Cell PRESS

How complex is GTPase signaling in trypanosomes?

Mark C. Field and Amanda J. O'Reilly

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

Many signaling pathways in higher eukaryotes use Ras-like small GTPases. Here, we ask how complex are these small GTPase signaling pathways in trypanosomes? We seek to address this issue by comparisons with the representation of both the GTPase molecules and their accessory factors in several genomes.

Signal transduction in trypanosomes

Cells respond to stimuli principally through changes to transcription patterns and protein abundance, location and modification. These changes are coordinated through transduction pathways that originate frequently at the cell surface and are targeted towards the nucleus and other cellular sites. Transduction is mediated by multiple factors, including kinases, phosphatases and small or heterotrimeric GTPases. Depending on the connectivity among pathway components, signals can be restricted to limited outputs or they can affect a more global response. Hence, the identification of factors that constitute signaling paths and understanding their interactions are essential for assessing the scope and flexibility of cellular responses. In the context of trypanosomes, this has an important bearing on cell-cycle progression and virulence.

The trypanosomatid kinome is considerable and, in common with higher eukaryotes, many kinases can potentially be organized into cascades [1]. Although several widely conserved kinase families are represented, there is evidence for trypanosomatid-specific kinases and, significantly, representation of several common kinase accessory domains is diminished in trypanosomes. Receptor-like tyrosine kinases appear completely absent. These findings suggest a sophisticated kinase-signaling apparatus in trypanosomes with multiple features distinct from higher eukaryotes.

Small GTPases are central regulators of many processes and can be subdivided into several families, including Ras, Rho, Rab, Arf and prokaryote related. All retain a GTP-binding domain and many possess accessory domains. The major function of a small GTPase is to act as a switch, the configuration of which is governed by the guanine nucleotide state. Intrinsic GTP hydrolysis and GDP–GTP exchange are accelerated by GTPase-activating proteins (GAPs) and guanine nucleotide-exchange factors (GEFs), respectively (Figure 1). GTPases also interact with additional factors and hence consideration of GTPase function must consider the associated regulatory and signaling apparatus. Rab and Arf families are comparatively well annotated and documented in trypanosomes [2,3]. Ras and Rho function in multiple pathways in higher eukaryotes, serving to coordinate cellular responses; for example, the Rho–Cdc42–Rac system in mammals facilitates cross-talk among these three GTPases, integrating several signaling pathways [4]. In trypanosomes, there is no functional information currently concerning Ras–Rho and, therefore, the signaling complexity and integration with kinase pathways or other GTPase-mediated pathways remains unknown. By contrast, heterotrimeric GTPases are completely absent.

Overall, the picture of trypanosome signaling systems is a confusion of common and lineage-specific elements: a considerable kinase contribution, only some of which might be predicted by experimental and informatics evidence; the complete absence of a major signaling platform, that is, heterotrimeric GTPases; and an unknown contribution from the small GTPases.

A restricted family of signaling GTPases in trypanosomes

A simple count and annotation of trypanosome small GTPases is probably insufficient to appreciate fully the complexity of the signaling capacity of these molecules. Also, comparison with other organisms is an essential component in evaluating this complexity and addressing the initial question. The sizes of four major GTPase classes, selected to represent signaling, trafficking and other functions, together with the associated GAP and GEF families, were estimated (Figure 2).

The largest subfamilies of small GTPase in most species are the Rab and Arf families (Figure 2), which are involved primarily in intracellular transport. The Arf family functions principally by interaction with membrane-coat complexes and includes Arf, Arl and Sar. Evidence suggests lineage-specific evolution of the Arf subfamily [3] and the functions of kinetoplastid Arfs do not resemble closely those of mammalian Arfs [5–8]. The Arl subfamily is conserved between trypanosomes and higher eukaryotes [5]. In higher eukaryotes, Arl1, 2, 3, 6 and 8 function with the microtubule-based cytoskeleton [6] and retention of these isoforms probably reflects the importance of microtubules to trypanosomes [3,9–11]. Sar1 is involved in transport of molecules from the endoplasmic reticulum (ER) to the Golgi complex and is conserved across evolution. The Rab family, which is responsible for the coordination of vesicle transport, docking and other functions, is the largest family of small GTPases in higher eukaryotes; this is also true in trypanosomes (Figure 2). Trypanosome

Corresponding author: Field, M.C. (mcf34@cam.ac.uk).

