better than for Vps23.

**Figure S7: Alignment of Tb11.01.4703, Vps23 and TSG101.** A Clustal alignment of the *H. sapiens* TSG101, *S. cerevisiae* Vps23 and Tb11.01.4703 is shown using default parameters. Fully conserved residues are marked '\*', strongly conserved are indicated by ':' and weakly conserved by '.'. Dashes indicate gaps introduced to optimize the alignment. A percent identity matrix is also shown.

**Figure S8: Neighbor-joining and Parsimony results for the SNF7-restricted data set.** This shows the output of neighbor-joining and parsimony analyses of the final SNF7 data set shown in Figure 3A.

**Figure S9: Neighbor-joining and Parsimony results for the Vps2-restricted data set.** This shows the output of neighbor-joining and parsimony analyses of the final Vps2 data set shown in Figure 3B.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

## References

- Haglund K, Sigismund S, Polo S, Szymkiewicz I, Di Fiore PP, Dikic I. Multiple monoubiquitination of RTKs is sufficient for their endocytosis and degradation. *Nat Cell Biol* 2003;5:461–466.
- Kerscher O, Felberbaum R, Hochstrasser M. Modification of proteins by ubiquitin and ubiquitin-like proteins. *Annu Rev Cell Dev Biol* 2006; 22:159–180.
- de Melker AA, van der Horst G, Calafat J, Jansen H, Borst J. c-Cbl ubiquitinates the EGF receptor at the plasma membrane and remains receptor associated throughout the endocytic route. *J Cell Sci* 2001; 114:2167–2178.
- Galan JM, Moreau V, Andre B, Volland C, Haguenaer-Tsapis R. Ubiquitination mediated by the Npi1p/Rsp5p ubiquitin-protein ligase is required for endocytosis of the yeast uracil permease. *J Biol Chem* 1996;271:10946–10952.
- Hicke L, Dunn R. Regulation of membrane protein transport by ubiquitin and ubiquitin-binding proteins. *Annu Rev Cell Dev Biol* 2003; 19:141–172.
- Hurley JH, Emr SD. The ESCRT complexes: structure and mechanism of a membrane-trafficking network. *Annu Rev Biophys Biomol Struct* 2006;35:277–298.
- Razi M, Futter CE. Distinct roles for TSG101 and HRS in multivesicular body formation and inward vesiculation. *Mol Biol Cell* 2006;17: 3469–3483.
- Katzmann DJ, Babst M, Emr SD. Ubiquitin-dependent sorting into the multivesicular body pathway requires the function of a conserved endosomal protein sorting complex, ESCRT-I. *Cell* 2001;106: 145–155.
- Rieder SE, Banta LM, Kohrer K, McCaffery JM, Emr SD. Multilamellar endosome-like compartment accumulates in the yeast Vps28 vacuolar protein sorting mutant. *Mol Biol Cell* 1996;7: 985–999.
- Raymond CK, Howald-Stevenson I, Vater CA, Stevens TH. Morphological classification of the yeast vacuolar protein sorting mutants: evidence for a prevacuolar compartment in class E vps mutants. *Mol Biol Cell* 1992;3:1389–1402.
- Welsch S, Muller B, Krausslich HG. More than one door—budding of enveloped viruses through cellular membranes. *FEBS Lett* 2007;581: 2089–2097.
- Shim JH, Xiao C, Hayden MS, Lee KY, Trombetta ES, Pypaert M, Nara A, Yoshimori T, Wilm B, Erdjument-Bromage H, Tempst P, Hogan BL, Mellman I, Ghosh S. CHMP5 is essential for late endosome function and down-regulation of receptor signaling during mouse embryogenesis. *J Cell Biol* 2006;172:1045–1056.
- Wagner KU, Krempler A, Qi Y, Park K, Henry MD, Triplett AA, Riedlinger G, Rucker IE, Hennighausen L. TSG101 is essential for cell growth, proliferation, and cell survival of embryonic and adult tissues. *Mol Cell Biol* 2003;23:150–162.
- Martin-Serrano J, Yarovoy A, Perez-Caballero D, Bieniasz PD. Divergent retroviral late-budding domains recruit vacuolar protein sorting factors by using alternative adaptor proteins. *Proc Natl Acad Sci U S A* 2003;100:12414–12419.

