CellPress

Trends in Parasitology

Opinion

Reductionist Pathways for Parasitism in Euglenozoans? Expanded Datasets Provide New Insights

Anzhelika Butenko,[1,2,*] Michael Hammond,[1] Mark C. Field,[1,3] Michael L. Ginger,[4] Vyacheslav Yurchenko,[2,5] and Julius Lukeš[1,6,*]

The unicellular trypanosomatids belong to the phylum Euglenozoa and all known species are obligate parasites. Distinct lineages infect plants, invertebrates, and vertebrates, including humans. Genome data for marine diplonemids, together with freshwater euglenids and free-living kinetoplastids, the closest known nonparasitic relatives to trypanosomatids, recently became available. Robust phylogenetic reconstructions across Euglenozoa are now possible and place the results of parasite-focused studies into an evolutionary context. Here we discuss recent advances in identifying the factors shaping the evolution of Euglenozoa, focusing on ancestral features generally considered parasite-specific. Remarkably, most of these predate the transition(s) to parasitism, suggesting that the presence of certain preconditions makes a significant lifestyle change more likely.

Highlights

Genome streamlining and the loss of certain metabolic pathways predate switches to parasitism in the evolution of Euglenozoa and are the result of a multistep process.

Numerous features previously considered trypanosomatid-specific are present also in the free-living euglenozoans. They include nutrient-triggered attachment to surfaces, polycistronic transcription, *trans*-splicing, trypanothione, the loss of glutathione reductase, genome compaction via almost complete loss of *cis*-spliced introns, presence and even diversification of surface proteins, subtilisins, and carboxypeptidases.

Rapid sequence evolution in Euglenozoa is not linked to parasitism, as it also occurs in free-living relatives of kinetoplastids.

Origins of Parasitism

Parasitism has emerged on many occasions during eukaryotic evolution. We have clear appreciations of the important cellular features of parasitic organisms due, in part, to the considerable attention they receive as disease agents. However, detailed information for both parasitic and free-living sister lineages is essential to distinguish genuine adaptations to parasitism from other specializations that are common across a lineage and represents a significant knowledge gap for many parasitic groups.

What drives the switch from a free-living to a parasitic lifestyle? A classic view holds that easy access to host nutrients per se is sufficient for metabolic streamlining, thereby locking previously free-living or symbiotic taxa into parasitism. But, are specific preconditions required [1], which are subsequently augmented and thereby 'seal the deal'? The era of inexpensive genome and transcriptome sequencing provides a platform to comprehensively address many long-standing and critical questions concerning the nature of drivers behind parasitism.

The phylum **Euglenozoa** (see Glossary) provides a suitable case study for comparative analysis. Euglenozoa encompasses a photosynthetic clade, multiple heterotrophic clades, independent examples of parasitism and mutualism, as well as a well studied exclusively parasitic clade of trypanosomatids (Figure 1A). Trypanosomatids are, and ancestrally were, parasites of invertebrates, but dixenous parasitism has evolved multiple times, further obscuring what drives the transition to obligate parasitism. Belonging to the class Kinetoplastea, so named because of an extraordinarily large mass of mitochondrial DNA hereafter called **kinetoplast DNA** (**kDNA**), the osmotrophic trypanosomatids have likely evolved from a free-living ancestor with a phagotrophic mode of nutrition as suggested by the presence of a **cytostome–cytopharynx complex** in free-living kinetoplastids and its

[1]Biology Centre, Institute of Parasitology, Czech Academy of Sciences, České Budějovice (Budweis), Czech Republic
[2]Faculty of Science, University of Ostrava, Ostrava, Czech Republic
[3]School of Life Sciences, University of Dundee, Dundee, UK
[4]School of Applied Sciences, University of Huddersfield, Huddersfield, UK
[5]Martsinovsky Institute of Medical Parasitology, Sechenov University, Moscow, Russia
[6]Faculty of Sciences, University of South Bohemia, České Budějovice (Budweis), Czech Republic

*Correspondence:
anzhelika.butenko@paru.cas.cz (A. Butenko) and
jula@paru.cas.cz (J. Lukeš).

 https://doi.org/10.1016/j.pt.2020.10.001

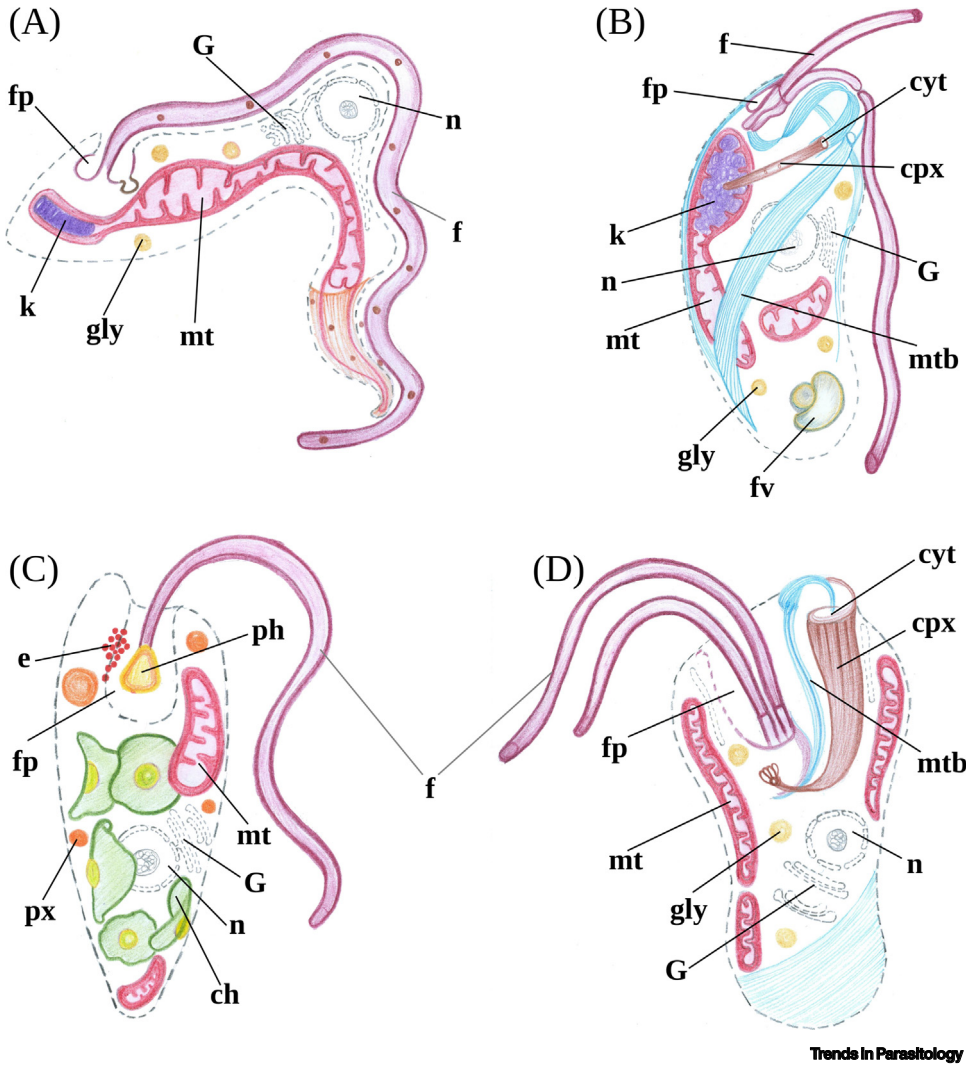

Figure 1. Schematic Morphologies of a Trypanosomatid, *Trypanosoma brucei* (A), a Bodonid, *Bodo saltans* (B), a Euglenid, *Euglena gracilis* (C), and a Diplonemid, *Diplonema japonicum* (D). Shared features are shown using the same color and/or broken lines; examples of features specific for individual clades are: the disk-shaped kinetoplast DNA for trypanosomatids; the bulky pro-kinetoplast DNA of bodonids; the green plastids of euglenids; and the prominent cytostome for diplonemids. Abbreviations: ch, chloroplast; cpx, cytopharynx; cyt, cytostome; e, eyespot; f, flagellum; fp, flagellar pocket; fv, food vacuole; G, Golgi apparatus; gly, glycosome; k, kinetoplast; mt, mitochondrion; mtb, microtubules; n, nucleus; ph, photoreceptor; px, peroxisome.

retention in some trypanosomatids (Figure 1B) [2,3]. Within kinetoplastids outside the Trypanosomatidae family (Parabodonida and Neobodonida, Prokinetoplastida), abundant in both freshwater and marine environments, multiple independent occurrences of parasitism have been documented [4]. Two other taxonomically diverse, broadly distributed lineages within Euglenozoa are the predominantly freshwater euglenids (Figure 1C) and almost exclusively marine diplonemids (Figure 1D), which are both generally considered to be free-living [4–8]. Symbiontids, euglenozoan protists covered with epibiotic bacteria, are understudied, and whole-genome/transcriptome sequencing data are lacking for this group. Therefore, we do not consider them here [9].

Glossary

Cytostome–cytopharynx complex: an oral apparatus of the cell, consisting of a deep invagination of the plasma membrane ('cytopharynx') and its opening on the cell surface ('cytostome').

Euglenozoa: a group of unicellular eukaryotes (protists) unifying Kinetoplastea, Diplonemea, Euglenida, and Symbiontida. Best known kinetoplastids are *Leishmania* and *Trypanosoma*, serious parasites of humans. Diplonemids are diverse and abundant marine planktonic flagellates. Euglenids incorporate both nonphotosynthetic (*Rhabdomonas*) and photosynthetic species with the secondary plastids of green algal origin (*Euglena*). Symbiontids are rare protists covered with epibiotic bacteria.

Extrusomes: membrane-bound extrusive organelles, which are discharged from cells under stimuli, serving a variety of functions.

Kinetochore: a complex machinery driving chromosome segregation during cell division by linking chromosomes/chromatids to spindle microtubules built of β-tubulin.

Kinetoplast DNA (kDNA): a highly complex mitochondrial DNA of Kinetoplastea, composed of relaxed or supercoiled DNA circles, either catenated into a single network, or noncatenated, or a combination of both arrangements.

Lacunae: hollow cytoplasmic spaces typically associated with the cytopharynx in Euglenozoa.

Nuclear pore complex (NPC): a large multiprotein complex spanning the nuclear envelope and governing transport of macromolecules between the nucleus and the cytoplasm.

