Ras superfamily GTPases and signal transduction in Euglena gracilis

Field, Mark C.

Published in:
Protist

DOI:
10.1016/j.protis.2024.126017

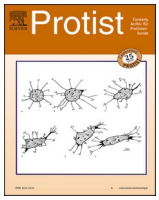
Publication date:
2024

Protist 175 (2024) 126017

Contents lists available at ScienceDirect

Protist

journal homepage: www.elsevier.de/protis



Original article

Ras superfamily GTPases and signal transduction in *Euglena gracilis*

Mark C. Field

School of Life Sciences, University of Dundee, Dundee, UK
Biology Centre, Czech Academy of Sciences, České Budějovice, Czech Republic

ARTICLE INFO

Monitoring Editor: Michael Ginger

Keywords:
Signal transduction
Euglena
Diversity
Biological complexity
Small GTPases
Ras

ABSTRACT

Biological complexity is challenging to define, but can be considered through one or more features, including overall genome size, number of genes, morphological features, multicellularity, number of life cycle stages and the ability to adapt to different environments. *Euglena gracilis* meets several of these criteria, with a large genome of ~38,000 protein coding genes and a considerable ability to survive under many different conditions, some of which can be described as challenging or harsh. Potential molecular exemplars of complexity tying these aspects together are signalling pathways, including GTPases, kinases and ubiquitylation, which increase the functionality of the gene-encoded proteome manyfold. Each of these examples can modulate both protein activity and gene expression. To address the connection between genome size and complexity I have undertaken a brief, and somewhat qualitative, survey of the small ras-like GTPase superfamily of *E. gracilis*. Unexpectedly, apart from Rab-GTPases which control intracellular transport and organelle identify, the size of the GTPase cohort is modest, and, for example, has not scaled with gene number when compared to the close relatives, trypanosomatids. I suggest that understanding the functions of this protein family will be vital to uncovering the complexity of *E. gracilis* biology.

1. Introduction

Euglena gracilis is notable for multiple features, including harbouring a secondary endosymbiont plastid, being capable of exploiting a considerably broad environmental range and possessing of a large and complex genome (Ebenezer et al., 2019). *E. gracilis* also utilises multiple forms of RNA splicing, including combining *cis*- and *trans*-spicing for transcript maturation and a large repertoire of protein coding genes that exceeds 38000. In part due to the large genome the true genetic complexity remains to be revealed, but it is clear that it is considerable, with gene models predicting large introns and complex gene structure, alternate splicing and post-translational control of protein expression uncoupled from RNA abundance (Tessier et al., 1991; McWatters and Russell, 2017; Gumińska et al., 2018; Ebenezer et al., 2019). Furthermore, understanding how this large genome is controlled is central to both biotechnological exploitation and optimisation as well as unravelling the complex contributions that euglenoids make to global ecosystems (Ebenezer et al., 2022).

Ras-like small guanosine triphosphatases (GTPases) are 21–30 kDa monomeric guanine nucleotide-binding proteins and related to the α-subunits of heterotrimeric G proteins. Members of the Ras superfamily function as key nodes within signaling networks in a remarkable range of cellular processes, including proliferation, differentiation, nuclear and vesicular transport, nutrient sensing and many additional functions (Homma et al., 2021). They are major disease genes in humans and other animals and determinants in infection, immunity and pathogenesis. Moreover, extensive analysis of members of several subfamilies across divergent lineages, and as a consequence of ancient origins of these individual subfamilies, has indicated a level of conservation of function across eukaryotes that can be used to predict what an organism may have 'on board' in terms of compartments, signalling pathways and life cycle complexity, *albeit* in a general sense (Fig. 1) (Klinger et al., 2016).

The basis of GTPase functionality consists of a GTP hydrolysis cycle (Alberts et al., 2013). As most GTPases are inefficient hydrolytic enzymes the cycle of GTP to GDP conversion is frequently accelerated by a GTPase-activating protein (GAP) and exchange of GDP for GTP is facilitated by a guanine nucleotide exchange factor (GEF). This essentially two state or binary switch mechanism is exploited for directional functionality within vesicular transport, activation of gene expression pathways, flagellum synthesis and maintenance and nucleocytoplasmic transport, as well as other functions.

The largest subfamily are the Rab/Ran group, and which are primarily functioning in vesicle transport and nuclear transport. Rabs are important specifiers of organelle identity and the differentiation of

E-mail addresses: mfield@mac.com, mcfield@dundee.ac.uk.

https://doi.org/10.1016/j.protis.2024.126017
Received 13 October 2023; Received in revised form 21 December 2023; Accepted 26 January 2024
Available online 28 January 2024