1471-4922/$ – see front matter doi:10.1016/j.pt.2008.03.005 Available online 6 May 2008

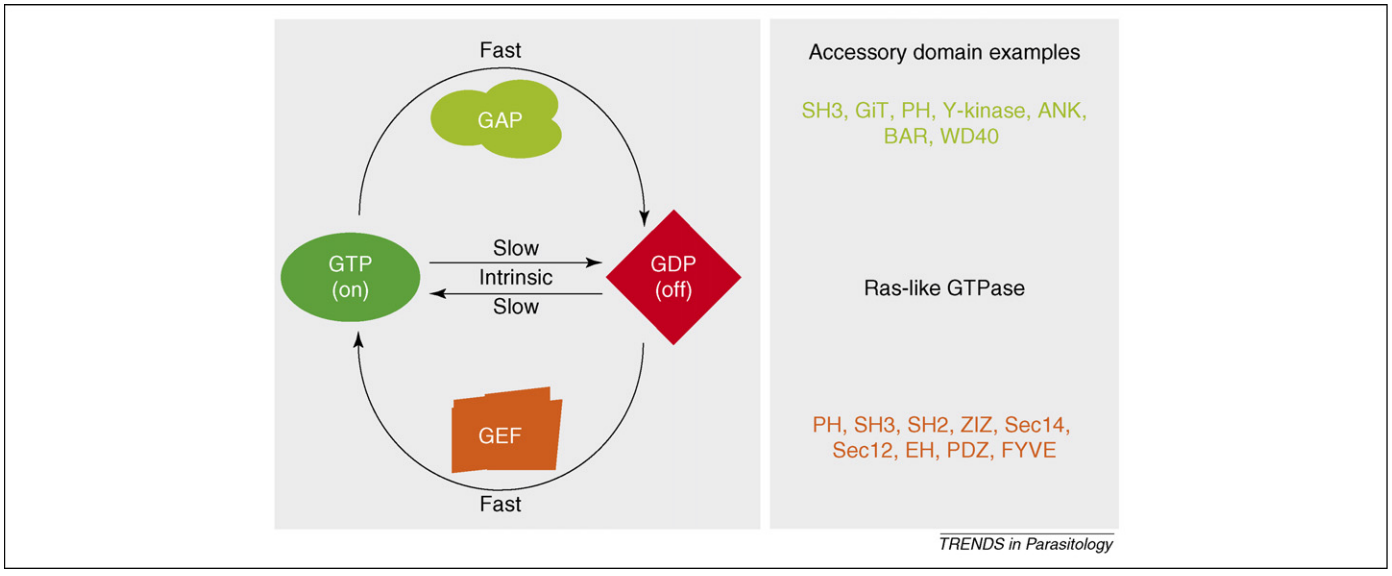

Figure 1. Small GTPase cycles and subsidiary factors. Most small GTPases function in the context of a GTP hydrolysis cycle, in which the molecule, typically, is considered active in the membrane-bound GTP-bound state (dark green) and inactive in the soluble GDP- or nucleotide-free form (red). GTP hydrolysis or GDP for GTP exchange engenders a conformational switch in the GTPase that is sensed by the cell as an altered ability to bind to cofactors or effector molecules. Typically, both intrinsic GTP hydrolysis rates and exchange of GDP for GTP in an unassisted manner are extremely slow for small GTPases. These reactions are accelerated by GTPase-activating proteins (GAPs, light green) or guanine nucleotide-exchange factors (GEFs, orange). The apparent simplicity of the GTPase cycle is overshadowed by the true diversity within these pathways – in particular, the numbers of GAPs or GEFs that can interact with a given GTPase might be large. In addition, the GAP and GEF protein families are structurally highly diverse, containing many recognized accessory domains; a selection of such accessory domains for GAP and GEF proteins are indicated; this contrasts with the high degree of structural conservation retained by the GTPase molecules. Coiled-coil domains and other structural features facilitate further interactions between the GAP and GEF molecules and additional factors.