Opisthokonts (Opisthokonta): a eukaryotic supergroup incorporating animals, fungi, and their unicellular relatives.

RNA editing: post-transcriptional changes in an RNA sequence, leading to differences from the DNA template.

Comparative analysis of trypanosomatid nuclear genomes against closely related free-living *Bodo saltans* indicated that a switch in the mode of nutrient gain, from phagotrophy to osmotrophy, rather than the loss of specific metabolic pathways, marked the transition to parasitism [10,11]. In some instances, cell surface innovations that influence the interactions of trypanosomatids with their hosts were inherited from a common ancestor with *B. saltans*. Genome analyses of monoxenous insect trypanosomatids, for example, *Leptomonas* and *Paratrypanosoma,* have reinforced the view that, with a few exceptions, wide-scale losses of metabolic pathways did not accompany the wider radiation of parasitism [12,13]. However, only recently has a thorough comparison between parasites and their free-living relatives become possible due to the availability of genomes and transcriptomes for representatives from each of these species-rich and ecologically significant groups [14]. This allows us to identify on an evolutionary scale which features can be considered early preconditions for parasitism and carried over during the transition, versus true parasitism-associated innovations.

Finally, comparisons between parasitic euglenozoans and free-living relatives can identify constraints associated with ubiquitous and highly conserved eukaryotic machineries. Indeed, for many examples, only one or two core subunits are identifiable, suggesting significant diversification of even apparent core functions. In total, we have analyzed the distribution of almost 30 features/traits across the Euglenozoa (Figure 2; see also Box S1 in the supplemental information online). The majority of these features fall into the following categories: (i) traits which were or still are considered lineage-specific within Euglenozoa, including but not restricted to presumably parasite-specific ones (e.g., kDNA network, cell-surface proteins, glycosomes, fatty acid biosynthesis using a set of elongases); (ii) traits with curious evolutionary history in Euglenozoa and/or other eukaryotes (e.g., heme biosynthesis enzymes, pentafunctional AROM protein, family A DNA polymerases); and (iii) protein complexes known to be divergent in trypanosomatids compared with **opisthokonts** (e.g., **kinetochore**, histones). The evolution of most metabolic and other features discussed herein was studied in parasitic trypanosomatids and their closest outgroup species with a sequenced genome, a eubodonid *B. saltans* [1,10,11,15]. The studies of other traits occasionally included diplonemids and/or euglenids [e.g., base J [16], **RNA editing** [17], heme biosynthesis [18], reactive oxygen species (ROS)-detoxifying systems [19], the presence of glycosomes [20], nonconventional introns [21], surface proteins [22]]. In this review, we summarize the insights brought by the recent analyses of diplonemid and euglenid genomes and transcriptomes [14,23] and expand the analyses of several traits (e.g., subtilisins, peptidases, histones, 1,3-β-glucan biosynthesis, amastins and their domain composition) using these newly available data. The emerging pattern of gains, losses, and reacquisitions [e.g., via horizontal gene transfer (HGT)] is complex, and one can anticipate that future dissection of more traits will add further complexity to this picture.

The Euglenozoan Common Ancestor: Lessons from Free-Living Relatives of Parasitic Trypanosomatids

Nuclear Genome Organization

Striking features of trypanosomatid nuclear genomes and gene expression include the organization of protein-coding genes into long polycistronic transcription units (PTUs), an apparent lack of transcription initiation regulation, deployment of 'ribosomal' RNA polymerase I to transcribe specialized protein-coding genes in African trypanosomes, as well as exploitation of subtelomeres for the isolation of key virulence genes [24,25]. From recent data, we find that the euglenozoan common ancestor (EUCA) likely possessed a large repetitive genome containing both conventional and nonconventional *cis*-spliced introns [14]. Genome streamlining and loss of most *cis*-spliced introns thus happened in the kinetoplastid common ancestor, although more genomes of free-living kinetoplastids need to be analyzed to confirm this (Figure 2). Protein-coding genes in

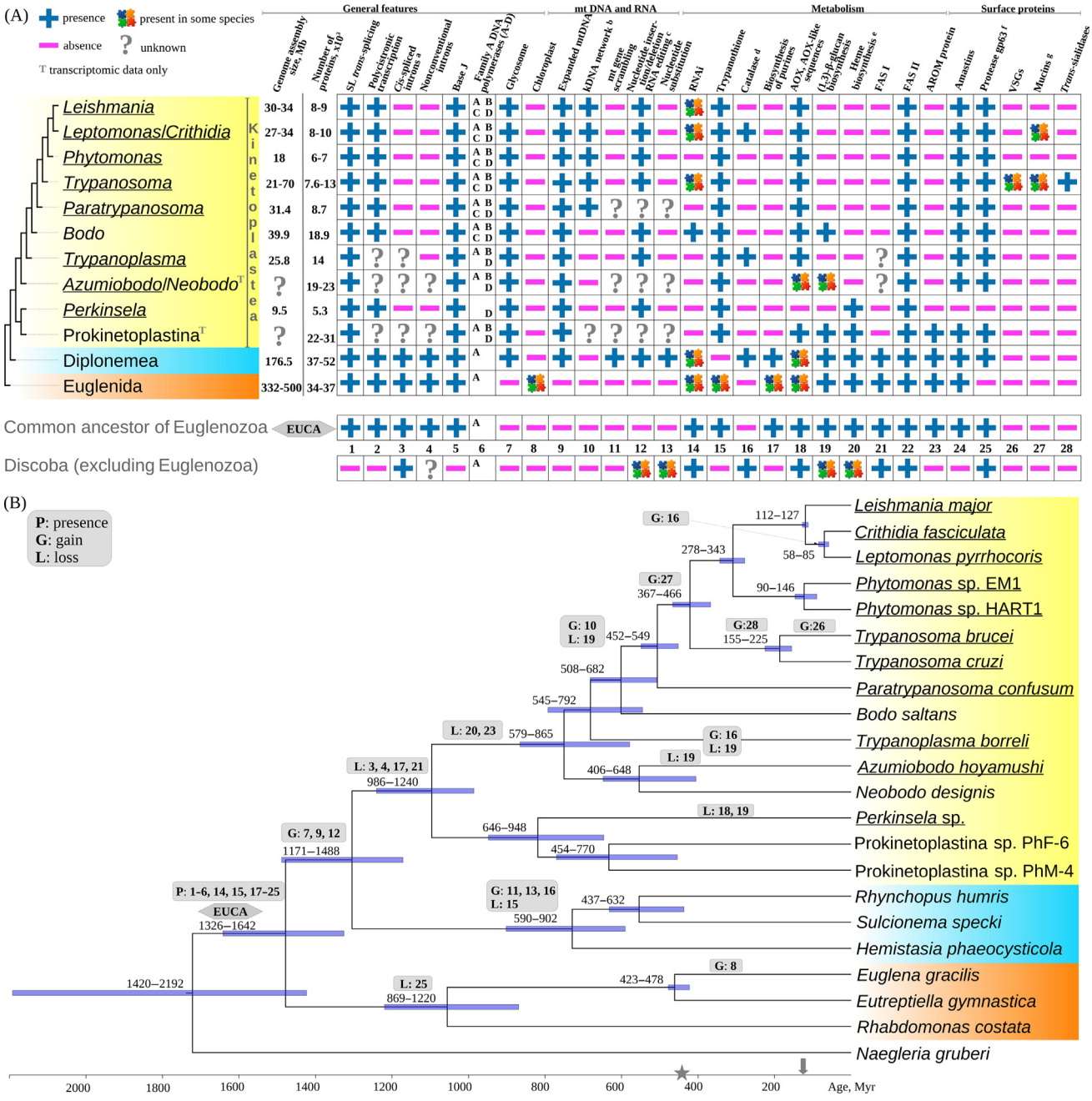

Figure 2. The Distribution of Various Features of Kinetoplastids, Diplonemids, and Euglenids and Their Putative State in the Common Ancestor of Euglenozoa. (A) Presence and absence are marked with '+' and '–' signs, respectively. Jigsaw puzzle icon indicates presence in only a subset of species within a group. Question marks indicate absence of data. Abbreviations: AOX, alternative oxidase; EUCA, euglenozoan common ancestor; FAS, fatty acid synthase; kDNA, kinetoplast DNA; mt, mitochondrial; RNAi, RNA interference; SL, splice leader RNA; VSGs, variant surface glycoproteins. [a]The presence of *cis*-spliced introns was confirmed for several trypanosomatid genomes and is confined to very few pre-mRNAs, including poly(A) polymerase and RNA helicase. [b]Pro-kDNA of *Bodo saltans* is composed of individual minicircles, with only a few very small catenanes. [c]The analysis of RNA editing in diplonemids revealed insertions, but not deletions. [d]We could not identify catalase among *Trypanoplasma borreli* transcripts, but its activity was demonstrated before by Opperdoes and colleagues [109]. [e]The last three enzymes of the pathway are encoded in the genomes of Leishmaniinae, while *Phytomonas* spp. have only ferrochelatase; *Perkinsela* possesses a near-complete set of heme path-

(Figure legend continued at the bottom of the next page.)

EUCA were also organized into PTUs of seemingly functionally unrelated genes, and the respective precursor mRNAs were subject to spliced leader RNA *trans*-splicing. The situation is different from other Discoba, a protist group incorporating, in addition to Euglenozoa, lineages such as Heterolobosea, Jakobida, and *Tsukubamonas* [26]. The genome of EUCA likely contained base J (glucosylated hydroxymethyluracil), which is differentially localized within the genomes of all three main euglenozoan clades [16] and participates in RNA polymerase II transcription termination in trypanosomatids [27]. Some euglenozoans are capable of post-transcriptional control of gene expression via RNA interference, which was almost certainly present in EUCA and differentially lost within specific lineages [28].