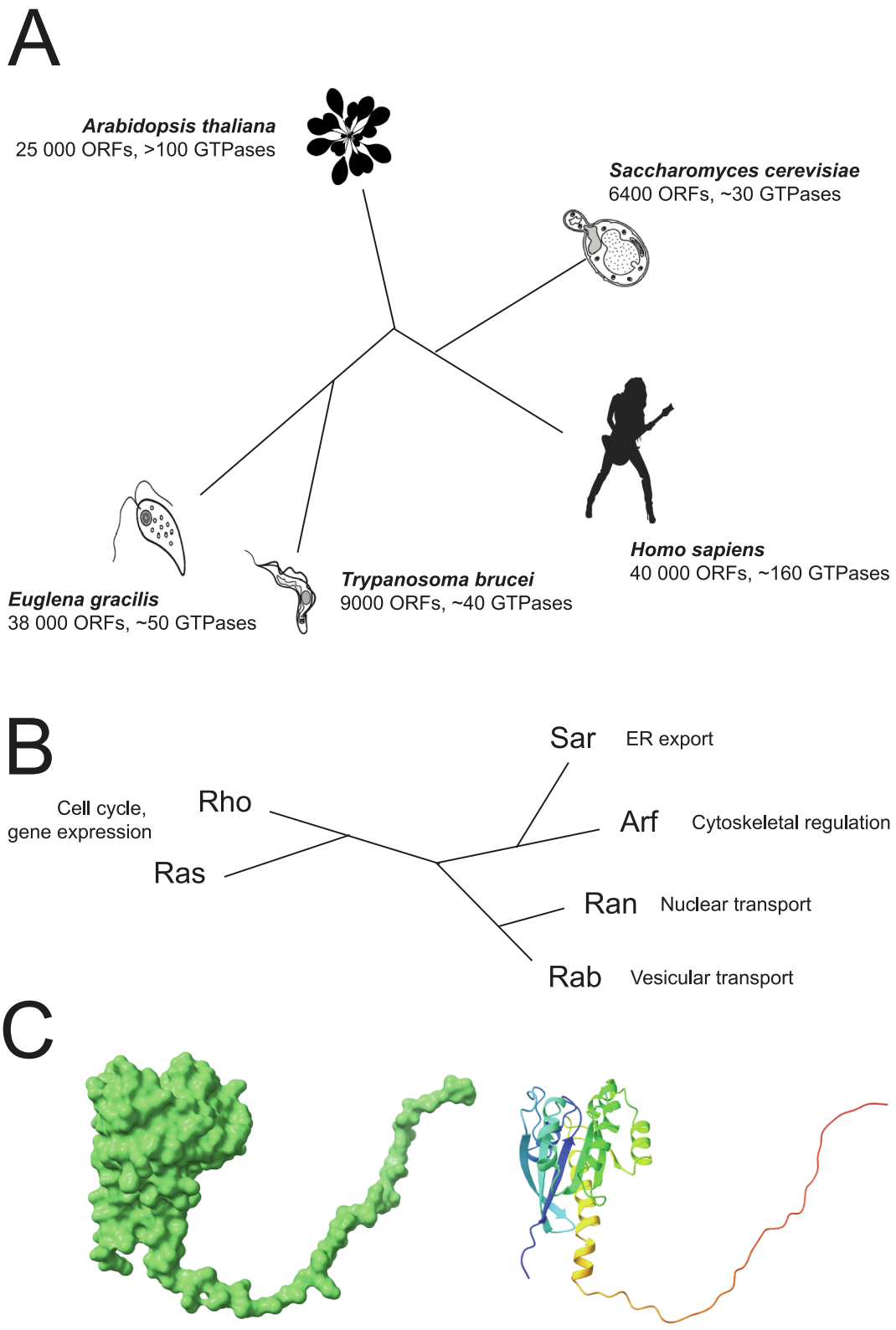


Fig. 1. Distribution and functional diversification of small GTPases. Panel A: Highly simplified phylogeny depicting the relationships between selected major lineages and exemplar species. For each example the approximate number of open reading frames (ORFs) and small GTPases encoded in the genome are given. Numbers are approximate due to strain variation in gene numbers, inaccuracies in gene annotation and other factors. Panel B: Broad classification of functions of small GTPase subfamilies. The dendrogram represents the phylogenetic relationships between GTPase subfamilies and indicates the major functional areas where each subfamily operates. Panel C: AlphaFold predicted structure for *T. brucei* Rab11. Left is a surface plot and right a peptide chain representation. Rab11 is used illustratively of the common overall fold and architecture of small GTPases. Note the presence of the extended C-terminal hyper-variable region which is involved in membrane anchoring via lipid modifications to the very C-terminus in many (but not all) small GTPase subfamilies.

transport pathways. Along with Rabs, the ARF subfamily also coordinate vesicular transport together with cytoskeletal modulation; the ARF/ARL family tends to be of modest size with four to six members in most organisms. Multiple small GTPases are involved in nucleolar processes, and in particular control of ribosome biogenesis. These latter include nucleostemins, while an important component of flagellum biogenesis is the small GTPase IFT27. GTPases primarily responsible for signal transduction include the Ras and Rho subfamilies, while Rheb and RagA/Gtr subfamily GTPases interact with the mTORC nutrient sensing machinery. While these designations are somewhat artificial as for the vast majority of GTPases a full understanding of functions remains elusive, they do serve as an indicator of where the major functional emphasis lies.

Complexity is a thrawn concept to define with any precision and especially in a biological context (Adami, 2002). Total gene number does not necessarily reflect overall sophistication, nor do the number of distinct life cycle stages correlate with gene number for example. However, several molecular systems do have at least some connection to sophistication in terms of the ability to respond to differential conditions or subcellular complexity, and include kinases, the ubiquitylation system and small GTPases. To consider the molecular basis of cellular complexity in *E. gracilis* I selected the small GTPases for several reasons. Their obvious central biological importance makes this an attractive family to understand, but they are also readily identifiable, containing a well recognised GTPase domain. Further, small GTPases have prokaryotic origins and hence predate eukaryogenesis (Verstraeten et al., 2011). Furthermore, the functions of most GTPases and their molecular interactions are understood at least in general terms, with strong evidence that functions are broadly conserved between orthologs, even when comparing highly divergent organisms (Klinger et al., 2016). There is also evidence for a loose scaling of Rab class GTPases with the numbers of ORFs in the genome (Gabernet-Castello et al., 2013), together with expansions of specific subfamilies, such as the Ras subfamily in metazoa. This latter is likely due to the role of these proteins in the control of developmental pathways and differentiation. I also chose to omit the larger GTPases, such as dynamin, to avoid an overly complex analysis and restricted searches to include essentially single domain GTPases. With this analysis I hope to at least address the question of how complex is *E. gracilis* when compared with the related kinetoplastids, and to begin to understand how the latter have become adapted to parasitism.

2. Methods

2.1. Databases

A predicted proteome for *E. gracilis* was used as the source of sequence data (https://fieldlab.org.s3-eu-west-2.amazonaws.com/reprints/reprints.html) (Ebenezer et al., 2019). This dataset represents ~ 98 % of open reading frames that have been described from various analyses (O'Neill et al., 2015; Yoshida et al., 2016; Ebenezer et al., 2019; Cordoba et al., 2021) and so was considered a sufficiently complete database for the purpose here.

