13 Chun, J. *et al*. (1995) Effect of mosquito age and reproductive status on melanization of Sephadex beads in *Plasmodium*-refractory and -susceptible strains of *Anopheles gambiae*. *J. Invertebr. Pathol.* 66, 11–17
14 Tabachnick, W.J. (2003) Reflections on the *Anopheles gambiae* genome sequence, transgenic mosquitoes and the prospect for controlling malaria and other vector borne diseases. *J. Med. Entomol.* 40, 597–606
15 Tahar, R. *et al*. (2002) Immune response of *Anopheles gambiae* to the early sporogonic stages of the human malaria parasite *Plasmodium falciparum*. *EMBO J.* 21, 6673–6680
16 Ito, J. *et al*. (2002) Transgenic anopheline mosquitoes impaired in transmission of a malaria parasite. *Nature* 417, 452–455

1471-4922/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.pt.2005.08.012

Genome Analysis

Signalling the genome: the Ras-like small GTPase family of trypanosomatids

Mark C. Field

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

The genomes of the three principle experimental-model species of Kinetoplastida – *Trypanosoma brucei brucei, Trypanosoma cruzi* and *Leishmania major* – are now complete, providing both a milestone for trypanosome biology and an opportunity to consider a multitude of questions at the genome level. Of the >40 members of the Ras-like GTPase family in *T. brucei*, at least 30 are involved in intracellular transport, whereas fewer than eight are likely to have a classical role in signal transduction. There are no true members of the Ras or Rho subfamilies but divergent Ras- or Rho-like GTPases are present, suggesting that signalling mechanisms in trypanosomatids are highly unusual. Comparisons of *T. brucei* with *T. cruzi* and *L. major* indicate a high degree of conservation among the species. These analyses provide a framework for the functional investigation of small-GTPase-mediated signalling processes in trypanosomes.

Trypanosomes: divergence and differentiation

Trypanosomes occupy a highly divergent position within the tree of life, and separated from the metazoan lineage three billion years ago. Many aspects of these organisms are extremely unusual, and the recent completion of the genome projects for three trypanosomatid species offers a superb opportunity to investigate the mechanisms available to trypanosomes at the molecular level and to carry this information forward into a true postgenomic era [1]. The comparative experimental accessibility of *Trypanosoma brucei brucei* also offers an attractive system for evolutionary cell biology [2]. One particular aspect of trypanosomatid biology that has not fully revealed itself is the mechanisms that are used for signal transduction and, by implication, developmental progression. The *T. brucei* life cycle consists of multiple developmental stages, with at least one proliferative stage in the mammalian and arthropod hosts: the bloodstream form (BSF) and the procyclic form (PCF), respectively. The BSF differs radically from the PCF, with alterations to cell structure, metabolism, surface-antigen expression and membrane-trafficking systems [3]. At least one terminally differentiated stage, the metacyclic form in the tsetse salivary glands, serves to complete the life cycle; a second such stage in the mammal, the stumpy form, might also be important for transmission [4]. Life-cycle progression requires environmental sensing; the BSF-to-PCF differentiation is stimulated by temperature and Krebs cycle intermediates, but the molecular mechanisms remain elusive. Molecular sequelae accompanying differentiation from BSF to PCF have been described but, again, their functional basis is unknown.

Gene expression and signalling

Gene expression in trypanosomatids is predominantly polycistronic, with subsequent resolution of nascent RNAs occurring through *trans*-splicing and polyadenylation. This mechanism removes the influence of promoters as mechanisms for controlling mRNA abundance, with major potential impact on signalling systems. In higher eukaryotes, signalling pathways are ultimately directed towards modulating gene expression, the final target being RNA polymerase, through influencing transcription factor recruitment to promoter elements. Apart from some exceptions, such a system is absent from trypanosomes. Attempts to investigate signalling by ablation of trypanosome paralogues of signalling factors have met with limited success [5], and few defined signals that result in specific alteration of protein expression have been described. Inhibition of transferrin accumulation is accompanied by rapid upregulation of the receptor protein [6], suggesting coupling between intracellular iron concentration and mRNA abundance (and/or translation efficiency), whereas expression of procyclin is sensitive to glycerol, suggestive of a signalling system [7]; however, a signalling pathway has not been elucidated for any of these examples. Perhaps most striking is the complete lack of heterotrimeric GTPases.