Organellar Genomes and Gene Expression

Despite encoding only a handful of genes, mitochondrial genomes and transcriptomes are notorious for evolving extreme sizes, topologies, editing, and other RNA processing events [29]. Still, euglenozoans stand out with euglenids harboring a small complement of linear DNA molecules in their mitochondria, while diplonemids evolved the largest mitochondrial genomes known so far in terms of overall DNA content. These genomes are composed of supercoiled noncatenated circles (Figure 2) [30]. In kinetoplastids, the amount of kDNA present is invariably expanded relative to EUCA, with a milestone achieved in *Perkinsela*, where over 90% of total cellular DNA is mitochondrial [31]. In trypanosomatids, the kDNA is composed of relaxed circles catenated into a single network (Figure 2) [32]. The mitochondrial genes of diplonemids are scrambled and transcripts require *trans*-splicing [17], while in kinetoplastids the genes are intact, but encrypted, with extensive RNA editing by uridine insertions and deletions required to produce translatable transcripts [33]. Unexpectedly, the post-transcriptional modification mechanism of kinetoplastids evolved solely in that lineage, while diplonemid mitochondrial transcripts are subject to two types of RNA editing: (i) uridine and adenosine appendage, and (ii) nucleotide substitution (cytidine-to-uridine, adenosine-to-inosine, and guanosine-to-adenosine) [17]. While uridine deletion editing observed in kinetoplastids was not found in diplonemids, uridine addition, carried out in a 'cut-add-reseal' mode in kinetoplastids, is performed by appending uridines to the 3′ ends of certain transcript modules prior to *trans*-splicing in diplonemids [17]. By contrast, euglenids lack obvious traces of editing machinery and transcript processing [34]. Hence, the baroquely complex RNA editing likely emerged in the kinetoplastid common ancestor, and other types of editing arose independently in diplonemids.

Replication and maintenance of kDNA is highly complex when compared with mitochondrial genomes of other eukaryotes. As a consequence, numerous dedicated enzymes, including an expanded family of A DNA polymerases (A–D) are required [35]. The EUCA apparently possessed only a single bacterial-related DNA polymerase A, the additional B and D isoforms arose in the common ancestor of kinetoplastids (Figure 2A), and the C isoform in the common ancestor of eubodonids and trypanosomatids [36]. Most mitochondrial proteins, including those involved in kDNA replication, are encoded by the nuclear genome and imported into the organelle via membrane protein translocation pores [23,37]. The genes encoding subunits of the translocase of

way enzymes. [f]Homologs were also identified in *Naegleria* and several other eukaryotic lineages and, thus, its absence from euglenids and *Perkinsela* are likely secondary losses. [g]We identified two low-confidence hits to mucin-like proteins (KAF0974799 and KAF0976983) in *Naegleria fowleri*. The features' presence/absence for Discoba outside Euglenozoa (an outgroup) is based on data from the literature. The presence of a mitochondrial family A polymerase-like sequence in Discoba is based on data for *Andalucia godoyi* [110]. (B) Time-calibrated tree of Euglenozoa inferred on the alignment of 43 conserved proteins using BEAST2 [111] under relaxed clock and calibrated Yule model as the tree prior (see Box S1 in the supplemental information online for details). *Naegleria gruberi* serves as an outgroup. Nodes are at mean divergence times. Bars and numbers at the nodes represent 95% highest posterior density interval. Ages on the X axis are in million years ago (Mya). Major events in the evolution of the features from panel (A) are marked with the respective numbers in rounded rectangles. Most species-specific gains/losses are omitted for simplicity. The arrow and the star correspond to two calibration points: ~120 Mya – *Paleoleishmania* fossil record [112]; ~450 Mya – fossil record of a euglenid *Moyeria* [113]. All nodes have posterior probability of 1. Names of parasitic/endosymbiotic species are underlined.

inner membrane (TIM) also demonstrate differential evolutionary patterns in Euglenozoa, with TIM17 and TIM23 complexes having undergone independent duplications in both euglenids and diplonemids, while being absent from all known kinetoplastids (Box 1).

Which Traits in Trypanosomatids Represent Parasite-Specific Features?

Increasing evidence indicates that parasites are comparable with their free-living relatives in complexity and metabolic capabilities, but additionally possess elaborate repertoires of virulence-related genes [10,38]. Earlier work involving *B. saltans*, the closest known free-living relative of

Box 1. Mitochondrial Membrane Protein Translocation Pores

The majority of mitochondrial proteins must be imported across the double membranes via dedicated translocation pore-based complexes. For protein transport to the intermembrane space, the EUCA is thought to have possessed an atypical translocase of the outer membrane pore (ATOM40) (Figure IA). Two import channels, formed by three integral proteins of the same family, were likely present on the inner membrane (IM), mediating transport of proteins into either the matrix or IM (Figure IA). Translocase of inner membrane (TIM) 22 traditionally inserts proteins lacking presequences into the IM, while the channel formed by TIM17 and TIM23 translocates presequence-containing proteins, the majority of which are matrix bound. ATOM40 is seemingly not present in the kinetoplastids *Azumiobodo hoyashimi* and *Perkinsela* sp. (Figure IB), nor is any homolog of the TOM40 import channel, which represents an intriguing absence since an outer membrane protein import pore appears essential in all surveyed mitochondriate lineages. TIM17 and TIM23 have been lost from *T. brucei*, with a homolog to TIM22 being utilized for both matrix import and IM insertion of proteins [37,100]. TIM17 and TIM23 are equally absent from all other kinetoplastids, suggesting that this reduction and repurposing occurred in their common ancestor (Figure IA). Such a reduction would unsurprisingly facilitate a streamlining of the import apparatus, which likely became irreversible through the transition to parasitism. Interestingly however, the surveyed mitochondrial proteome of *T. brucei* shows an unreduced complexity with ~1200 verified proteins, the majority of which (956) are also seen in anaerobic bloodstream form cells [101], suggesting that the quantity of imported proteins has not noticeably diminished, as seen in the reduced mitochondria of other parasites [102,103]. The detection of over 1700 proteins in the mitochondrion of *E. gracilis* suggests that other euglenozoan lineages exhibit a complexity unexpected in unicellular organisms [23]. Euglenids and diplonemids additionally retain all three IM integral proteins, with duplications seen in both TIM17 and TIM23 (Figure IA), perhaps responding to increased import capacity for a greater variety of proteins.

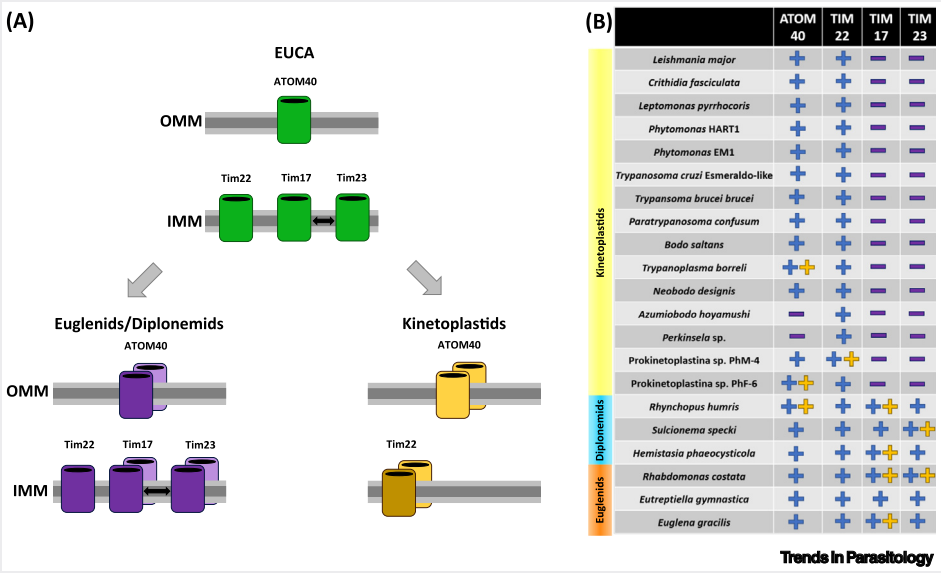

Figure I. Mitochondrial Import Channels in Euglenozoa. (A) Schematic diagram demonstrating mitochondrial evolution of protein translocation from the euglenozoan common ancestor (EUCA), residing on either the outer mitochondrial membrane (OMM) or inner mitochondrial membrane (IMM). Lighter shaded pores represent proteins not seen in all species of a particular clade; black arrows show interaction of proteins within the same import complex. (B) Import channel within euglenozoans: 'blue +' = presence; '–' = absence; 'yellow +' = duplication. Abbreviation: ATOM40, atypical translocase of the outer membrane pore 40.

parasitic trypanosomatids, suggested that losses of metabolic capacity, such as amino acid, purine, folate, and ubiquinone biosynthesis pathways, previously assumed as parasitism-associated reductions, predate the emergence of parasitism in trypanosomatids [1,10,11]. Extensive genomic and transcriptomic datasets from a variety of euglenozoans now support this model [14]. Indeed, analyses of free-living and parasitic kinetoplastids showed that they are auxotrophic for several amino acids, in contrast to diplonemids and euglenids [14]. Furthermore, the EUCA likely possessed a penta-functional AROM protein catalyzing several steps of the shikimate pathway, previously identified in only a few other eukaryotes (Figure 2). For fatty acid biosynthesis, an unusual rendition, involving a set of elongases, is not restricted to trypanosomatids [15] but is likely also functional in free-living bodonids [14], albeit the EUCA likely possessed a conventional fatty acid synthase I, present in diplonemids and euglenids (Figure 2). The cytochrome-independent, cyanide-insensitive alternative oxidases described in trypanosomatids [39] and other eukaryotes, such as fungi and animals [40,41] are not an exclusive feature of euglenozoan parasites as they are identified in diplonemids and euglenids and were most likely present in the EUCA and other early eukaryotes (Figure 2).

By contrast, the unusual compartmentalization of the first six or seven enzymes of glycolysis and other carbohydrate metabolic enzymes into peroxisomes, remodeling this organelle into glycosomes [42], occurred in the common ancestor of diplonemids and kinetoplastids [20,43]. *Euglena gracilis* lacks peroxisome-targeted isoforms of glycolytic enzymes [44], and of ~50 high-confidence glycosomal proteins of *Trypanosoma brucei* [45], none have recognizable homologs in *E. gracilis*, suggesting that the evolution of glycosomes in Euglenozoa goes beyond carbohydrate metabolism [22,44]. The only known plastid acquisition event in Euglenozoa took place in the ancestor of Euglenophyceae, which possess secondary plastids of *Pyramimonas*-like green algal origin [46].

Although the loss of amino acid, nucleotide, vitamin, heme biosynthesis, and many other pathways is not directly connected with parasitism, some of the losses clearly are a consequence of this change. For example, the EUCA was likely endowed with the ability to synthesize paramylon (β-1,3-glucan), a structural and storage polysaccharide in a variety of organisms [47,48]. In Euglenozoa the loss of the paramylon synthetic capacity correlates with the parasitic/endosymbiotic lifestyle (Box 2).