2.2. Survey and data validation

Queries were a reference dataset that was previously used to identify Ras-like GTPases from *Trypanosoma brucei* and contains representatives of all major known Ras subfamilies and from a broad range of taxa (Field and O'Reilly, 2008). BLASTp (BLAST version 2.10.1) was performed against the *E. gracilis* predicted proteome using default settings. This raw dataset was parsed to exclude clear mishits (sequences shorter than 170 residues or longer that 800 residues). The resulting dataset were aligned using Clustal (https://www.ebi.ac.uk/Tools/msa/clustalo/) (Sievers and Higgins, 2014) and visualised using JalView (https://www.jalview.org/) (Procter et al., 2021) and FigTree (https://tree.bio.ed.ac.uk/software/figtree/ and https://github.com/rambaut/figtree/releases). Sequences were rejected that fell into clusters that, by BLAST (https://blast.ncbi.nlm.nih.gov/blast.cgi), did not return a Ras-related domain at CCDB, a Ras-related protein or lacked a reference set entry. Sequences were retained, even if they failed these criteria but contained a CAAX box (a signature of prenylation, a major feature of many Ras-like GTPases), for at least one additional round of validation. A total of eight rounds were performed and each of the *E. gracilis* sequences included in the final summation were validated and annotated manually by BLAST against the nr database and CCDB.

3. Results and discussion

3.1. The GTPase repertoire of Euglena gracilis

The present study is a survey, and not an in depth comparative genomics analysis. However, initial searches were relaxed, and deliberately so, to ensure that a full representation of the euglena GTPase repertoire was obtained. This strategy proved to indeed have significantly oversampled and the majority of initial sequences were rejected based on one or more criteria (see methods). The final list was manually curated, and is surprisingly small when compared with related taxa and considering the size of the genome and the number of ORFs encoded.

A total of 56 small GTPases were identified (Fig. 1, Table S1). Despite high confidence in the search strategy and completeness of the predicted proteome, caution is warranted; several sequences are clearly truncated, for example lacking an initiation methionine or the expected C-terminal prenylation motif (Supp data 1). Regardless, representatives of all of the expected Ras subfamilies were retrieved, including Ran, Ras, Rab, RagA and ARF proteins, *albeit* at a greater number to a previous analysis (Ebenezer et al., 2019). The distribution of these GTPases by subfamily suggests a considerable emphasis on intracellular compartmental complexity, with Rab proteins representing two thirds of the assignable Ras subfamily members (Fig. 2). I will suggest below that this reflects both a complex endosomal system coupled with a less elaborate GTPase repertoire elsewhere, *albeit* with some specific emphasis on metabolic control.

3.2. Signalling and related GTPases

For an organism with both an impressive number of protein coding genes and non-coding RNA genes (Cordoba et al., 2021), together with the ability to respond to considerable and wide ranging changes in the environment, the presence of single representatives of the Ras and Rho subfamily is surprising, and perhaps suggests that control of such responses is distinct from Ras/Rho GTPases. It is clear that cAMP-mediated signalling and kinases for example are important for phototaxis (Daiker et al., 2011) and a considerable adenylate cyclase family is encoded in the genome (Ebenezer et al., 2019); how these processes and other signalling mechanisms interface with Ras-mediated pathways is unknown in *E. gracilis*. Significantly, additional signalling type GTPases are also present, including Rheb, RagA and a double Ras-domain protein.

Rheb and RagA clades are, at least in part, involved in nutrient responses, with, in metazoa, both functioning with mTORC and controlling cellular growth and autophagy (Deng et al., 2019; Cui et al., 2023). RagA/Gtr GTPase dimers coordinate mTORC by facilitating the recruitment of Rheb to the mTORC complex (Kim and Kim, 2016; Gollwitzer et al., 2022; Cui et al., 2023). Two RagA and two Rheb GTPases were identified in the genome of *E. gracilis,* but none of the RagA group appear orthologous to *T. brucei,* while only one is possibly orthologous to a *Leishmania* RagA/Gtr GTPase as determined by BLAST of the *E. gracilis* sequences into the *T. brucei* or *Leishmania* genomes. This suggests an independent expansion of the RagA clade of GTPases in trypanosomes and euglenoids and is perhaps a reflection of differential demands for environmental responses in the highly flexible *E. gracilis,* which can survive in a multitude of conditions, compared to the more