Corresponding author: Field, M.C. (mcf34@cam.ac.uk).
Available online 19 August 2005

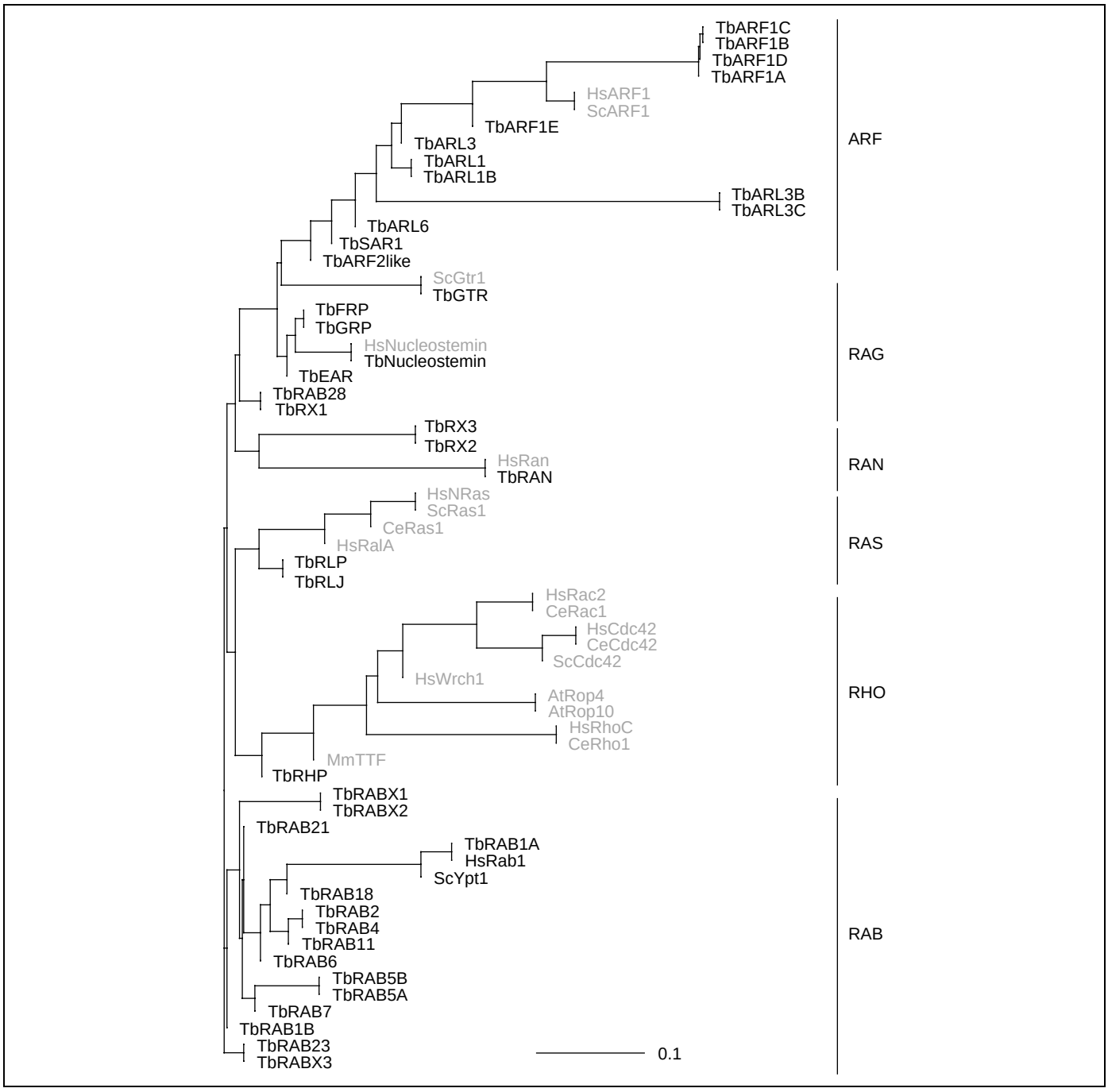

Figure 1. Phylogenetic reconstruction of all small GTPases from *Trypanosoma brucei brucei*. Trypanosome proteins are shown in black, designated by their proposed systematic names, and non-trypanosomal proteins used to assign the gene families (from *Homo sapiens*, *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, *Arabidopsis thaliana* and *Mus musculus*) are shown in grey. Gross Ras subfamily groupings are shown on the right. The reconstruction faithfully reproduces the small GTPase subfamilies, together with several unique and deeply divergent clades populated by trypanosome sequences. The genetic distance is shown as a line. Tree was generated with ClustalX and the neighbour-joining method, together with 10 000 bootstrap replicates. Highly similar trees were obtained using other methods, indicating that the tree is robust. (See supplementary material online for details of GeneDB accession numbers.)