Rabs have been discussed extensively elsewhere [2]; briefly, the majority of trypanosome Rabs are orthologous to higher eukaryote Rab proteins, although a minority have novel sequence signatures. In terms of their number, the sizes of the Rab and Arf families are similar to other unicellular organisms, for example, *Saccharomyces cerevisiae* and *Plasmodium falciparum,* and appear consistent with trypanosome cellular complexity.

Ras and Rho are responsible for a considerable fraction of the signaling system of higher eukaryotes and plants. There is no evidence for an extensive Rho system in trypanosomes (Figures 2 and 3). In *Trypanosoma cruzi,* a Rho-like GTPase with related functions to higher eukaryote Rho is present, however, there is no orthologue in *Trypanosoma brucei* or *Leishmania major* [12,13]. A divergent Rho-like protein is present in all kinetoplastids [14]. The trypanosome Ras repertoire is similarly minimized; two divergent Ras-like proteins, RLP (Ras-like protein) and RLJ (Ras-like protein with J-domain), are present, although these are clearly distinct from classical Ras proteins [14]. Hence, there is significant divergence among species in the distributions of small GTPases. In Metazoa, *S. cerevisiae* and *Dictyostelium discoideum*, there is a sizeable Ras and Rho family. These families are much smaller in trypanosomes (Figures 2 and 3).

Low complexity in trypanosome GTPase cycles

The GAP and GEF factors, being considerably more heterogeneous than the GTPases, are considered most conveniently by functional and structural class.

The major Arf family GAPs and GEFs are defined by ArfGAP and Sec7 domains, respectively. In higher eukaryotes, these families are extensive, with 27 ArfGAPs and 16 Sec7 domain proteins in *Homo sapiens* [5,15] (Figure 2). There are, thus, sufficient family members for high specificity among GAPs, GEFs and their respective Arf and Arl GTPases. Substantially fewer ArfGAP- and Sec7-containing proteins are encoded in the trypanosome genome compared with the Arfs themselves. A restricted ArfGAP and ArfGEF configuration is also found in additional taxa (Figure 2). Sec7-domain proteins are divided into high and low molecular-weight families; both are present in metazoa but only the high molecular weight class is found in trypanosomes. This pattern suggests preferential subfamily expansion and is clearly insufficient to provide the level of specificity observed in metazoa.

The ratio of detectable ARFs to ArfGEFs to ArfGAPs in trypanosomes is 1:0.2:0.4 and, in humans, the ratio is 1:0.6:1, suggesting considerable GTPase promiscuity within trypanosome ArfGAP and GEFs or the presence of novel Arf-control factors. A similarly low number of ArfGAP and GEFs is found in *P. falciparum* and *Tetrahymena thermophila*, suggesting that this feature is general among protist lineages. Therefore, despite the considerable size of the trypanosome Arf and Arl families, the small number of GAP and GEFs suggests restricted signaling flexibility.