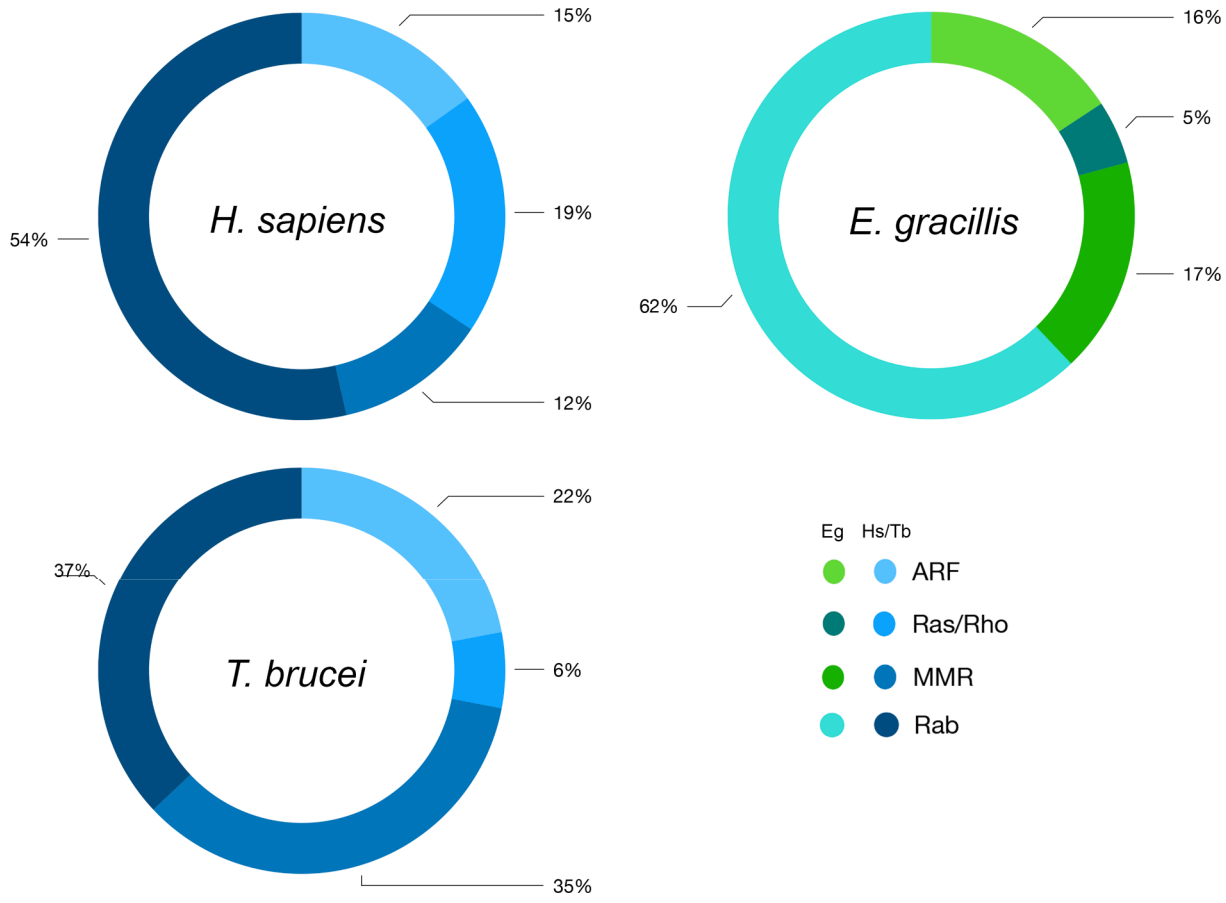


Fig. 2. Distribution of small GTPases between subfamilies in selected species. Punch through pie charts for *Homo sapiens*, *Trypanosoma brucei* and *Euglena gracilis*, divided into ARF, Rab, Ras/Rho and MMR subclasses. Data for *H. sapiens* and *T. brucei* are taken from Field (2005) and for *E. gracilis* from the present analysis. Comparisons of some specific subfamily paralog numbers are given in Table S2.

limited demands for trypanosomes that live within highly controlled host environments, *albeit* that recent studies have suggested greater metabolic flexibility than previously considered (e.g. Kovářová et al., 2018).

The double Ras-domain protein is unusual in possessing two GTPase domains in tandem followed by an EF-hand domain at the C-terminus; this latter domain implies a role in calcium signalling (Fig. 3). Alphafold was used to confirm the domain assignments for this unusual protein, but significantly proteins with similar architecture are present in multiple protist lineages as well as some prokaryotic taxa (Fig. 3) (Field, 2005); it is unclear if these forms have arisen with common ancestry or, due to the very patchy nature of their presence, via convergence. This protein is most similar architecturally to TbEAR from *T. brucei, albeit* that TbEAR lacks an EF-hand, nor is it the top sequence retrieved by BLAST. High throughput RNAi screens indicate essentiality for TbEAR in trypanosomes and a localisation to the mitochondrion (Alsford et al., 2012; Billington et al., 2023), but the unusual phylogenetic distribution of the *E. gracilis* protein and the absence of robust evidence for orthology suggests that TbEAR and the *E. gracilis* double Ras-domain protein are not closely related. *E. gracilis* also possesses an ortholog of RLJ, a GTPase with a J-domain C-terminal extension (Field, 2005), but the function of this protein in any taxon remains unclear.

3.3. Nuclear and nucleolar GTPases

Multiple GTPases operate to ensure the transcription and processing of rRNA, together with ribosomal assembly, mRNA maturation and export. Prominent amongst these is nucleostemin, which is localised to the nucleolus and well conserved across eukaryotes. The protein carries both N- and C-terminal extensions beyond the central GTPase domain (Tsai and Meng, 2009). Interestingly in invertebrates there is a single paralog and which is expanded to two genes in vertebrates, with division of labour between nucleostemin, which has a role in genome protection and the paralog GNL3-like, responsible for ribosomal synthesis (Lin et al., 2014). In common with most eukaryotes, *E. gracilis* has a single nucleostemin paralog which presumably supports both roles. Supporting this are three additional GTPases, two orthologs of Large Subunit GTPase 1 (Lsg1), as well as an ortholog of Nog1, and which are all involved in ribosomal assembly in other organisms (Klingauf-Nerurkar et al., 2020). Overall, this suggests a conserved role for GTPases in ribosomal biogenesis. There is also a single Ran ortholog, and which is highly conserved and presumably fulfils the role of coordinating nucelocytoplasmic import and export.

3.4. Intracellular transport GTPases

Small GTPases are critical to the definition, synthesis and movement of proteins and lipids between intracellular compartments. These include the flagellum, endomembrane system and interactions with the cytoskeleton, most are subtended by an evolutionary common origin, underpinning the basic mechanisms behind each of these subcellular systems (Field and Rout, 2022).

There is a single IFT27 ortholog, previously identified as associated with the flagellum by proteomics and hence confirming a conserved role in intraflagellar transport (Hammond et al., 2021). There are two Rab23 paralogs; this may also suggest a flagellum-based signalling pathway similar to mammals that is associated with the flagellum, although Rab23 also supports flagellum-independent functions (Hor et al., 2018).