Thus, metabolic reduction in Kinetoplastea was a multistep process, likely unrelated to the singular origin of parasitism in trypanosomatids [14]. In some instances, gene family expansions happened within trypanosomatids, notably cell-surface permeases and transporters [10,49,50]. Only where adaptation to specific niches occurred were switches to parasitism/endosymbiosis obviously accompanied by striking metabolic losses, illustrated by the plant-pathogen *Phytomonas* and the amoeba endosymbiont *Perkinsela* [51–53].

Complex Evolutionary Patterns of Specific Metabolic Features in Euglenozoa

Heme Biosynthesis

Heme biosynthesis is a near universal essential part of metabolism but has a complex history in kinetoplastids. Despite dispensing with so much of the biology that otherwise characterizes kinetoplastids [18] the early-branching *Perkinsela* synthesizes heme [53] while the entire pathway was lost in other kinetoplastids studied in this respect (Figure 2). It is assumed that the free-living bodonids acquire heme from their bacterial prey, whereas the parasitic trypanosomatids obtain it from their host [18]. However, the picture is further complicated by a reacquisition of the last three steps, via HGT, by *Leishmania* and related flagellates, which possibly have access to coproporphyrinogen III of the host. Another source of heme for some trypanosomatids is their

Box 2. Parasitism and (1,3)-β-Glucan: 'Hide and Use It or Lose It'

1,3-β-Glucan (also known as paramylon) is a widely distributed polysaccharide found in some bacteria, euglenozoans, the SAR (Stramenopila, Alveolata, Rhizaria) supergroup, haptophytes, fungi, and plants, where it usually plays storage and structural roles [47,48]. In addition, 1,3-β-glucan is an immunomodulator active in both vertebrates and invertebrates [104], even in plants [105]. Given this role, it might be disadvantageous for parasites to synthesize 1,3-β-glucan, unless it is shielded from the host immune system. While the EUCA likely possessed the ability to synthesize 1,3-β-glucan in a reaction catalyzed by an enzyme belonging to family 48 of glycosyl transferases, this capacity was lost multiple times in parasitic lineages (Figure I). Indeed, the same holds true for the genus *Naegleria*, where the free-living *Naegleria gruberi* preserves the 1,3-β-glucan synthase, and the causative agent of amoebic meningoencephalitis, *Naegleria fowleri*, lost it, while its loss in *Perkinsela* might rather be a consequence of massive genome reduction [53]. Interestingly, since diplonemids synthesize 1,3-β-glucan in nutrient-poor conditions [47], we suggest that, similar to diatoms [106], diplonemids use it for lowering cell buoyancy, leading to submersion in more nutrient-rich layers of the ocean. In contrast to Euglenozoa, parasites from several eukaryotic lineages, where 1,3-β-glucan plays an important structural role, tend to preserve its biosynthesis and hide it from the host's immune system. For example, 1,3-β-glucan is found in the inner layer of the oocyst wall of *Toxoplasma* but is lacking in its tissue stages [107]. Similarly, pathogenic filamentous fungi use α-1,3-glucan to conceal 1,3-β-glucan in their cell wall [108]. With the growing number of available genome sequences, it would be interesting to study the distribution of glucan synthase genes and its correlation with the lifestyle, evolution of glucan synthase domain architecture, and possible functions of 1,3-β-glucan.

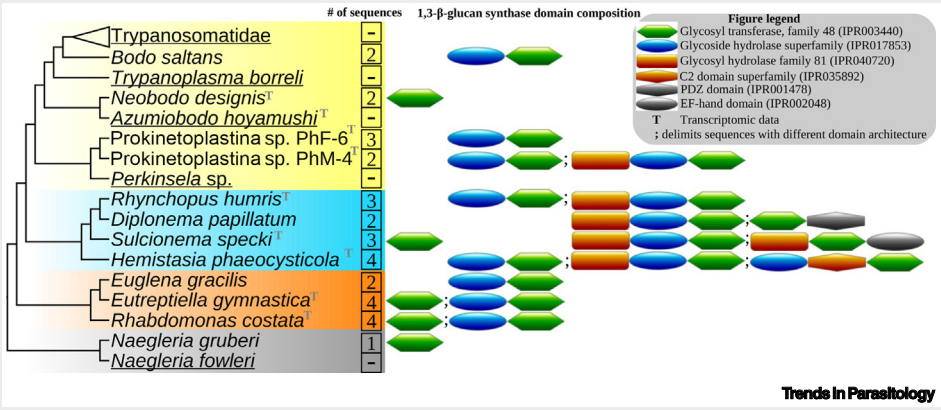

Figure I. Presence/Absence Pattern of Genes Encoding 1,3-β-Glucan Synthase in Euglenozoa and *Naegleria*. The number of identified protein sequences and their domain composition is shown. The ability to synthesize (1,3)-β-glucan in Euglenozoa, and its closest relatives with sequenced genomes, *Naegleria* spp., appears to correlate with the lifestyle: free-living organisms possess it while pathogens and symbionts have lost it. Species names of pathogenic and symbiotic organisms are underlined. Abbreviation: T, transcriptomic data.

endosymbiont [54] . Finally, the enzyme of the last step of heme biosynthesis, ferrochelatase, was repeatedly acquired via HGT, but its role remains uncertain [18].

ROS Detoxifying Systems

Aerobic metabolism is always connected to the production of dangerous ROS, which can oxidize proteins, lipids, and nucleic acids [55,56]. With varying life cycles and environments across Euglenozoa, it is likely that specific taxa have significantly different levels of ROS production, as well as mechanisms for coping with them [57].

Evolution of ROS-detoxifying systems in Euglenozoa is complex and is shaped by a plethora of factors [14,57–59]. The ROS protection system of *Euglena* relies on a combination of glutathione, thioredoxin, and trypanothione NADPH-dependent thiol-redox systems but lacks a heme-containing catalase, and likely resembles the ancestral state in the EUCA (Figure 2). Parasitic trypanosomatids, dependent solely on the trypanothione-based system for ROS protection, have lost both the glutathione and thioredoxin reductase-based systems [14]. Other kinetoplastids show an intermediate state on the way to the minimized NADPH-dependent

thiol-redox systems of trypanosomatids, having lost glutathione but retained thioredoxin reductase.

The distribution of catalase, a unique enzyme for direct ROS detoxification, is patchy and its evolution is complex. Its absence from euglenids [60] could imply that it was not present in the EUCA (Figure 2) and was acquired via HGT from a eukaryote by a common ancestor of diplonemids and kinetoplastids. While retained by several diplonemids, the catalase gene was replaced by an α-proteobacterial homolog in other diplonemids [57]. Moreover, catalase was absent from the common ancestor of kinetoplastids, yet independently acquired at least twice by trypanosomatids from different bacterial groups [57,59]. Taken together, this testifies to the importance of catalase in euglenozoan evolution and documents a tendency to functionally substitute catalase with bacterial homologs, likely from those that they prey upon.

Composition and Evolution of Selected Protein Complexes

Nuclear Pore Complex and Lamina

From the nucleus, we focus on the **nuclear pore complex** (**NPC**), the site of nucleocytoplasmic transport, the nuclear lamina, the structural and spatial organizer, and the kinetochore, the link between chromosomes and the mitotic spindle, since each has been the subject of recent investigations in kinetoplastids [61,62].

The diplonemid lamina is at least partly conserved in kinetoplastids due to the presence of homologs for the major components NUP-1 and NUP-2, which, by contrast, are missing from *E. gracilis* [63]. The diplonemid NPC (Figure 3) demonstrates similarities with the *E. gracilis* state [22] as both β-propeller proteins (Seh1 and ALADIN) are present and, importantly, the former is lost from trypanosomes (Figure 3A,B). The gating NPC components of diplonemids more closely resemble those in *E. gracilis* than in *T. brucei* (Figure 3C). However, this interpretation needs to be treated with caution as precise homology for NPC proteins is extremely difficult to assign, even for closely related species [64,65]. Most significant is the absence of the mRNA export factor DBP5 from diplonemids and kinetoplastids, a key component of late steps of RNA processing [66], present in *E. gracilis*. Hence, the NUP-1 lamina is a shared feature of diplonemids and kinetoplastids, indicating its origin at the common ancestor of these two lineages, and additionally suggesting that this configuration is neither parasitism-associated nor associated with the absence of *cis*-splicing.

Kinetochore

The kinetochore, a complex of over 60 proteins, is organized in a modular fashion and is well conserved across most eukaryotes [67]. For detailed information on the structure and functions of kinetochores in kinetoplastids and other eukaryotes the reader is referred to the excellent reviews on the topic [67,68]. The sole known example of wholesale kinetochore replacement are kinetoplastids, where a simpler structure of ~30 proteins replaces the canonical form [62,69]. In contrast to kinetoplastids, homologs of the centromeric histone H3 are present in diplonemids and possibly involved in kinetochore assembly [14]. Little else is currently known about the apparently enigmatic diplonemid kinetochores, with their composition and structure waiting to be elucidated. Intriguingly, the presence of a few kinetoplastid kinetochore (KKT) genes, as well as specific regulatory proteins of kinase (KKT10, KKT19), and phosphatase (KKIP7) families in diplonemids and euglenids suggests that parts of the kinetoplastid system may have been represented more broadly amongst the discobid protists. However, it remains unclear whether these proteins are components of diplonemid or euglenid kinetochores.