Over 90% of known RabGAP proteins contain the Tre2/Bub2/Cdc16 (TBC) domain [16]; exceptions include Rab3A-GAP [17]; a PtdIns 3-kinase and Rab4, 5 GAP [18]; and tuberin, which is a Rab5 GAP [19]. Few TBC GAPs are characterized functionally [19]. There are typically a similar number of TBC proteins and Rabs in most organisms. In *S. cerevisiae*, with 11 Rabs and six TBCs, TBC specificity is low because they can interact with multiple Rabs [20]. With 14 TBC RabGAPs and 16 Rabs in *T. brucei,* the complexity of Rab versus RabGAP is probably retained across evolution (Figure 2). A potential orthologue for

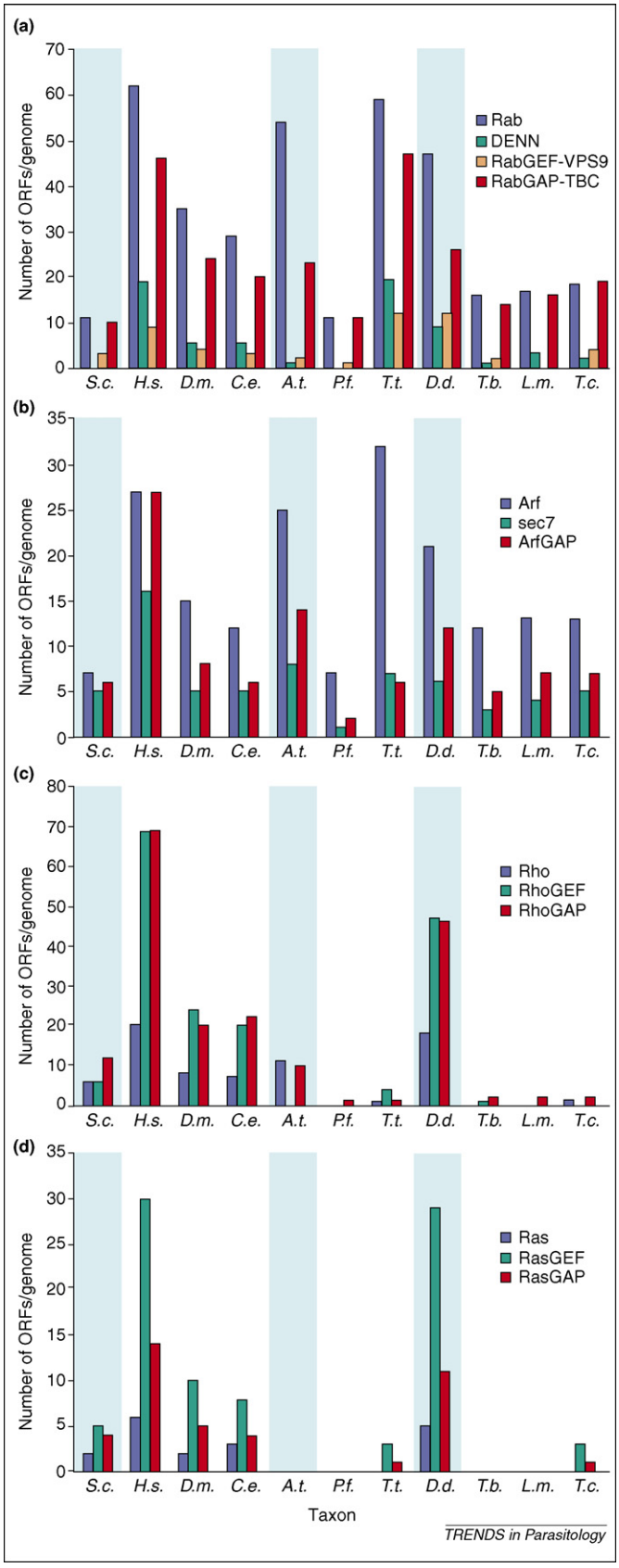

Figure 2. GTPase families in selected taxa. **(a)** Rab family, **(b)** Arf family, **(c)** Rho family and **(d)** Ras family show the total number of small GTPase and major GAP and GEF protein families returned from searches of the *Saccharomyces cerevisiae, Homo sapiens, Drosophila melanogaster, Caenorhabditis elegans, Arabidopsis thaliana, Plasmodium falciparum, Tetrahymena thermophila, Dictyostelium discoideum, Trypanosoma brucei, Leishmania major* and *Trypanosoma cruzi* predicted proteomes. In each case, the GTPase is in blue, GEFs are in green or orange and GAPs are in red. Note that the apparent absence of the RhoGEF system from *A. thaliana* is due to substitution of this family by the distinct PRONE domain. Datasets were extracted from UniProt, except *Tetrahymena thermophila*, which was from The Institute for Genomic Research. Known GEF and GAP sequences were collected from Uniprot by keyword, GO (gene ontology) and Pfam annotation; this set was extended by subsequent database searching. Assignments were validated using BLAST (basic local alignment search tool) and domain prediction. Additional information was obtained from [20,26,29,31,32]. Light blue shading is to aid in visualization.

Rab3AGAP is present unexpectedly in *T. brucei*, even though there is no evidence for regulated exocytosis in trypanosomes [2,3], however, neither the PtdIns 3-kinase nor tuberin GAPs are represented in *T. brucei*. For Rab5, the GAP repertoire might be more limited in trypanosomes than higher eukaryotes.