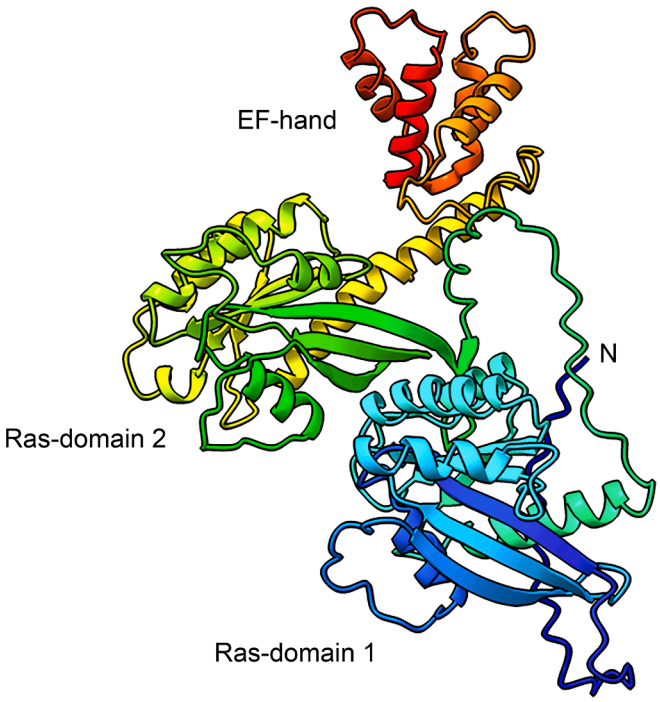


Fig. 3. Structure of double Ras-domain protein from *Euglena gracilis.* The best prediction is shown from an Alphafold analysis of the double Ras-domain protein encoded by EG transcript 6904. The individual domains are indicated in a rainbow coloured ribbon plot, with blue at the N-terminus (indicated) through red at the C-terminus. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

As might be expected, from a correlation between gene number and size of the Rab GTPase cohort (Gabernet-Castello et al., 2013), *E. gracilis* has an expanded set of Rab GTPases compared to trypanosomes, while the Arf family is of a more modest size, containing four members. The *E. gracilis* Rab repertoire was initially described several years ago, but the reevaluation here reveals additional complexity and pathways that were not captured previously. The cohort is larger and perhaps more significantly, more consistent with expectations of an organism with flexible biology (Ebenezer et al., 2019). Of the Rab subfamily over thirty members were identified and suggest that there are both more intracellular pathways, and hence compartments, than in trypanosomes where 16 Rabs have been described. There are also multiple paralogs for several conserved pathways, which suggests a level of sophistication and potential division of labour between essentially parallel pathways that is considerable.

In the exocytic pathway, there are three Rab1 paralogs and a single Rab2 member; this may suggest differential packaging of proteins exiting the ER. There are also three Rab6, one Rab8 and two Rab43 paralogs, all of which are associated with transiting the Golgi complex, while there is also a single Rab14, which is also involved in the later, post-Golgi stages of exocytosis. Overall, this repertoire suggests that there is limited complexity within the early exocytic pathway of *E. gracilis* but potentially considerable diversification at the level of the Golgi apparatus. Significantly, Sar1, the GTPase that mediates recruitment of the COPII coat to ER exit sites, was not found here, nor in previous work (Ebenezer et al., 2019), and which suggests the possibility of mechanistic divergence in this aspect of anterograde transport.

More significant are expansions within the endosomal system, *albeit* that this is a common observation in protists and elsewhere. There are two Rab5 paralogs in *E. gracilis*, equal to African trypanosomes and which clearly support distinct functions (Hall et al., 2004). However, more remarkable appears the presence of up to seven Rab11 paralogs and five distinct Rab7 genes, a level of complexity that surpasses most organisms with the exception of some higher plants (Table S1). Rab11 in animals, plants, fungi and trypanosomes is involved in recycling pathways, i.e. endocytosis and return to the surface of specific cargo, as well as in late stages of exocytosis of *de novo* synthesised cargo proteins. This pathway is frequently associated with receptor-mediated endocytosis, and is supported by a single paralog in African trypanosomes (Jeffries et al., 2001). The presence of multiple Rab11-mediated pathways suggests a highly sophisticated system for the sorting, recognition and segregation of material that is taken into the *E. gracilis* cell and its return to the surface, and again is a possible reflection of the complexity of surface protein families, and provides a potential mechanism for the control of membrane trafficking under differential conditions. Moreover, multiple Rab7 paralogs suggest a complex set of routes to the terminal lysosome. It is significant that a Rab4 isoform was not identified, which is distinct from trypanosomes, and which suggests that a rapid recycling pathway is absent from *E. gracilis.*

There are also Rab proteins indicating that recycling pathways between endosomes and the Golgi complex are in place. Specifically, there are single paralogs for Rab22 and Rab21 is absent. Rab21 is involved in retromer recycling and also late endocytosis to lysosome trafficking and is present in trypanosomes (Ali et al., 2014; Pei et al., 2023), while Rab22 has a distinct role in endosome/Golgi bidirectional transport. This may suggest a slight simpler form of the interface between the Golgi/endosome system, and the absence of Rab21 is likely a secondary loss.

Finally, the possession of a Rab14 paralog is consistent with the presence of a contractile vacuole in *E. gracilis*, while two Rab32 paralogs may indicate either trafficking to lysosome-like organelles (Ohbayashi et al., 2017) or an additional component of the mTORC system of nutrient sensing (Drizyte-Miller et al., 2020). Given that several other GTPases, specifically Rheb and RagA, also interact with mTORC, this level of GTPase complexity suggests that metabolic regulation may have considerable flexibility.

4. Conclusions

This brief survey of the *E. gracilis* small GTPase repertoire is intended as a spur for further investigation, and also to provide some insights into the intracellular compartment and signal transduction complexity (summarised in Fig. 4). A number of general conclusions can be drawn.

Firstly, the repertoire is comparatively small for a genome of the size of *E. gracilis*, being only about 30 % larger that *T. brucei*, despite a near five-fold increase in the number of protein-coding genes. This is reflected in a small Ras and Rho cohort and may well indicate that environmental sensing and control of differential gene expression lies within alternate mechanisms. Secondly, in contrast to the small Ras/Rho cohort there is considerable complexity predicted for the mTORC amino acid sensor, which may well consist of conserved elements such as Rab32, Rheb and RagA GTPases. It will be of considerable interest to address how these factors interact and respond to altering conditions and nutrient availability. Third and finally, there is a high level of pathway complexity within the endocytic system, with a multitude of potential recycling and exocytosis pathways, together with several late transport pathways to the lysosome. Overall, the survey does suggest an emphasis on flexibility in interacting with the environment and adapting to distinct conditions reflected by the GTPase repertoire.

CRediT authorship contribution statement

Mark C. Field: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Validation, Visualization, Writing – original draft, Writing – review & editing.