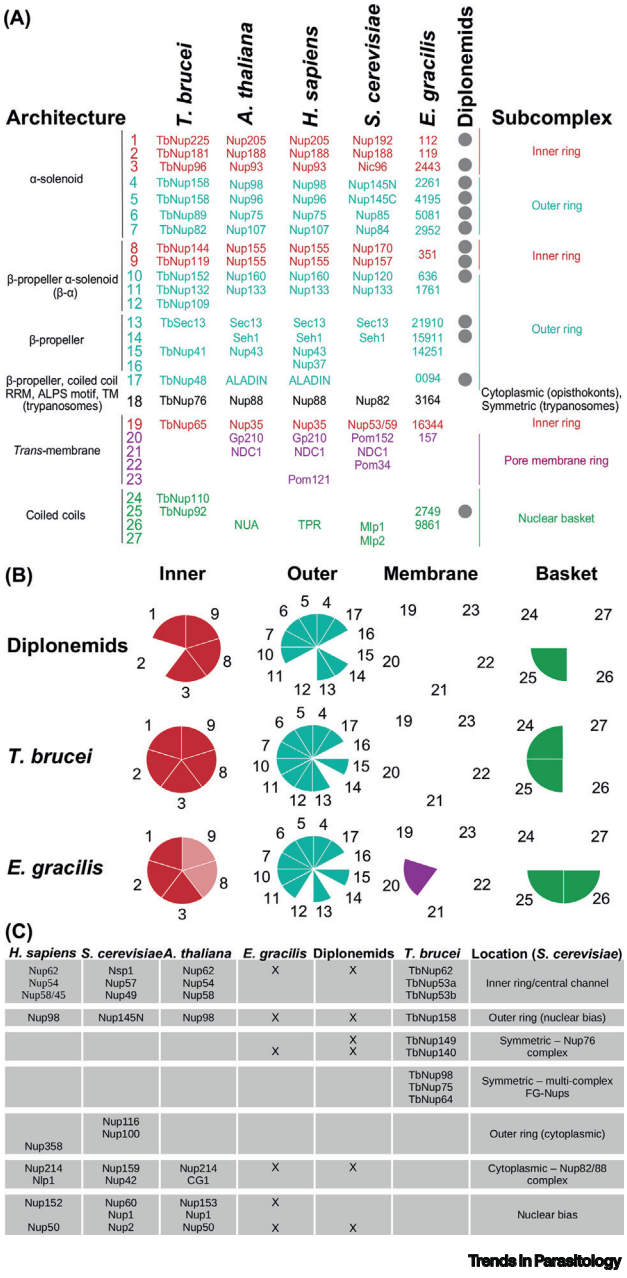

Figure 3. Nuclear Pore Complex (NPC). (A) An architecture of the NPC subunits in Euglenozoa, opisthokonts (*Homo sapiens* and *Saccharomyces cerevisiae*) and plants (*Arabidopsis thaliana*). The presence of the respective proteins in diplonemids is marked with gray circles. (B) A composition of the inner and outer ring complexes, pore membrane ring, and nuclear basket of the NPC in diplonemids, *Trypanosoma brucei*, and *Euglena gracilis*. Transparency of segments is lowest for the best hits. Subunit numbering at panel A corresponds to the numbering at panel B. (C) Phenylalanine–glycine family of the NPC proteins in *T. brucei*, *E. gracilis* and diplonemids, S. *cerevisiae*, *H. sapiens*, and *A. thaliana*. The presence of the respective homologs in *E. gracilis* and diplonemids is indicated with the 'x' sign. Diplonemid NPC appears to be more similar to its euglenid counterpart than to the NPC of *T. brucei*. On the contrary, the lamina of diplonemids resembles more the lamina of *T. brucei* than that of *E. gracilis*. Abbreviation: TM, transmembrane domain.

Histone Octamer and H1

Trypanosomatid histones, reviewed in [70,71], are amongst the most divergent, with histone H1 lacking even the histone fold domain. In contrast to canonical euglenid sequences, diplonemid histone H1 also appears to be highly divergent since it could not be identified in our dataset, even when employing sensitive bioinformatic searches. Moreover, diplonemid core histones H2B and H4 are divergent from both their respective homologs in other euglenozoans, as well as animals, fungi, and plants (Figures S1–S5 in the supplemental information online). Diplonemid histones also demonstrate lineage-specific expansions, resulting in examples of considerably sized paralog families, indicating novel chromatin-regulatory mechanisms.

Insights into Evolution of Virulence-Related Genes

Cell-Surface Proteins

Across eukaryotes, surface composition is highly attuned towards supporting lifestyles, participating in immune evasion, cell invasion, migration, and environmental sensing, yet cell surfaces exhibit considerable diversity, even between closely related taxa [72,73]. Amastins are a family of surface glycoproteins that are present in all kinetoplastids (except *Perkinsela*) and, while having poorly understood functions, are evolutionarily flexible [74]. Amastins encoded by a large gene family in trypanosomatids typically contain four transmembrane domains and are expressed mainly during the intracellular amastigote stage in *Trypanosoma cruzi* and *Leishmania* spp. [74]. The presence of amastin domain-containing proteins of unknown function (all with predicted transmembrane helices characteristic for trypanosomatid amastins) in euglenids and diplonemids (Figure 2, Figure S6 in the supplemental information online) prompts us to suggest their presence in the EUCA. Furthermore, the gp63 protease family (leishmanolysin), an important trypanosomatid virulence factor playing multiple roles in both an insect vector and a vertebrate host [75], is expanded in diplonemids but was not identified in euglenids and *Perkinsela* (Figure S7 in the supplemental information online). Given its paneukaryotic presence [76], gp63 was likely present in the EUCA and secondarily lost from both euglenids and the minimized *Perkinsela*. In addition to amastins and gp63 proteases, trypanosomes incorporate variant surface glycoproteins (VSGs), mucins, *trans*-sialidases and proteins encoded by expression site-associated genes to their membrane [72,77]. Of major surface proteins described in trypanosomes, diplonemids carry only ESP23, which is a conserved multipass protein of unknown function [78]. We failed to detect ancestral homologs of mucins, *trans*-sialidases, and the VSG families in diplonemids and euglenids, which rose to prominence in the African trypanosomes [79] (Figure 2A).

Subtilisin and Metallocarboxypeptidases

Secreted peptidases are important contributors to kinetoplastid parasitism. The broad families of subtilisins and metallocarboxypeptidases are involved in a variety of functions, including proteolysis and enzyme activation. Subtilisins mediate survival of *Leishmania* in macrophages with a molecular function as a maturase for tryparedoxin peroxidase [80]. Some diplonemids massively expanded their subtilisins, euglenids show large variation in total counts between species, while kinetoplastids have reduced numbers (Figure 4). It is striking that this protein family, which contributes towards parasitism, is counterintuitively subject to reduction, particularly in parasitic species. The subcellular distribution of subtilisins is predicted as primarily in the secretory pathway, tethered to the cell membrane or cytoplasm, though a number are additionally predicted in organellar compartments. Kinetoplastids have a specific reduction in these latter subtilisins, while those associated with parasitism appear targeted to the surface membrane (Figure 4).

The M14 metallocarboxypeptidases are a smaller family. On average, diplonemids have the largest number of paralogs but prokinetoplastids show the highest total (Figure 4). Varied numbers of carboxypeptidases are detected in both euglenids and free-living kinetoplastids, followed by a noticeable reduction within their parasitic kin (Figure 4). Parasitism is also accompanied by targeting changes, with secreted carboxypeptidases being almost entirely absent, and favoring cytoplasmic and nuclear paralogs (Figure 4). Since some parasites, such as *Plasmodium*, make use of the host's processing enzymes [81], it is plausible that trypanosomatids have developed a similar dependency that accompanies this loss in their genomes.

Of particular interest is the M32 metallocarboxypeptidase family, scarcely seen in eukaryotes, but noticeably reported in trypanosomes and considered attractive drug targets [82]. Several parasitic lineages show significant expansions of this family, localized primarily to the cytoplasm

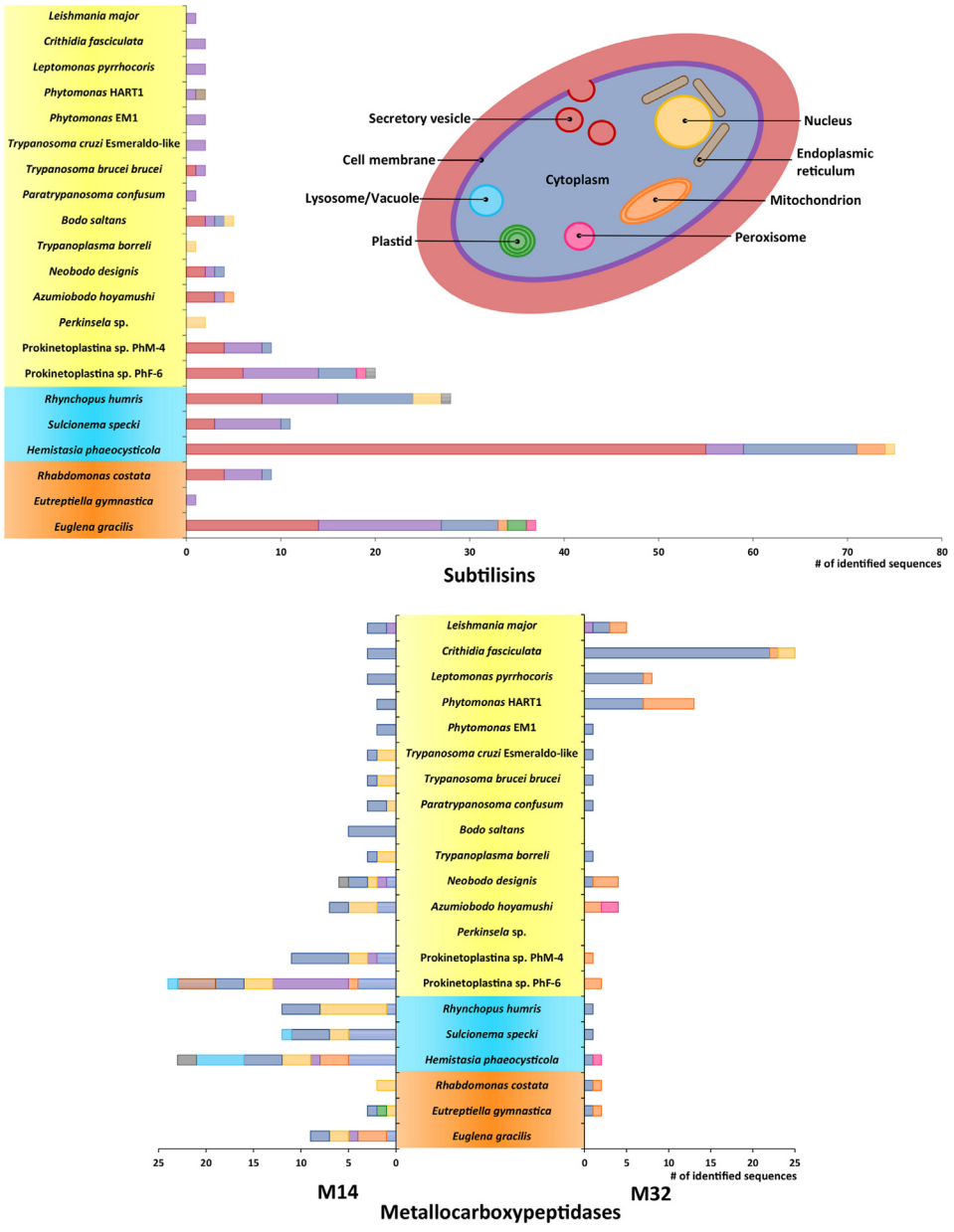

Figure 4. Comparative Distribution of Various Peptidases within Euglenozoa. Distribution of subtilisins and M14 and M32 metallocarboxypeptidases among selected species, with predicted subcellular localization indicated (blue: cytoplasm; brown: endoplasmic reticulum; green: plastid; light blue: lysosome/vacuole; orange: mitochondrion; pink: peroxisome; purple: cell membrane; red: secreted; yellow: nucleus; gray: other). Subtilisins and metallocarboxypeptidases are virulence factors described in trypanosomatids. Counterintuitively, the family of subtilisins is reduced in the kinetoplastid parasites compared with the free-living diplonemids, euglenids, and prokinetoplastids. A similar trend is observed for M14 metallocarboxypeptidases, while the M32 family is expanded in several parasitic lineages. The *in silico*-determined distribution of subtilisins localized them primarily in the secretory pathway, tethered to the cell membrane or cytoplasm. The M32 peptidases appear localized primarily to the cytoplasm or mitochondrion.

and mitochondrion, which contrasts with their minimal counts or even absence in the free-living representatives (Figure 4).