The RabGEF family is heterogenous and many RabGEFs probably remain to be described. Several RabGEFs (e.g. Sec2, Ric1/Rgp1 and Ccz1) are only known from *S. cerevisiae* [21–23] and Mss4 is restricted to Opisthokonts. Vps9-domain-containing RabGEF candidates are widespread evolutionarily but are few compared with the number of Rabs [24]; nine in *H. sapiens* and possibly two in *T. brucei*. The Vps9 family includes Rabex5, a major Rab5-GEF in higher eukaryotes; however, neither of the *T. brucei* Vps9 proteins are closely related to Rabex5. A differentially expressed in neoplastic versus normal cells (DENN) domain-containing candidate RabGEF and two conserved GEFs, the transport protein particle (TRAPP) complex I and homotypic fusion and vacuole protein sorting (HOPS) complex, which in *S. cerevisiae* are GEFs for Ypt1p and Ypt7p, respectively, are also present in trypanosomes [25]. Overall, although there are some clear cases of divergence, the RabGAP and GEF repertoire is well conserved between trypanosomes and higher eukaryotes and is probably of similar complexity.

Finally, the Ras and Rho GAPs and GEFs are extremely reduced in trypanosomes. *H. sapiens* has 20 Rho proteins, approximately 70 potential Rho GAPs and approximately 80 potential Rho GEFs [26,27]; such complexity is due, in part, to tissue-specific expression; however, it also indicates the extreme integration of Rho-dependent signaling pathways with many cellular events. In Metazoa, RhoGAPs are defined by the presence of the RhoGAP domain [26] and RhoGEFs, principally either the RhoGEF (DH) or DOCK180 (dedicator of cytokinesis 1) domains [23]. In plants, RopGEFs possess the plant-specific ROP nucleotide exchanger (PRONE) domain [27]. The presence of a divergent Rho-like protein in *T. brucei* is consistent with predicted RhoGAP domain-containing proteins in trypanosomes. One putative *T. brucei* RhoGAP (Tb09.160.4180) shares a conserved additional phosphatase domain with *D. discoideum* and metazoa. A single RhoGEF domain-containing protein is also found in *T. brucei*. Notwithstanding, this putative Rho-like system of trypanosomes is highly reduced and is unlikely to contribute in a major fashion to trypanosome signaling. The Ras system in higher eukaryotes is less extensive than Rho but, again, potential RasGAPs and RasGEFs are recognized readily by the presence of specific domains, the RasGAP and RasGEF domains, respectively. Apart from the presence of divergent Ras-like GTPases in the trypanosome genome [14] and a small number of divergent