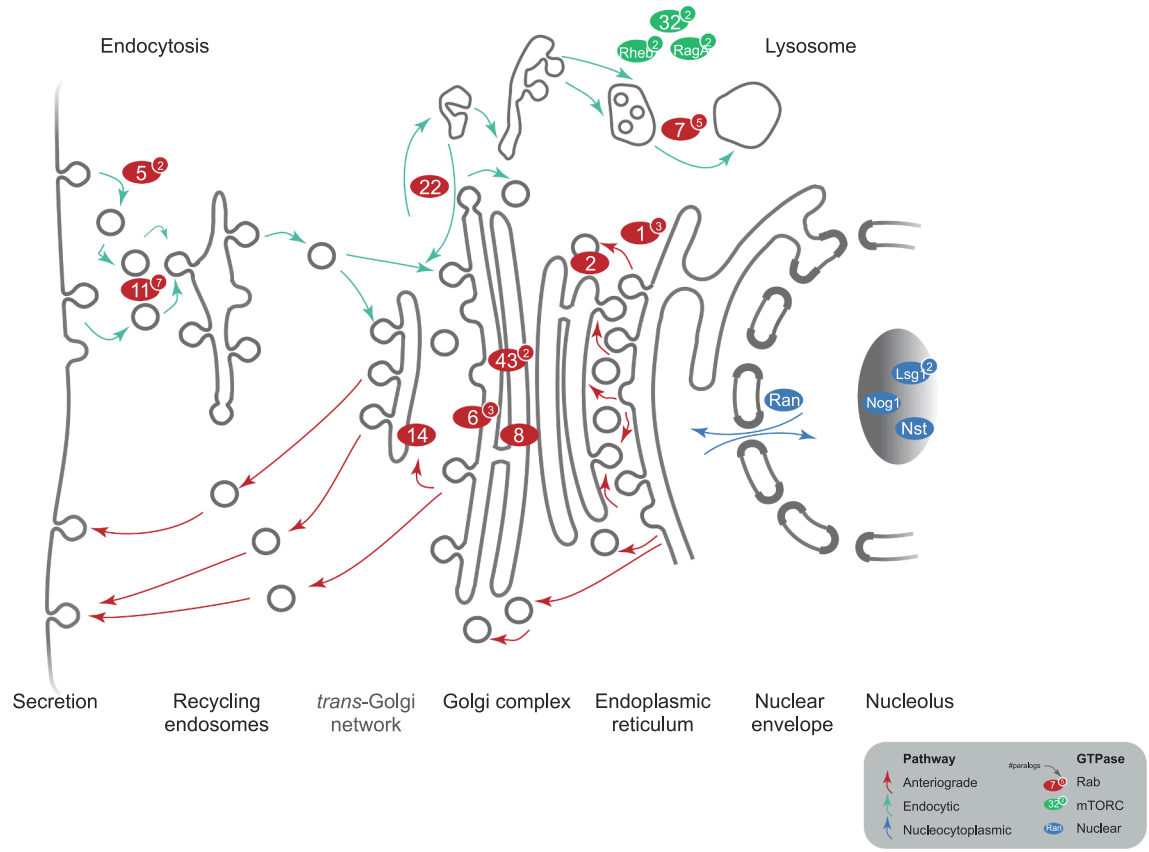


Fig. 4. Predicated locations of small GTPases and subcellular compartments in *Euglena gracilis*. The general figure is adapted from Field and Rout (2022). Rab proteins are indicated with red lozenges, and the number of paralogs by small circles with numbers. Other GTPases likely associated with mTORC are indicated with green lozenges and those with nuclear functions in blue, and their annotation indicated. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

I am hugely grateful to Erin Butterfield (Dundee) for extracting candidate GTPase protein sequences from the *E. gracilis* predicted proteome. Work in my laboratory is supported by the Wellcome Trust (204697/Z/16/Z to MCF).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.protis.2024.126017.

References

Adami, C., 2002. What is complexity? Bioessays 24 (12), 1085–1094. https://doi.org/10.1002/bies.10192. PMID: 12447974.

Alberts, B., Bray, D., Hopkin, K., Johnson, A.D., Lewis, J., Raff, M., Roberts, K., Walter, P., 2013. Essential Cell Biology, 4th ed. Garland Publishing.

Ali, M., Leung, K.F., Field, M.C., 2014. The ancient small GTPase Rab21 functions in intermediate endocytic steps in trypanosomes. Eukaryot Cell 13 (2), 304–319. https://doi.org/10.1128/EC.00269-13. Epub 2013 Dec 27. PMID: 24376004; PMCID: PMC3910970.

Alsford, S., Eckert, S., Baker, N., Glover, L., Sanchez-Flores, A., Leung, K.F., Turner, D.J., Field, M.C., Berriman, M., Horn, D., 2012. High-throughput decoding of antitrypanosomal drug efficacy and resistance. Nature 482 (7384), 232–236. https://doi.org/10.1038/nature10771. PMID: 22278056; PMCID: PMC3303116.

Billington, K., Halliday, C., Madden, R., Dyer, P., Barker, A.R., Moreira-Leite, F.F., Carrington, M., Vaughan, S., Hertz-Fowler, C., Dean, S., Sunter, J.D., Wheeler, R.J., Gull, K., 2023. Genome-wide subcellular protein map for the flagellate parasite *Trypanosoma brucei*. Nat. Microbiol. 8 (3), 533–547. https://doi.org/10.1038/s41564-022-01295-6. Epub 2023 Feb 20. PMID: 36804636; PMCID: PMC9981465.

Cordoba, J., Perez, E., Van Vlierberghe, M., Bertrand, A.R., Lupo, V., Cardol, P., Baurain, D., 2021. De Novo transcriptome meta-assembly of the mixotrophic freshwater microalga *Euglena gracilis*. Genes (Basel) 12 (6), 842. https://doi.org/10.3390/genes12060842. PMID: 34072576; PMCID: PMC8227486.

Cui, Z., Napolitano, G., de Araujo, M.E.G., Esposito, A., Monfregola, J., Huber, L.A., Ballabio, A., Hurley, J.H., 2023. Structure of the lysosomal mTORC1-TFEB-Rag-Ragulator megacomplex. Nature 614 (7948), 572–579. https://doi.org/10.1038/s41586-022-05652-7. Epub 2023 Jan 25. PMID: 36697823; PMCID: PMC9931586.