Adenylate cyclase, the downstream target of several higher-eukaryote heterotrimeric GTPases, is fused directly to transmembrane receptors in trypanosomes, suggesting that GTPase-mediated receptor coupling is circumvented [8]. Overall, it seems that trypanosomatids might have evolved distinctive signalling systems.

The small GTPase family of *Trypanosoma brucei*

The small GTPases are highly conserved and comprise several subfamilies that are responsible for mediating signal transduction (Ras, Rho and Cdc42), cytoskeletal control [Rho, Cdc42, Rac and ADP ribosylation factors (ARFs)] and transport processes (Rabs, Ran and ARF). All small GTPases interact with effector molecules that propel the GTPases through a cycle of GTP binding, hydrolysis and exchange. Despite the apparent functional discrimination that accompanies the sequence specificity inherent within subfamily groupings [9], all of these molecules function as information transducers.

The *T. brucei* genome encodes 41 small GTPases, the vast majority of which has clear orthologues in *Trypanosoma cruzi* and *Leishmania major*. The largest class is the Rab family – with 16 members that control vesicle transport [10] – followed by ARF, with 12 members [1]. Together with Ran, which mediates nucleocytoplasmic transport, and Sar1, which controls the recruitment of coat proteins to the endoplasmic reticulum, >70% of the trypanosome small-GTPase repertoire is, therefore, devoted to transport processes. Domain structures and phylogenetic relationships of the small GTPase family are shown in Figure 1 and Table 1, and a systematic nomenclature is proposed later and in the supplementary material online. There is no true Ras homologue, and although two members of the trypanosome family segregate into the Ras clade both diverge at the base of the clade and lack critically conserved residues (Figure 1). Furthermore, the Rho clade is also absent, and the trypanosomal protein TbRHP is present only within the base of the clade. No paralogous clade replaces Rho as does the Rop family in higher plants [11]. These observations effectively rule out phagocytic activity in these organisms by currently understood mechanisms.

New GTPases identified from the genome

Rho clade

There is only one member of this clade: TbRHP. The protein has a long C terminus, which discriminates it from a Rho family GTPase, but it retains a CVIM prenylation motif and an RK-rich region close to the C terminus, which is suggestive of a membrane-targeting signal.

Ras clade

Two trypanosomal sequences are found at the base of this clade: TbRLP and TbRLJ. The former was identified as a Ras candidate [12] and is, indeed, the most Ras-like in the genome, containing both a prenylation motif and a C-terminal lysine cluster that are common in Ras proteins. However, TbRLP is of sufficient divergence for functional assignment to be equivocal without additional data (Figure S1 in the supplementary material online). TbRLJ is a short protein that terminates in AAFM and that is most homologous to a divergent group of Ras-like molecules with a C-terminal DnaJ domain [13,14]. TbRLJ lacks this domain and is, therefore, unlikely to have a similar function (Figure S2 in the supplementary material online).

Table 1. Structure and function of the trypanosome Ras-like GTPases[a]

Name	No. of family members	Predicted domain structure[b,c]	Function
Sar1	1		Transport
Ran	1		Transport or signalling
RLJ	1		Signalling?
RX1	1		Not known
EAR	1		Not known
ARF and ARL	12		Transport or cytoskeleton
RAB	16		Transport
RLP	1		Signalling?
GTR and GRP	2		Transport
RHP	1		Not known
RX2 and RX3	2		Not known
FRP	1		Not known
NST	1		Differentiation

[a]Proposed systematic names for the entire small GTPase family are given, together with the number of family members encoded by the genome (see Tables S1 and S2 in supplementary material online). In some cases, a function is either known from empirical studies or can be confidently inferred based on sequence similarity with other trypanosome family members or orthologues in other systems.
[b]Structures: red oval, GTPase domain; yellow circle, myristoylation motif; blue circle, prenylation motif; green square, FYVE domain; brown square, nuclear-targeting or p53-interaction domain; blue square, acidic domain.
[c]Regions without clear structural features are shown as black bars.