15. Bowers K, Lottridge J, Helliwell SB, Goldthwaite LM, Luzio JP, Stevens TH. Protein-protein interactions of ESCRT complexes in the yeast *Saccharomyces cerevisiae*. *Traffic* 2004;5:194–210.
16. Williams RL, Urbe S. The emerging shape of the ESCRT machinery. *Nat Rev Mol Cell Biol* 2007;8:355–368.
17. Obita T, Saksena S, Ghazi-Tabatabai S, Gill DJ, Perisic O, Emr SD, Williams RL. Structural basis for selective recognition of ESCRT-III by the AAA ATPase Vps4. *Nature* 2007;449:735–739.
18. Nickerson DP, Russell MR, Odorizzi G. A concentric circle model of multivesicular body cargo sorting. *EMBO Rep* 2007;8:644–650.
19. Dupre S, Haguenaer-Tsapis R. Deubiquitination step in the endocytic pathway of yeast plasma membrane proteins: crucial role of Doa4p ubiquitin isopeptidase. *Mol Cell Biol* 2001;21:4482–4494.
20. Babst M, Sato TK, Banta LM, Emr SD. Endosomal transport function in yeast requires a novel AAA-type ATPase, Vps4p. *EMBO J* 1997;16:1820–1831.
21. Babst M, Wendland B, Estepa EJ, Emr SD. The Vps4p AAA ATPase regulates membrane association of a Vps protein complex required for normal endosome function. *EMBO J* 1998;17:2982–2993.
22. Dacks JB, Field MC. Evolution of the eukaryotic membrane-trafficking system: origin, tempo and mode. *J Cell Sci* 2007;120:2977–2985.
23. 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.
24. 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.
25. Dacks JB, Poon PP, Field MC. Phylogeny of endocytic components yields insight into the process of non-endosymbiotic organelle evolution. *Proc Natl Acad Sci U S A* 2008;105:588–593.
26. Field MC, Natesan SK, Gabernet-Castello C, Koumandou VL. Intracellular trafficking in the trypanosomatids. *Traffic* 2007;8:629–639.
27. Morgan GW, Hall BS, Denny PW, Field MC, Carrington M. The endocytic apparatus of the kinetoplastida. Part II: machinery and components of the system. *Trends Parasitol* 2002;18:540–546.
28. Slater R, Bishop NE. Genetic structure and evolution of the Vps25 family, a yeast ESCRT-II component. *BMC Evol Biol* 2006;6:59.
29. 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;117:3831–3838.
30. 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;156:175–190.
31. Walker G, Dacks JB, Embley TM. Ultrastructural description of *Breviata anathema*, n. gen., n. sp., the organism previously studied as “*Mastigamoeba invertens*”. *J Eukaryot Microbiol* 2006;53:65–78.
32. Tse YC, Mo B, Hillmer S, Zhao M, Lo SW, Robinson DG, Jiang L. Identification of multivesicular bodies as prevacuolar compartments in *Nicotiana tabacum* BY-2 cells. *Plant Cell* 2004;16:672–693.
33. Adl SM, Simpson AG, Farmer MA, Andersen RA, Anderson OR, Barta JR, Bowser SS, Brugerolle G, Fensome RA, Fredericq S, James TY, Karpov S, Kugrens P, Krug J, Lane CE et al. The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *J Eukaryot Microbiol* 2005;52:399–451.
34. Field MC, Carrington M. Intracellular membrane transport systems in *Trypanosoma brucei*. *Traffic* 2004;5:905–913.
35. Engstler M, Bangs JD, Field MC. Intracellular transport systems in Trypanosomes: function, evolution and virulence. In: Barry JD, Mottram J, McCulloch R, Acosta-Serrano A, editors. *Trypanosomes – After the Genome*. Horizon Press, Norfolk, UK 2007;281–318.
36. Chung WL, Carrington M, Field MC. Cytoplasmic targeting signals in transmembrane invariant surface glycoproteins of trypanosomes. *J Biol Chem* 2004;279:54887–54895.
37. Chung WL, Leung KF, Carrington M, Field MC. Ubiquitylation is required for degradation of *trans*-membrane surface proteins in trypanosomes. doi: 10.1111/j.1600-0854.2008.00785.x.
38. Carlton JM, Hirt RP, Silva JC, Delcher AL, Schatz M, Zhao Q, Wortman JR, Bidwell SL, Alsmark UC, Besteiro S, Sicheritz-Ponten T, Noel CJ, Dacks JB, Foster PG, Simillion C et al. Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. *Science* 2007;315:207–212.
39. Marti M, Regos A, Li Y, Schraner EM, Wild P, Muller N, Knopf LG, Hehl AB. An ancestral secretory apparatus in the protozoan parasite *Giardia intestinalis*. *J Biol Chem* 2003;278:24837–24848.
40. Morrison HG, McArthur AG, Gillin FD, Aley SB, Adam RD, Olsen GJ, Best AA, Cande WZ, Chen F, Cipriano MJ, Davids BJ, Dawson SC, Elmendorf HG, Hehl AB, Holder ME et al. Genomic minimalism in the early diverging intestinal parasite *Giardia lamblia*. *Science* 2007;317:1921–1926.
41. Lanfredi-Rangel A, Attias M, de Carvalho TM, Kattenbach WM, De Souza W. The peripheral vesicles of trophozoites of the primitive protozoan *Giardia lamblia* may correspond to early and late endosomes and to lysosomes. *J Struct Biol* 1998;123:225–235.
42. Bishop N, Woodman PG. TSG101/mammalian VPS23 and mammalian VPS28 interact directly and are recruited to VPS4-induced endosomes. *J Biol Chem* 2001;276:11735–11742.
43. Alsford S, Kawahara T, Glover L, Horn D. Tagging a *T. brucei* RNA locus improves stable transfection efficiency and circumvents inducible expression position effects. *Mol Biochem Parasitol* 2005;144:142–148.
44. Chen Y, Hung CH, Burderer T, Lee GS. Development of RNA interference revertants in *Trypanosoma brucei* cell lines generated with a double stranded RNA expression construct driven by two opposing promoters. *Mol Biochem Parasitol* 2003;126:275–279.
45. Allen CL, Goulding D, Field MC. Clathrin-mediated endocytosis is essential in *Trypanosoma brucei*. *EMBO J* 2003;22:4991–5002.
46. Alexander DL, Schwartz KJ, Balber AE, Bangs JD. Developmentally regulated trafficking of the lysosomal membrane protein p67 in *Trypanosoma brucei*. *J Cell Sci* 2002;115:3253–3263.
47. Jeffries TR, Morgan GW, Field MC. A developmentally regulated rab11 homologue in *Trypanosoma brucei* is involved in recycling processes. *J Cell Sci* 2001;114:2617–2626.
48. Pal A, Hall BS, Nesbeth DN, Field HI, Field MC. Differential endocytic functions of *Trypanosoma brucei* Rab5 isoforms reveal a glycosylphosphatidylinositol-specific endosomal pathway. *J Biol Chem* 2002;277:9529–9539.
49. Garrus JE, von Schwedler UK, Pornillos OW, Morham SG, Zavitz KH, Wang HE, Wettstein DA, Stray KM, Cote M, Rich RL, Myszka DG, Sundquist WI. Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. *Cell* 2001;107:55–65.
50. 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;163:237–243.
51. Hall BS, Gabernet-Castello C, Voak A, Goulding D, Natesan SK, Field MC. TbVps34, the trypanosome orthologue of Vps34, is required for Golgi complex segregation. *J Biol Chem* 2006;281:27600–27612.
52. Patron NJ, Inagaki Y, Keeling PJ. Multiple gene phylogenies support the monophyly of cryptomonad and haptophyte host lineages. *Curr Biol* 2007;17:887–891.
53. Rachel R, Wyschkony I, Riehl S, Huber H. The ultrastructure of *Ignicoccus*: evidence for a novel outer membrane and for intracellular vesicle budding in an archeon. *Archaea* 2002;1:9–18.
54. Sanderfoot A. Increases in the number of SNARE genes parallels the rise of multicellularity among the green plants. *Plant Physiol* 2007;144:6–17.
55. Bonifacino JS, Glick BS. The mechanisms of vesicle budding and fusion. *Cell* 2004;116:153–166.

56. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997;25:3389–3402.
57. Smith TF, Waterman MS. Identification of common molecular subsequences. *J Mol Biol* 1981;147:195–197.
58. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 1997;25:4876–4882.
59. Abascal F, Zardoya R, Posada D. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* 2005;21:2104–2105.
60. Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 2003;19:1572–1574.
61. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 2003;52:696–704.
62. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 2006;22:2688–2690.
63. Horii M, Shibata H, Kobayashi R, Katoh K, Yorikawa C, Yasuda J, Maki M. CHMP7, a novel ESCRT-III-related protein, associates with CHMP4b and functions in the endosomal sorting pathway. *Biochem J* 2006;400:23–32.
64. Felsenstein J. PHYLIP-Phylogeny Inference Package (version 3.2). *Cladistics* 1993;5:164–166.
65. Hirumi H, Hirumi K. Axenic culture of African trypanosome bloodstream forms. *Parasitol Today* 1994;10:80–84.
66. Wirtz E, Leal S, Ochatt C, Cross GA. A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei*. *Mol Biochem Parasitol* 1999;99:89–101.
67. Redmond S, Vadivelu J, Field MC. RNAi: an automated web-based tool for the selection of RNAi targets in *Trypanosoma brucei*. *Mol Biochem Parasitol* 2003;128:115–118.
68. Nolan DP, Jackson DG, Biggs MJ, Brabazon ED, Pays A, Van Laethem F, Paturiaux-Hanocq F, Elliott JF, Voorheis HP, Pays E. Characterization of a novel alanine-rich protein located in surface microdomains in *Trypanosoma brucei*. *J Biol Chem* 2000;275:4072–4080.