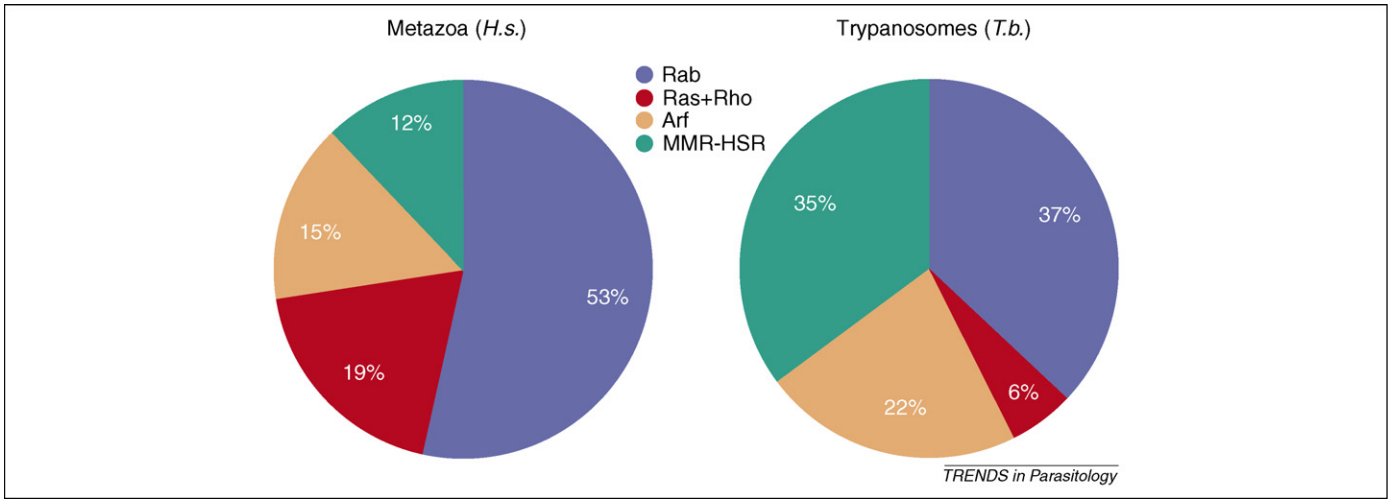

Figure 3. Contributions of small GTPase subfamilies to the total repertoire in Metazoa and trypanosomes. Percentage representation for each family is based principally on counts derived from *H. sapiens* and *T. brucei*. Colors are as follows: blue, Rab; red, combined Ras and Rho; yellow, Arf family; green, MMR-HSR. Predicted proteomes were searched for small GTPase and corresponding GAP and GEF families.

RasGAP and RasGEF domain proteins in *T. cruzi*, the remaining components of the Ras GTPase cycle are absent. Significantly, TCTP (translationally controlled tumor protein), an orthologue of a probable RhebGEF from *Drosophila* [28] is present (Rheb is a further distant Ras-related GTPase) and has been duplicated in both the *T. brucei* and *L. major* genomes.

Modest representation of prokaryote-related and other small GTPases in trypanosomes

Several additional categories of small GTPases are present in the trypanosome genome [3] (accessions available at http://homepage.mac.com/mfield/lab/publications.html). Two nuclear-localized GTPase subfamilies, GRP (Gtr/Rag-1-like protein) and nucleostemin, are conserved across eukaryotes but their precise functions remain unknown [14]. An MMR_HSR1 (Pfam 01926) family, shared with prokaryotes, is well represented; this includes proteins bearing GTPase domains that fall into the EngA, EngB, Mmr-Hsr1, Nog, Ngp and Obg subfamilies and, for each class, there is at least one trypanosome representative. In higher eukaryotes, the presence of these molecules is somewhat overshadowed by expansion of the Ras and Rho subfamilies, however, in trypanosomes, this ancient family makes a substantial contribution to the total GTPase repertoire (Figure 3). This group and nucleostemin can be differentiated from classical Ras GTPases by a divergent ordering of the conserved GTP-binding motifs, specifically G4, 1, 2 and 3, rather than the conventional G1, 2, 3 and 4 [29]. Significantly, the presence of a nucleostemin in trypanosomes disproves a hypothesis that nucleostemins arose with the deuterostomes [30].