Ran clade

In addition to TbRAN, this clade contains two closely related GTPases with long C termini that terminate in prenylation signals. The absence of database hits of significance and the apparent unusual domain structure suggest that these are novel proteins and, hence, they have been named RX2 and RX3, for Ras-like unknowns.

Rag clade

This clade contains the largest number of new members of the trypanosome family, but no clear functional emphasis is apparent. TbRX1 is extremely similar to Rab family GTPases but several vital motifs are mutated and, most importantly, the Rab-diagnostic WDTAGQ box is represented by FDVSGQ and there is no prenylation signal. Interestingly, TbRAB28 is extremely similar to TbRX1, raising doubts as to the true status of the former as a Rab.

Four proteins are recovered in a subclade that is defined by nucleostemin – an evolutionarily conserved protein that has a role in differentiation and interacts with p53 in mammals [15]. There is a clear nucleostemin orthologue in trypanosomatids: TbNST, which retains a domain structure comprising a central GTPase with C- and N-terminal extensions, and multiple KR-rich nuclear-localization signals in the N-terminal region (Figure S5 in the supplementary material online). In humans, nucleostemins are associated with an undifferentiated cell state, being expressed in the nucleolus of stem cells and in many cancers but not in normal differentiated tissues [15].

TbEAR is a member of a conserved protein family containing two tandem GTPase domains (Figures S3 and S4 in the supplementary material online). The TbEAR family is similar to a family of proteins from proteobacteria that includes the essential EngA proteins [16]. The tandem GTPase arrangement seems to be restricted to proteobacteria, protozoa and metazoans [17,18] (Table S2 in the supplementary material online).

TbFRP contains a C-terminal GTPase domain and an N-terminal FYVE domain. The coupling of a FYVE domain to a Ras-like GTPase seems to be unique to trypanosomatid systems and suggests that TbFRP might be recruited to phosphatidylinositol (3,4,5)-trisphosphate (PIP_3)-containing membranes (which are endosomal in higher eukaryotes) by the FYVE domain, with the GTPase mediating a signalling function. The final member of this clade is TbGRP, which has a conserved domain structure and significant homology with TbGTR.

ARF clade

Besides the ARF, ARL and Sar1 homologues, this clade also contains TbGTR, an orthologue of yeast Gtr1p. Gtr proteins contain a conserved C-terminal extension and lack a prenylation motif. In yeast, Gtr1 is involved in nucleocytoplasmic transport and interacts with RCC1, which is also an interaction partner of Ran. Gtr1 also forms a complex with the related protein Gtr2 and itself [19]. Despite some divergence, it is possible that TbGTR and TbGRP associate, functioning in a coordinated manner. In humans, the Gtr family comprises at least four members, all of which seem to interact with each other [20].

Future perspectives

The great similarity among the GTPase complements in *T. brucei*, *T. cruzi* and *L. major* [1,10] suggests that there are no radical differences among these systems and, perhaps more importantly, that most data obtained from one organism can be applied to the others with a high degree of confidence. Any differences are, of course, of special interest. A lack of major redundancy in these systems also suggests a comparatively simplistic signalling system in trypanosomes and/or a de-emphasis on GTPase-mediated signalling pathways. Several absentees, principally Ras and Rho, have major implications and confirm suspicions that signalling pathways in kinetoplastids are radically different from those in metazoans. This is contrasted by the high degree of similarity among the Rab and ARF families and their counterparts in higher eukaryotes. Hence, a directed investigation of signalling in these parasites is not only of academic interest as a new model system but also vital for understanding the manner in which they can respond to changing environments and undergo developmental changes that accompany differentiation. Furthermore, the distinct nature of the putative signalling GTPases suggests an independent evolutionary origin of some of these pathways. Several rather unusual GTPases have also been detected, and the functions of these cannot be predicted based on *in silico* studies alone but, instead, require direct experimentation. Perhaps this is the major finding; there is much novel and highly informative biology within the trypanosome genomes. The time has now come for the genomes to be exploited fully.