Daiker, V., Häder, D.P., Richter, P.R., Lebert, M., 2011. The involvement of a protein kinase in phototaxis and gravitaxis of *Euglena gracilis*. Planta 233 (5), 1055–1062. https://doi.org/10.1007/s00425-011-1364-5. Epub 2011 Feb 1 PMID: 21286747.

Deng, L., Chen, L., Zhao, L., Xu, Y., Peng, X., Wang, X., Ding, L., Jin, J., Teng, H., Wang, Y., Pan, W., Yu, F., Liao, L., Li, L., Ge, X., Wang, P., 2019. Ubiquitination of Rheb governs growth factor-induced mTORC1 activation. Cell Res. 29 (2), 136–150. https://doi.org/10.1038/s41422-018-0120-9. Epub 2018 Dec 4. PMID: 30514904; PMCID: PMC6355928.

Drizyte-Miller, K., Chen, J., Cao, H., Schott, M.B., McNiven, M.A., 2020. The small GTPase Rab32 resides on lysosomes to regulate mTORC1 signaling. J. Cell Sci. 133 (11), jcs236661 https://doi.org/10.1242/jcs.236661. PMID: 32295849; PMCID: PMC7295596.

Ebenezer, T.E., Zoltner, M., Burrell, A., Nenarokova, A., Novák Vanclová, A.M.G., Prasad, B., Soukal, P., Santana-Molina, C., O'Neill, E., Nankissoor, N.N., Vadakedath, N., Daiker, V., Obado, S., Silva-Pereira, S., Jackson, A.P., Devos, D.P., Lukeš, J., Lebert, M., Vaughan, S., Hampl, V., Carrington, M., Ginger, M.L., Dacks, J. B., Kelly, S., Field, M.C., 2019. Transcriptome, proteome and draft genome of *Euglena gracilis*. BMC Biol. 17 (1), 11. https://doi.org/10.1186/s12915-019-0626-8. PMID: 30732613; PMCID: PMC6366073.

Ebenezer, T.E., Low, R.S., O'Neill, E.C., Huang, I., DeSimone, A., Farrow, S.C., Field, R. A., Ginger, M.L., Guerrero, S.A., Hammond, M., Hampl, V., Horst, G., Ishikawa, T., Karnkowska, A., Linton, E.W., Myler, P., Nakazawa, M., Cardol, P., Sánchez-

Thomas, R., Saville, B.J., Shah, M.R., Simpson, A.G.B., Sur, A., Suzuki, K., Tyler, K. M., Zimba, P.V., Hall, N., Field, M.C., 2022. Euglena International Network (EIN): driving euglenoid biotechnology for the benefit of a challenged world. Biol. Open 11 (11), bio059561. https://doi.org/10.1242/bio.059561. Epub 2022 Nov 22. PMID: 36412269; PMCID: PMC9836076.

Field, M.C., 2005. Signalling the genome: the Ras-like small GTPase family of trypanosomatids. Trends Parasitol. 21 (10), 447–450. https://doi.org/10.1016/j.pt.2005.08.008. PMID: 16112905.

Field, M.C., O'Reilly, A.J., 2008. How complex is GTPase signaling in trypanosomes? Trends Parasitol. 24 (6), 253–257. https://doi.org/10.1016/j.pt.2008.03.005. Epub 2008 May 6 PMID: 18467174.

Field, M.C., Rout, M.P., 2022. Coatomer in the universe of cellular complexity. Mol. Biol. Cell. 33 (14), pe8 https://doi.org/10.1091/mbc.E19-01-0012. PMID: 36399624; PMCID: PMC9727805.

Gabernet-Castello, C., O'Reilly, A.J., Dacks, J.B., Field, M.C., 2013. Evolution of Tre-2/Bub2/Cdc16 (TBC) Rab GTPase-activating proteins. Mol. Biol. Cell 24 (10), 1574–1583. https://doi.org/10.1091/mbc.E12-07-0557. Epub 2013 Mar 13. PMID: 23485563; PMCID: PMC3655817.

Gollwitzer, P., Grützmacher, N., Wilhelm, S., Kümmel, D., Demetriades, C., 2022. A Rag GTPase dimer code defines the regulation of mTORC1 by amino acids. Nat. Cell Biol. 24 (9), 1394–1406. https://doi.org/10.1038/s41556-022-00976-y. Epub 2022 Sep 12. Erratum in: Nat Cell Biol. 2023 Feb;25(2):366. PMID: 36097072; PMCID: PMC9481461.

Gumińska, N., Płecha, M., Zakryś, B., Milanowski, R., 2018. Order of removal of conventional and nonconventional introns from nuclear transcripts of *Euglena gracilis*. PLoS Genet. 14 (10), e1007761 https://doi.org/10.1371/journal.pgen.1007761. PMID: 30365503; PMCID: PMC6221363.

Hall, B., Allen, C.L., Goulding, D., Field, M.C., 2004. Both of the Rab5 subfamily small GTPases of *Trypanosoma brucei* are essential and required for endocytosis. Mol. Biochem. Parasitol. 138 (1), 67–77. https://doi.org/10.1016/j.molbiopara.2004.07.007. PMID: 15500917.

Hammond, M., Zoltner, M., Garrigan, J., Butterfield, E., Varga, V., Lukeš, J., Field, M.C., 2021. The distinctive flagellar proteome of *Euglena gracilis* illuminates the complexities of protistan flagella adaptation. New Phytol. 232 (3), 1323–1336. https://doi.org/10.1111/nph.17638. Epub 2021 Aug 16 PMID: 34292600.

Homma, Y., Hiragi, S., Fukuda, M., 2021. Rab family of small GTPases: an updated view on their regulation and functions. FEBS J. 288 (1), 36–55. https://doi.org/10.1111/febs.15453. Epub 2020 Jul 1. PMID: 32542850; PMCID: PMC7818423.

Hor, C.H.H., Tang, B.L., Goh, E.L.K., 2018. Rab23 and developmental disorders. Rev. Neurosci. 29 (8), 849–860. https://doi.org/10.1515/revneuro-2017-0110. PMID: 29727300.