There are also trypanosome-specific GTPases. TbFRP contains both a PtdIns 3-phosphate-binding Fab1/YOTB/Vac1/EEA1 (FYVE) domain as well as a GTPase domain [14]; this combination is not found outside the kinetoplastida. In addition, three orphan GTPases (RX1–3) result from a *T. brucei*-specific duplication and the trypanosomatid common ancestor contained a single RX2/3 gene. Because there are limited functional data for this final category, there is no information concerning their GAPs and GEFs.

Conclusion

How complex is GTPase signaling in trypanosomes? Although fuller functional analysis is required for a clear answer, a provisional conclusion is possible. From comparisons of the sizes and compositions of the small GTPase superfamily and their associated GAPs and GEFs between trypanosomes and other species, there is apparent decreased representation of Ras- and Rho-like subfamilies in trypanosomes but with retention of Rab and Arf subfamily size. There are also several unusual and prokaryote-related GTPases, many conserved with higher eukaryotes. Alone, this is sufficient evidence to predict a partial de-emphasis on classical GTPase-mediated signaling in trypanosomes, while maintaining complexity within the trafficking system.

Acknowledgements

This work was supported by a program grant from the Wellcome Trust. We are grateful to the sequencing projects for making data freely available, to the Cambridge eScience Centre for support of CamGRID, to Joel Dacks, Lila Koumandou and Ka-Fai Leung for comments on the manuscript. Accession numbers for trypanosome small GTPases, GAPs and GEFs are available at http://homepage.mac.com/mfield/lab/publications.html.

References

1 Parsons, M. *et al.* (2005) Comparative analysis of the kinomes of three pathogenic trypanosomatids: *Leishmania major, Trypanosoma brucei* and *Trypanosoma cruzi*. *BMC Genomics* 6, 127
2 Ackers, J.P. *et al.* (2005) A bioinformatic analysis of the RAB genes of *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* 141, 89–97
3 Berriman, M. *et al.* (2005) The genome of the African trypanosome *Trypanosoma brucei*. *Science* 309, 416–422
4 Hall, A. (2005) Rho GTPases and the control of cell behaviour. *Biochem. Soc. Trans.* 33, 891–895
5 Li, Y. *et al.* (2004) Functional genomic analysis of the ADP-ribosylation factor family of GTPases: phylogeny among diverse eukaryotes and function in *C. elegans*. *FASEB J.* 18, 1834–1850
6 Burd, C.G. *et al.* (2004) Arf-like GTPases: not so Arf-like after all. *Trends Cell Biol.* 14, 687–694
7 Price, H.P. *et al.* (2005) Functional analysis of TbARL1, an N-myristoylated Golgi protein essential for viability in bloodstream trypanosomes. *J. Cell Sci.* 118, 831–841
8 Price, H.P. *et al.* (2007) *Trypanosoma brucei* ARF1 plays a central role in endocytosis and Golgi-lysosome trafficking. *Mol. Biol. Cell* 18, 864–873