Acknowledgements

I thank the Sanger Institute and the Institute for Genomic Research for providing kinetoplastid sequence data.

Supplementary data

Supplementary data associated with this article can be found at doi:10.1016/j.pt.2005.08.008

References

1 Berriman, M. *et al*. (2005) The genome of the African trypanosome *Trypanosoma brucei*. *Science* 309, 416–422
2 Dacks, J.B. and Field, M.C. (2004) Eukaryotic cell evolution from a comparative genomic perspective: the endomembrane system. In *Organelles, Genomes and Eukaryote Phylogeny: an Evolutionary Synthesis in the Age of Genomics* (Hirt, R. and Horner, D., eds), pp. 309–334, Taylor & Francis
3 Field, M.C. and Carrington, M. (2004) Intracellular membrane transport systems in *Trypanosoma brucei*. *Traffic* 5, 905–913
4 Matthews, K.R. *et al*. (2004) Molecular regulation of the life cycle of African trypanosomes. *Trends Parasitol.* 20, 40–47
5 Muller, I.B. *et al*. (2002) Stage-specific requirement of a mitogen-activated protein kinase by *Trypanosoma brucei*. *Mol. Biol. Cell* 13, 3787–3799
6 Mussmann, R. *et al*. (2004) Factors affecting the level and localization of the transferrin receptor in *Trypanosoma brucei*. *J. Biol. Chem.* 279, 40690–40698
7 Vassella, E. *et al*. (2000) A major surface glycoprotein of *Trypanosoma brucei* is expressed transiently during development and can be regulated post-transcriptionally by glycerol or hypoxia. *Genes Dev.* 14, 615–626
8 Seebeck, T. *et al*. (2004) cAMP signalling in the kinetoplastid protozoa. *Curr. Mol. Med.* 4, 585–599
9 Pereira-Leal, J.B. and Seabra, M.C. (2001) Evolution of the Rab family of small GTP-binding proteins. *J. Mol. Biol.* 313, 889–901
10 Ackers, J.P. *et al*. (2005) A bioinformatic analysis of the RAB genes of *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* 141, 89–97
11 Li, H. *et al*. (2001) The Rop GTPase switch controls multiple developmental processes in *Arabidopsis*. *Plant Physiol.* 126, 670–684
12 Sowa, M.P. *et al*. (1999) A novel gene encoding a *ras*-like GTP-binding protein from *Trypanosoma brucei*: an evolutionary ancestor of the *ras* and *rap* genes of higher eukaryotes? *Gene* 230, 155–161
13 Nepomuceno-Silva, J.L. *et al*. (2004) RJLs: a new family of Ras-related GTP-binding proteins. *Gene* 327, 221–232
14 Nepomuceno-Silva, J.L. *et al*. (2004) Characterization of *Trypanosoma cruzi* TcRjl locus and analysis of its transcript. *Parasitology* 129, 325–333
15 Tsai, R.Y. and McKay, R.D. (2002) A nucleolar mechanism controlling cell proliferation in stem cells and cancer cells. *Genes Dev.* 16, 2991–3003
16 Mittenhuber, G. (2001) Comparative genomics of prokaryotic GTP-binding proteins (the Era, Obg, EngA, ThdF (TrmE), YchF and YihA families) and their relationship to eukaryotic GTP-binding proteins (the DRG, ARF, RAB, RAN, RAS and RHO families). *J. Mol. Microbiol. Biotechnol.* 3, 21–35
17 Caldon, C.E. and March, P.E. (2003) Function of the universally conserved bacterial GTPases. *Curr. Opin. Microbiol.* 6, 135–139
18 Robinson, V.L. *et al*. (2002) Domain arrangement of Der, a switch protein containing two GTPase domains. *Structure (Camb)* 10, 1649–1658
19 Nakashima, N. *et al*. (1999) *Saccharomyces cerevisiae* putative G protein, Gtr1p, which forms complexes with itself and a novel protein designated as Gtr2p, negatively regulates the Ran/Gsp1p G protein cycle through Gtr2p. *Genetics* 152, 853–867
20 Sekiguchi, T. *et al*. (2001) Novel G proteins, Rag C and Rag D, interact with GTP-binding proteins, Rag A and Rag B. *J. Biol. Chem.* 276, 7246–7257

1471-4922/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.pt.2005.08.008