Cell Form and Differentiation

Striking features of trypanosomatid cell biology are the cellular morphological and biochemical alterations interpreted as a facilitating adaptation to the specific host niches, attachment to surfaces [3,83], or involved in meiosis and recombination [84]. Trypanosomatid cell forms are historically described by whether the kDNA and associated flagellar basal bodies lie anterior or posterior to the nucleus and the extent to which the flagellum is attached to the cell body beyond the flagellar pocket. Thus, pivotal to changes in cell form are flagella and flagellum attachment-zone lengths, ranging from the short sensory flagellum built by intracellular amastigotes [85] to the long attached flagella of mesocyclic and long epimastigote stages of *T. brucei* found within tsetse flies [86]. By contrast, we tend to consider the phagotrophic bodonids from the perspective of less dynamic cell morphologies: at one extreme the endosymbiont *Perkinsela* has dispensed with a flagellum-associated cytoskeleton, making morphological transitions difficult [53], but even the demonstration of cyst-like morphologies are limited to a handful of reports [87,88] and these forms appear not to be true cysts. This leaves an open question of whether encystment is a trait generally lacking from euglenozoans.

Yet, do the complex patterns of life cycle differentiation and cytoskeletal rearrangements, which are seen in trypanosomatids, represent traits derived following, or even facilitating, the transition to obligate parasitism or perhaps complex dixenous parasitism? Recent acquisition into culture and resultant cell biology characterizations of novel diplonemid taxa [89–91] suggest that the answer to this question is probably 'no' since patterns of cellular differentiation and dynamic changes in cell morphology are evident as these diplonemids age in batch culture or are subject to starvation. Thus, in response to environmental cues, some diplonemids adopt sessile conformations, attaching to surfaces via gelatinous cell coats, remodeling flagellum ultrastructure with concomitant changes to swimming behavior, or assembling **lacunae** and **extrusomes** in response to extrinsic factors. In other diplonemids, extrusomes are always evident [89]. Collectively, however, the changes in form that have been documented for several diplonemid taxa provide a morphological diversity that begins to approach that seen in trypanosomatids. Observations that some diplonemids build a paraflagellar rod, a key synapomorphy for Euglenozoa, only at some points in their life cycles, and of starvation-triggered attachment of diplonemids to surfaces, suggest that the roots of trypanosomatid attachment to epithelial surfaces in invertebrate vectors lie deep within euglenozoan evolution.

Returning to the context of trypanosomatid life cycles, there are examples of cytoskeletal structures, such as the cytostome–cytopharynx complex, retained in only some trypanosomatid species or assembled in only some life stages [3,92]. However, regarding the different morphologies of trypomastigote, epimastigote, and promastigote cell shapes, providing the flagellum–cell surface attachment remains intact, then changes in the expression of only a single cytoskeletal protein can be sufficient to drive changes in cell morphology [93–95]. Moreover, changes in the expression of a single RNA-binding protein – a principal mechanism through which gene expression is regulated in trypanosomatids [96] – can be sufficient to trigger the completion of complex differentiation programs in their entireties [97]. This suggests that, in trypanosomatids, regulatory hierarchies underpinning complex changes in the expression levels of many proteins may be triggered by relatively simple 'master switches'. This attests to a longstanding view that regulatory networks controlling gene expression in parasites are generally streamlined compared with those in the free-living taxa. With the exception of some surface coat proteins in the African trypanosomes, trypanosomatids, and seemingly also bodonids, lack gene regulation at the level of

transcription. It will, thus, be fascinating to determine the extent to which regulatory cascades controlling wide-scale changes in gene expression within euglenozoans differ in complexity between the gene-plentiful free-living taxa and the parasites which typically encode fewer genes within their nuclear genomes.

Concluding Remarks

Euglenozoa, unifying parasitic trypanosomatids with presumably free-living heterotrophic diplonemids and photosynthetic or heterotrophic euglenids, represent a tractable group for understanding lifestyle evolution. Having in hand a robust phylogenetic framework as well as sequence data from all these groups, we can readily study the evolution of underlying features making the switches between lifestyles possible. We can also distinguish parasite-specific features from the ancestral euglenozoan traits present already in the free-living predecessors of the contemporary parasites.

Thus, polycistronic transcription, ubiquitous *trans*-splicing, base J, certain surface molecules (amastin and gp63 protease), subtilisins and peptidases, a trypanothione-based ROS-detoxifying system, the nutrient-triggered attachment to surfaces, and alternative energy-generating mechanisms are examples of features likely already present in the EUCA, which was metabolically versatile, similar to diplonemids. The common ancestor of kinetoplastids apparently sustained metabolic losses and possibly a loss of most introns, while genome streamlining, and more drastic metabolic losses occurred subsequently in the obligatory parasitic trypanosomatids. Although for other prominent features, such as expanded kDNA with bizarre architecture, complex RNA editing, and divergent mRNA processing mechanisms, no clear evolutionary pattern can currently be derived, the comparative analysis is informative. It is evident that the emergence of some features so far associated with parasitism in trypanosomatids actually predated its emergence. The homologs of certain trypanosomatid virulence-related genes are not only present, but sometimes even expanded in diplonemids and euglenids. While the role of respective proteins remains unknown in the free-living euglenozoans, they have been repurposed in the obligatory parasitic trypanosomatids to facilitate host–pathogen interactions. We propose that the emergence of parasitism in euglenozoans may have been facilitated by the capacity of generalists to 'spin-out' parasitic specialists with great efficiency. While there are a few ancestral euglenozoan gene families, which have undergone expansion [1,10] as the switch to parasitism happened in trypanosomatids (e.g., amastins, gp63, subtilisins and peptidases, cell-surface permeases and transporters), it is challenging to delineate those features the emergence of which coincides with the origin of parasitism in this family. Some traits, however, were lost presumably as a consequence of the parasitic lifestyle (e.g., 1,3-β-glucan biosynthesis in parasitic Euglenozoa and glutathione-based ROS-protection system in trypanosomatids). Widely variable lifestyles of kinetoplastids and the scarcity of information on free-living representatives of the clade make assumptions about the lifestyle of their common ancestor highly speculative. We suggest that the common ancestor of Kinetoplastea was a free-living organism obtaining all necessary nutrients from its prey [possibly bacteria and/or other protists, given its existence ~1000 million years ago (Mya) by our estimates], which enabled metabolic losses. Common ancestors of currently existing parasitic groups could form accidental, most likely transient, symbiotic or commensal relationships with an unknown host or different hosts. Such interactions could have been facilitated by the ancestrally present ability of nutrient-triggered attachment to cell surfaces, as well as by ancestral subtilisins, peptidases, cell-surface proteins which, in certain cases, have evolved into host–parasite relationships.

A number of questions remain unanswered (see Outstanding Questions). The evolutionary forces shaping euglenozoan genomes and causality of events triggering multiple switches to parasitism,

Outstanding Questions

What were the drivers for genome streamlining and the loss of various metabolic pathways in supposedly free-living ancestral kinetoplastids? Could the ancestor of kinetoplastids have been a symbiont? Was it a specialist in a narrow-niche environment, getting all the required nutrients from its prey? What factors facilitated multiple switches to parasitism/symbiosis in kinetoplastids?

What roles do the homologs of proteins connected to virulence in the parasitic trypanosomatids play in the free-living euglenozoans?

What evolutionary forces led to the development of bizarre mitochondrial genome architectures and modes of RNA editing in diplonemids and kinetoplastids?

What is the composition, structure, and origin of the diplonemid kinetochore?

What are the mechanisms controlling gene expression and life cycle progression in the free-living euglenozoans?

as well as the development of a number of bizarre features of the euglenozoan mitochondrion, are not clear. The concurrent presence of many highly derived traits, such as kinetochore, histones, and the prereplication complex in kinetoplastids and diplonemids, implies that the high level of divergence of these (otherwise conserved) systems is not directly related to a relaxed selective pressure associated with parasitism. Among trypanosomatids, unconstrained evolution of at least one essential cytoskeletal protein, such that homologs are recognizable only on the basis of gene synteny and functional characterization, rather than on sequence conservation [98], raises a prospect that a notable proportion of seemingly unrelated hypothetical genes in diverse euglenozoans may turn out to share a common ancestry. The availability of high-quality genomic and transcriptomic data for a representative set of diplonemids, as well as experimental data from genetically tractable *Diplonema papillatum* [20,99], will be crucial for answering several important evolutionary questions connected to the origin of parasitism.

Acknowledgments

We are grateful to Ambar Kachale (Institute of Parasitology) for help with the figures, Gertraud Burger (University of Montreal) for sharing *D. papillatum* histone sequences, and Fred Opperdoes (de Duve Institute) for helpful discussions. This work was supported by the Czech Grant Agency 18-15962S and 20-07186S, ERC CZ LL1601, and the ERD Funds project OPVVV 0000759. M.C.F. is a Welcome Trust Investigator.