9 Cuvillier, A. *et al.* (2000) LdARL-3A, a *Leishmania* promastigote-specific ADP-ribosylation factor-like protein, is essential for flagellum integrity. *J. Cell Sci.* 113, 2065–2074
10 Cuvillier, A. *et al.* (2003) Abortive infection of *Lutzomyia longipalpis* insect vectors by aflagellated LdARL-3A-Q70L overexpressing *Leishmania amazonensis* parasites. *Cell. Microbiol.* 5, 717–728
11 Sahin, A. *et al.* (2004) Trypanosomatid flagellum biogenesis: ARL-3A is involved in several species. *Exp. Parasitol.* 108, 126–133
12 De Melo, L.D. *et al.* (2004) TcRho1 of *Trypanosoma cruzi*: role in metacyclogenesis and cellular localization. *Biochem. Biophys. Res. Commun.* 323, 1009–1016
13 De Melo, L.D. *et al.* (2006) TcRho1, the *Trypanosoma cruzi* Rho homologue, regulates cell-adhesion properties: evidence for a conserved function. *Biochem. Biophys. Res. Commun.* 345, 617–622
14 Field, M.C. (2005) Signalling the genome: the Ras-like small GTPase family of trypanosomatids. *Trends Parasitol.* 21, 447–450
15 Jiang, S.Y. and Ramachandran, S. (2006) Comparative and evolutionary analysis of genes encoding small GTPases and their activating proteins in eukaryotic genomes. *Physiol. Genomics* 24, 235–251
16 Rak, A. *et al.* (2000) Crystal structure of the GAP domain of Gyp1p: first insights into interaction with Ypt/Rab proteins. *EMBO J.* 19, 5105–5113
17 Nagano, F. *et al.* (1998) Molecular cloning and characterization of the non-catalytic subunit of the Rab3 subfamily-specific GTPase-activating protein. *J. Biol. Chem.* 273, 24781–24785
18 Chamberlain, M.D. *et al.* (2004) The p85 α-subunit of phosphatidylinositol 3′-kinase binds to and stimulates the GTPase activity of Rab proteins. *J. Biol. Chem.* 279, 48607–48614
19 Xiao, G.H. *et al.* (1997) The tuberous sclerosis 2 gene product, tuberin, functions as a Rab5 GTPase activating protein (GAP) in modulating endocytosis. *J. Biol. Chem.* 272, 6097–6100
20 Albert, S. and Gallwitz, D. (1999) Two new members of a family of Ypt/Rab GTPase activating proteins. Promiscuity of substrate recognition. *J. Biol. Chem.* 274, 33186–33189
21 Walch-Solimena, C. *et al.* (1997) Sec2p mediates nucleotide exchange on Sec4p and is involved in polarized delivery of post-Golgi vesicles. *J. Cell Biol.* 137, 1495–1509
22 Siniossoglou, S. *et al.* (2000) Ric1p and Rgp1p form a complex that catalyses nucleotide exchange on Ypt6p. *EMBO J.* 19, 4885–4894
23 Kucharczyk, R. *et al.* (2001) The Ccz1 protein interacts with Ypt7 GTPase during fusion of multiple transport intermediates with the vacuole in *S. cerevisiae*. *J. Cell Sci.* 114, 3137–3145
24 Delprato, A. and Lambright, D.G. (2007) Structural basis for Rab GTPase activation by VPS9 domain exchange factors. *Nat. Struct. Mol. Biol.* 14, 406–412
25 Koumandou, V.L. *et al.* (2007) Control systems for membrane fusion in the ancestral eukaryote; evolution of tethering complexes and SM proteins. *BMC Evol. Biol.* 7, 29
26 Tcherkezian, J. and Lamarche-Vane, N. (2007) Current knowledge of the large RhoGAP family of proteins. *Biol. Cell* 99, 67–86
27 Rossman, K.L. and Sondek, J. (2005) Larger than Dbl: new structural insights into RhoA activation. *Trends Biochem. Sci.* 30, 163–165
28 Hsu, Y.C. *et al.* (2007) *Drosophila* TCTP is essential for growth and proliferation through regulation of dRheb GTPase. *Nature* 445, 785–788
29 Reynaud, E.G. *et al.* (2005) Human Lsg1 defines a family of essential GTPases that correlates with the evolution of compartmentalization. *BMC Biol.* 3, 21
30 Bateman, A. *et al.* (2004) The Pfam protein families database. *Nucleic Acids Res.* 32, D138–D141
31 Bernards, A. (2005) Ras superfamily and interacting proteins database. *Methods Enzymol.* 407, 1–9
32 Rutherford, S. and Moore, I. (2002) The *Arabidopsis* Rab GTPase family: another enigma variation. *Curr. Opin. Plant Biol.* 5, 518–528