Jeffries, T.R., Morgan, G.W., Field, M.C., 2001. A developmentally regulated rab11 homologue in *Trypanosoma brucei* is involved in recycling processes. J. Cell Sci. 114 (Pt 14), 2617–2626. https://doi.org/10.1242/jcs.114.14.2617. PMID: 11683389.

Kim, J., Kim, E., 2016. Rag GTPase in amino acid signaling. Amino Acids 48 (4), 915–928. https://doi.org/10.1007/s00726-016-2171-x. Epub 2016 Jan 18 PMID: 26781224.

Klingauf-Nerurkar, P., Gillet, L.C., Portugal-Calisto, D., Oborská-Oplová, M., Jäger, M., Schubert, O.T., Pisano, A., Peña, C., Rao, S., Altvater, M., Chang, Y., Aebersold, R., Panse, V.G., 2020. The GTPase Nog1 co-ordinates the assembly, maturation and quality control of distant ribosomal functional centers. Elife 7 (9), e52474. https://doi.org/10.7554/eLife.52474. PMID: 31909713; PMCID: PMC6968927.

Klinger, C.M., Ramirez-Macias, I., Herman, E.K., Turkewitz, A.P., Field, M.C., Dacks, J.B., 2016. Resolving the homology-function relationship through comparative genomics of membrane-trafficking machinery and parasite cell biology. Mol. Biochem. Parasitol. 209 (1–2), 88–103. https://doi.org/10.1016/j.molbiopara.2016.07.003. Epub 2016 Jul 19. PMID: 27444378; PMCID: PMC5140719.

Kovářová, J., Nagar, R., Faria, J., Ferguson, M.A.J., Barrett, M.P., Horn, D., 2018. Gluconeogenesis using glycerol as a substrate in bloodstream-form *Trypanosoma brucei*. PLoS Pathog. 14 (12), e1007475 https://doi.org/10.1371/journal.ppat.1007475. PMID: 30589893; PMCID: PMC6307712.

Lin, T., Meng, L., Lin, T.C., Wu, L.J., Pederson, T., Tsai, R.Y., 2014. Nucleostemin and GNL3L exercise distinct functions in genome protection and ribosome synthesis, respectively. J. Cell Sci. 127 (Pt 10), 2302–2312. https://doi.org/10.1242/jcs.143842. Epub 2014 Mar 7. PMID: 24610951; PMCID: PMC6519424.

McWatters, D.C., Russell, A.G., 2017. Euglena transcript processing. Adv. Exp. Med. Biol. 979, 141–158. https://doi.org/10.1007/978-3-319-54910-1_8. PMID: 28429321.

Ohbayashi, N., Fukuda, M., Kanaho, Y., 2017. Rab32 subfamily small GTPases: pleiotropic Rabs in endosomal trafficking. J. Biochem. 162 (2), 65–71. https://doi.org/10.1093/jb/mvx027. PMID: 28430987.

O'Neill, E.C., Trick, M., Hill, L., Rejzek, M., Dusi, R.G., Hamilton, C.J., Zimba, P.V., Henrissat, B., Field, R.A., 2015. The transcriptome of *Euglena gracilis* reveals unexpected metabolic capabilities for carbohydrate and natural product biochemistry. Mol. Biosyst. 11 (10), 2808–2820. https://doi.org/10.1039/c5mb00319a. PMID: 26289754.

Pei, Y., Lv, S., Shi, Y., Jia, J., Ma, M., Han, H., Zhang, R., Tan, J., Zhang, X., 2023. RAB21 controls autophagy and cellular energy homeostasis by regulating retromer-mediated recycling of SLC2A1/GLUT1. Autophagy 19 (4), 1070–1086. https://doi.org/10.1080/15548627.2022.2114271. Epub 2022 Aug 21. PMID: 35993307; PMCID: PMC10012929.

Procter, J.B., Carstairs, G.M., Soares, B., Mourão, K., Ofoegbu, T.C., Barton, D., Lui, L., Menard, A., Sherstnev, N., Roldan-Martinez, D., Duce, S., Martin, D.M.A., Barton, G. J., 2021. Alignment of biological sequences with Jalview. Methods Mol. Biol. 2231, 203–224. https://doi.org/10.1007/978-1-0716-1036-7_13. Erratum. In: Methods Mol Biol. 2021;2231:C1. PMID: 33289895; PMCID: PMC7116599.

Sievers, F., Higgins, D.G., 2014. Clustal omega. Curr. Protoc. Bioinf. 48, 3.13.1–3.13.16. https://doi.org/10.1002/0471250953.bi0313s48. PMID: 25501942.

Tessier, L.H., Keller, M., Chan, R.L., Fournier, R., Weil, J.H., Imbault, P., 1991 Sep. Short leader sequences may be transferred from small RNAs to pre-mature mRNAs by trans-splicing in Euglena. EMBO J. 10 (9), 2621–2625. https://doi.org/10.1002/j.1460-2075.1991.tb07804.x. PMID: 1868836; PMCID: PMC452961.

Tsai, R.Y., Meng, L., 2009. Nucleostemin: a latecomer with new tricks. Int. J. Biochem. Cell Biol. 41 (11), 2122–2124. https://doi.org/10.1016/j.biocel.2009.05.020. Epub 2009 Jun 6. PMID: 19501670; PMCID: PMC2753700.

Verstraeten, N., Fauvart, M., Versées, W., Michiels, J., 2011. The universally conserved prokaryotic GTPases. Microbiol. Mol. Biol. Rev. 75 (3), 507–542. https://doi.org/10.1128/MMBR.00009-11 second and third pages of table of contents. PMID: 21885683; PMCID: PMC3165542.

Yoshida, Y., Tomiyama, T., Maruta, T., Tomita, M., Ishikawa, T., Arakawa, K., 2016. De novo assembly and comparative transcriptome analysis of *Euglena gracilis* in response to anaerobic conditions. BMC Genomics 3 (17), 182. https://doi.org/10.1186/s12864-016-2540-6. PMID: 26939900; PMCID: PMC4778363.