Supplemental Information

Supplemental information associated with this article can be found online at https://doi.org/10.1016/j.pt.2020.10.001

References

1. Janouškovec, J. and Keeling, P.J. (2016) Evolution: causality and the origin of parasitism. *Curr. Biol.* 26, R174–R177
2. Harmer, J. *et al.* (2018) Farming, slaving and enslavement: histories of endosymbioses during kinetoplastid evolution. *Parasitology* 145, 1311–1323
3. Skalický, T. *et al.* (2017) Extensive flagellar remodeling during the complex life cycle of *Paratrypanosoma*, an early-branching trypanosomatid. *Proc. Natl. Acad. Sci. U. S. A.* 114, 11757–11762
4. Yazaki, E. *et al.* (2017) Global Kinetoplastea phylogeny inferred from a large-scale multigene alignment including parasitic species for better understanding transitions from a free-living to a parasitic lifestyle. *Genes Genet. Syst.* 92, 35–42
5. Lukeš, J. *et al.* (2014) Evolution of parasitism in kinetoplastid flagellates. *Mol. Biochem. Parasitol.* 195, 115–122
6. Leander, B.S. *et al.* (2007) Macroevolution of complex cytoskeletal systems in euglenids. *BioEssays* 29, 987–1000
7. Novák Vanclová, A.M.G. *et al.* (2020) Metabolic quirks and the colourful history of the *Euglena gracilis* secondary plastid. *New Phytol.* 225, 1578–1592
8. Flegontova, O. *et al.* (2016) Extreme diversity of diplonemid eukaryotes in the ocean. *Curr. Biol.* 26, 3060–3065
9. Yubuki, N. and Leander, B.S. (2018) Diversity and evolutionary history of the Symbiontida (Euglenozoa). *Front. Ecol. Evol.* 6, 100
10. Jackson, A.P. *et al.* (2016) Kinetoplastid phylogenomics reveals the evolutionary innovations associated with the origins of parasitism. *Curr. Biol.* 26, 161–172
11. Opperdoes, F.R. *et al.* (2016) Comparative metabolism of free-living *Bodo saltans* and parasitic trypanosomatids. *J. Euk. Microbiol.* 63, 657–678
12. Flegontov, P. *et al.* (2013) *Paratrypanosoma* is a novel early-branching trypanosomatid. *Curr. Biol.* 23, 1787–1793
13. Flegontov, P. *et al.* (2016) Genome of *Leptomonas pyrrhocoris*: a high-quality reference for monoxenous trypanosomatids and new insights into evolution of *Leishmania*. *Sci. Rep.* 6, 23704
14. Butenko, A. *et al.* (2020) Evolution of metabolic capabilities and molecular features of diplonemids, kinetoplastids, and euglenids. *BMC Biol.* 18, 23
15. Lee, S.H. *et al.* (2006) Fatty acid synthesis by elongases in trypanosomes. *Cell* 126, 691–699
16. Dooijes, D. *et al.* (2000) Base J originally found in kinetoplastida is also a minor constituent of nuclear DNA of *Euglena gracilis*. *Nucleic Acids Res.* 28, 3017–3021
17. Kaur, B. *et al.* (2020) Gene fragmentation and RNA editing without borders: eccentric mitochondrial genomes of diplonemids. *Nucleic Acids Res.* 48, 2694–2708
18. Cenci, U. *et al.* (2016) Heme pathway evolution in kinetoplastid protists. *BMC Evol. Biol.* 16, 109
19. Montrichard, F. *et al.* (1999) Evidence for the co-existence of glutathione reductase and trypanothione reductase in the non-trypanosomatid Euglenozoa: *Euglena gracilis* Z. *FEBS Lett.* 442, 29–33
20. Morales, J. *et al.* (2016) Differential remodelling of peroxisome function underpins the environmental and metabolic adaptability of diplonemids and kinetoplastids. *Proc. R. Soc. B* 283, 20160520
21. Milanowski, R. *et al.* (2014) Distribution of conventional and nonconventional introns in tubulin (alpha and beta) genes of euglenids. *Mol. Biol. Evol.* 31, 584–593
22. Ebenezer, T.E. *et al.* (2019) Transcriptome, proteome and draft genome of *Euglena gracilis*. *BMC Biol.* 17, 11
23. Hammond, M.J. *et al.* (2020) A uniquely complex mitochondrial proteome from *Euglena gracilis*. *Mol. Biol. Evol.* 37, 2173–2191
24. Lukeš, J. *et al.* (2018) Trypanosomatids are much more than just trypanosomes: clues from the expanded family tree. *Trends Parasitol.* 34, 466–480
25. Maslov, D.A. *et al.* (2019) Recent advances in trypanosomatid research: genome organization, expression, metabolism, taxonomy and evolution. *Parasitology* 146, 1–27
26. Burki, F. *et al.* (2020) The new tree of eukaryotes. *Trends Ecol. Evol.* 35, 43–55
27. Reynolds, D. *et al.* (2014) Regulation of transcription termination by glucosylated hydroxymethyluracil, base J, in *Leishmania major* and *Trypanosoma brucei*. *Nucleic Acids Res.* 42, 9717–9729
28. Lye, L.F. *et al.* (2010) Retention and loss of RNA interference pathways in trypanosomatid protozoans. *PLoS Pathog.* 6, e1001161
29. Smith, D.R. and Keeling, P.J. (2015) Mitochondrial and plastid genome architecture: reoccurring themes, but significant differences at the extremes. *Proc. Natl. Acad. Sci. U. S. A.* 112, 10177–10184
30. Burger, G. and Valach, M. (2018) Perfection of eccentricity: mitochondrial genomes of diplonemids. *IUBMB Life* 70, 1197–1206

31. Lukeš, J. *et al.* (2018) Massive mitochondrial DNA content in diplonemid and kinetoplastid protists. *IUBMB Life* 70, 1267–1274
32. Lukeš, J. *et al.* (2002) Kinetoplast DNA network: evolution of an improbable structure. *Eukaryot. Cell* 1, 495–502
33. Aphasizheva, I. *et al.* (2020) Lexis and grammar of mitochondrial RNA processing in trypanosomes. *Trends Parasitol.* 36, 337–355
34. Dobáková, E. *et al.* (2015) Unexpectedly streamlined mitochondrial genome of the euglenozoan *Euglena gracilis*. *Genome Biol. Evol.* 7, 3358–3367
35. Jensen, R.E. and Englund, P.T. (2012) Network news: the replication of kinetoplast DNA. *Annu. Rev. Microbiol.* 66, 473–491
36. Harada, R. *et al.* (2020) Inventory and evolution of mitochondrion-localized family A DNA polymerases in Euglenozoa. *Pathogens* 9, 257
37. Harsman, A. and Schneider, A. (2017) Mitochondrial protein import in trypanosomes: expect the unexpected. *Traffic* 18, 96–109
38. Janouškovec, J. *et al.* (2015) Factors mediating plastid dependency and the origins of parasitism in apicomplexans and their close relatives. *Proc. Natl. Acad. Sci. U. S. A.* 112, 10200–10207
39. Chaudhuri, M. *et al.* (2006) Trypanosome alternative oxidase: from molecule to function. *Trends Parasitol.* 22, 484–491
40. Ebiloma, G.U. *et al.* (2019) Alternative oxidase inhibitors: mitochondrion-targeting as a strategy for new drugs against pathogenic parasites and fungi. *Med. Res. Rev.* 39, 1553–1602
41. Weaver, R.J. (2019) Hypothesized evolutionary consequences of the alternative oxidase (AOX) in animal mitochondria. *Integr. Comp. Biol.* 59, 994–1004
42. Gabaldón, T. *et al.* (2016) Peroxisomes in parasitic protists. *Mol. Biochem. Parasitol.* 209, 35–45
43. Durrani, H. *et al.* (2020) A global analysis of enzyme compartmentalization to glycosomes. *Pathogens* 9, 281
44. Szöőr, B. *et al.* (2014) Evolution, dynamics and specialized functions of glycosomes in metabolism and development of trypanosomatids. *Curr. Opin. Microbiol.* 22, 79–87
45. Guther, M.L.S. *et al.* (2014) High-confidence glycosome proteome for procyclic form *Trypanosoma brucei* by epitope-tag organelle enrichment and SILAC proteomics. *J. Proteome Res.* 13, 2796–2806
46. Leander, B.S. (2004) Did trypanosomatid parasites have photosynthetic ancestors? *Trends Microbiol.* 12, 251–258
47. Škodová-Sveráková, I. *et al.* (2020) Unique dynamics of paramylon storage in the marine euglenozoan *Diplonema papillatum*. *Protist* 171, 125717
48. Harris, P.J. and Stone, B.A. (2009) Evolutionary aspects of (1,3)-β-glucans and related polysaccharides. In *Chemistry, Biochemistry, and Biology of (1,3)-β-Glucans and Related Polysaccharides* (Bacic, A. *et al.*, eds), pp. 655–662, Academic Press
49. Barrett, M.P. *et al.* (1995) Functional expression and characterization of the *Trypanosoma brucei* procyclic glucose transporter, THT2. *Biochem. J.* 312, 687–691
50. Burchmore, R.J. *et al.* (2003) Genetic characterization of glucose transporter function in *Leishmania mexicana*. *Proc. Natl. Acad. Sci. U. S. A.* 100, 3901–3906
51. Porcel, B.M. *et al.* (2014) The streamlined genome of *Phytomonas* spp. relative to human pathogenic kinetoplastids reveals a parasite tailored for plants. *PLoS Genet.* 10, e1004007
52. Butler, C.E. *et al.* (2017) Genome sequence of *Phytomonas françai*, a cassava (*Manihot esculenta*) latex parasite. *Genome Announc.* 5, e01266-01216
53. Tanifuji, G. *et al.* (2017) Genome sequencing reveals metabolic and cellular interdependence in an amoeba-kinetoplastid symbiosis. *Sci. Rep.* 7, 11688
54. Kořený, L. *et al.* (2013) Make it, take it, or leave it: heme metabolism of parasites. *PLoS Pathog.* 9, e1003088
55. Speijer, D. *et al.* (2020) Comparing early eukaryotic integration of mitochondria and chloroplasts in the light of internal ROS challenges: timing is of the essence. *mBio* 11, e00955-20
56. Gill, S.S. and Tuteja, N. (2010) Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. *Plant Physiol. Biochem.* 48, 909–930
57. Škodová-Sveráková, I. *et al.* (2020) Catalase and ascorbate peroxidase in euglenozoan protists. *Pathogens* 9, 317
58. Kraeva, N. *et al.* (2017) Catalase in Leishmaniinae: with me or against me? *Infect. Genet. Evol.* 50, 121–127
59. Bianchi, C. *et al.* (2019) An enigmatic catalase of *Blastocrithidia*. *Mol. Biochem. Parasitol.* 232, 111199
60. Vlasits, J. *et al.* (2007) Hydrogen peroxide oxidation by catalase-peroxidase follows a non-scrambling mechanism. *FEBS Lett.* 581, 320–324
61. Rout, M.P. *et al.* (2017) Specialising the parasite nucleus: pores, lamins, chromatin, and diversity. *PLoS Pathog.* 13, e1006170
62. Akiyoshi, B. and Gull, K. (2014) Discovery of unconventional kinetochores in kinetoplastids. *Cell* 156, 1247–1258
63. Kořený, L. and Field, M.C. (2016) Ancient eukaryotic origin and evolutionary plasticity of nuclear lamina. *Genome Biol. Evol.* 8, 2663–2671
64. Mans, B.J. *et al.* (2004) Comparative genomics, evolution and origins of the nuclear envelope and nuclear pore complex. *Cell Cycle* 3, 1612–1637
65. Chopra, K. *et al.* (2019) Evolutionary divergence of the nuclear pore complex from fungi to metazoans. *Protein Sci.* 28, 571–586
66. Lin, D.H. *et al.* (2018) Structural and functional analysis of mRNA export regulation by the nuclear pore complex. *Nat. Commun.* 9, 2319
67. Hamilton, G. *et al.* (2019) Seeing is believing: our evolving view of kinetochore structure, composition, and assembly. *Curr. Opin. Cell Biol.* 60, 44–52
68. Akiyoshi, B. (2016) The unconventional kinetoplastid kinetochore: from discovery toward functional understanding. *Biochem. Soc. Trans.* 44, 1201–1217
69. Nerusheva, O.O. *et al.* (2019) Identification of four unconventional kinetoplastid kinetochore proteins KKT22-25 in *Trypanosoma brucei*. *Open Biol.* 9, 190236
70. Alsford, S. and Horn, D. (2004) Trypanosomatid histones. *Mol. Microbiol.* 53, 365–372
71. Saha, S. (2020) Histone modifications and other facets of epigenetic regulation in trypanosomatids: leaving their mark. *mBio* 11, e01079-20
72. Jackson, A.P. *et al.* (2013) A cell-surface phylome for African trypanosomes. *PLoS Negl. Trop. Dis.* 7, e2121
73. Jackson, A.P. (2016) Gene family phylogeny and the evolution of parasite cell surfaces. *Mol. Biochem. Parasitol.* 209, 64–75
74. Jackson, A.P. (2010) The evolution of amastin surface glycoproteins in trypanosomatid parasites. *Mol. Biol. Evol.* 27, 33–45
75. Castro Neto, A.L. *et al.* (2019) *In silico* characterization of multiple genes encoding the GP63 virulence protein from *Leishmania braziliensis*: identification of sources of variation and putative roles in immune evasion. *BMC Genomics* 20, 118
76. Manna, P.T. *et al.* (2013) Adaptin evolution in kinetoplastids and emergence of the variant surface glycoprotein coat in African trypanosomatids. *Mol. Phylogenet. Evol.* 67, 123–128
77. Pech-Canul, A.C. *et al.* (2017) A brief view of the surface membrane proteins from *Trypanosoma cruzi*. *J. Parasitol. Res.* 2017, 3751403
78. Gadelha, C. *et al.* (2015) Architecture of a host–parasite interface: complex targeting mechanisms revealed through proteomics. *Mol. Cell. Proteom.* 14, 1911–1926
79. Jackson, A.P. *et al.* (2012) Antigenic diversity is generated by distinct evolutionary mechanisms in African trypanosome species. *Proc. Natl. Acad. Sci. U. S. A.* 109, 3416–3421
80. Santos-de-Souza, R. *et al.* (2019) Serine proteinases in *Leishmania* (*Viannia*) *braziliensis* promastigotes have distinct subcellular distributions and expression. *Int. J. Mol. Sci.* 20, 1315
81. Khamsehnejad, M.I. *et al.* (2019) Identification, molecular characterization, and *in silico* structural analysis of carboxypeptidase B2 of *Anopheles stephensi*. *J. Med. Entomol.* 56, 72–85
82. Frasch, A.P. *et al.* (2018) Substrate specificity profiling of M32 metallocarboxypeptidases from *Trypanosoma cruzi* and *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* 219, 10–16

83. Kelly, F.D. *et al.* (2020) Touching the surface: diverse roles for the flagellar membrane in kinetoplastid parasites. *Microbiol. Mol. Biol. Rev.* 84, e00079-19
84. Peacock, L. *et al.* (2014) Meiosis and haploid gametes in the pathogen *Trypanosoma brucei*. *Curr. Biol.* 24, 181–186
85. Gluenz, E. *et al.* (2010) Beyond 9+0: noncanonical axoneme structures characterize sensory cilia from protists to humans. *FASEB J.* 24, 3117–3121
86. Sharma, R. *et al.* (2009) The heart of darkness: growth and form of *Trypanosoma brucei* in the tsetse fly. *Trends Parasitol.* 25, 517–524
87. Romeiro, A. *et al.* (2000) Ultrastructural and biochemical characterization of promastigote and cystic forms of *Leptomonas wallacei* n. sp. isolated from the intestine of its natural host *Oncopeltus fasciatus* (Hemiptera: Lygaeidae). *J. Euk. Microbiol.* 47, 208–220
88. Nawata, A. *et al.* (2015) Encystment and excystment of kinetoplastid *Azumiobodo hoyamushi*, causal agent of soft tunic syndrome in ascidian aquaculture. *Dis. Aquat. Org.* 115, 253–262
89. Prokopchuk, G. *et al.* (2019) Morphological, ultrastructural, motility and evolutionary characterization of two new Hemistasiidae species. *Protist* 170, 259–282
90. Tashyreva, D. *et al.* (2018) Life cycle, ultrastructure, and phylogeny of new diplonemids and their endosymbiotic bacteria. *mBio* 9, e02447-02417
91. Tashyreva, D. *et al.* (2018) Phylogeny and morphology of new diplonemids from Japan. *Protist* 169, 158–179
92. Vidal, J.C. *et al.* (2016) Loss of the cytostome-cytopharynx and endocytic ability are late events in *Trypanosoma cruzi* metacyclogenesis. *J. Struct. Biol.* 196, 319–328
93. An, T. *et al.* (2020) FAZ27 cooperates with FLAM3 and ClpGM6 to maintain cell morphology in *Trypanosoma brucei*. *J. Cell Sci.* 133, jcs245258
94. Sunter, J.D. *et al.* (2015) Modulation of flagellum attachment zone protein FLAM3 and regulation of the cell shape in *Trypanosoma brucei* life cycle transitions. *J. Cell Sci.* 128, 3117–3130
95. Hayes, P. *et al.* (2014) Modulation of a cytoskeletal calpain-like protein induces major transitions in trypanosome morphology. *J. Cell Biol.* 206, 377–384
96. Clayton, C. (2019) Regulation of gene expression in trypanosomatids: living with polycistronic transcription. *Open Biol.* 9, 190072
97. Kolev, N.G. *et al.* (2012) Developmental progression to infectivity in *Trypanosoma brucei* triggered by an RNA-binding protein. *Science* 338, 1352–1353
98. Dean, S. *et al.* (2019) Basalin is an evolutionarily unconstrained protein revealed via a conserved role in flagellum basal plate function. *eLife* 8, e42282
99. Kaur, B. *et al.* (2018) Transformation of *Diplonema papillatum*, the type species of the highly diverse and abundant marine microeukaryotes Diplonemida (Euglenozoa). *Environ. Microbiol.* 20, 1030–1040
100. Fukasawa, Y. *et al.* (2017) Origin and evolutionary alteration of the mitochondrial import system in eukaryotic lineages. *Mol. Biol. Evol.* 34, 1574–1586
101. Zíková, A. *et al.* (2017) A paradigm shift: the mitoproteomes of procyclic and bloodstream *Trypanosoma brucei* are comparably complex. *PLoS Pathog.* 13, e1006679
102. Santos, H.J. *et al.* (2018) Reinventing an organelle: the reduced mitochondrion in parasitic protists. *Trends Parasitol.* 34, 1038–1055
103. Voleman, L. and Doležal, P. (2019) Mitochondrial dynamics in parasitic protists. *PLoS Pathog.* 15, e1008008
104. Vetvicka, V. and Fernandez-Botran, R. (2018) β-glucan and parasites. *Helminthologia* 55, 177–184
105. Melida, H. *et al.* (2018) Non-branched β-1,3-glucan oligosaccharides trigger immune responses in *Arabidopsis*. *Plant J.* 93, 34–49
106. Myklestad, S.M. and Granum, E. (2009) Biology of (1,3)-β-glucans and related glucans in protozoans and chromistans. In *Chemistry, Biochemistry, and Biology of (1,3)-β-Glucans and Related Polysaccharides* (Bacic, A. *et al.*, eds), pp. 1–677, Academic Press
107. Bushkin, G.G. *et al.* (2012) β-1,3-glucan, which can be targeted by drugs, forms a trabecular scaffold in the oocyst walls of *Toxoplasma* and *Eimeria*. *mBio* 3, e00258-12
108. Yoshimi, A. *et al.* (2017) Function and biosynthesis of cell wall α-1,3-glucan in fungi. *J. Fungi* 3, 63
109. Opperdoes, F.R. *et al.* (1988) Demonstration of glycosomes (microbodies) in the Bodonid flagellate *Trypanoplasma borelli* (Protozoa, Kinetoplastida). *Mol. Biochem. Parasitol.* 30, 155–163
110. Gray, M.W. *et al.* (2020) The draft nuclear genome sequence and predicted mitochondrial proteome of *Andalucia godoyi*, a protist with the most gene-rich and bacteria-like mitochondrial genome. *BMC Biol.* 18, 22
111. Bouckaert, R. *et al.* (2019) BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* 15, e1006650
112. Poinar, G. and Poinar, R. (2004) *Paleoleishmania proterus* n. gen., n. sp., (Trypanosomatidae: Kinetoptastida) from Cretaceous Burmese amber. *Protist* 155, 305–310
113. Gray, J. and Boucot, A.J. (1989) Is *Moyeria* a euglenoid? *Lethaia* 22, 447–456