From Flegontov, P., & Lukeš, J., (2012). Mitochondrial Genomes of Photosynthetic Euglenids and Alveolates. In: Maréchal-Drouard, L., *Mitochondrial Genome Evolution* (pp. 127–153). Elsevier Ltd.: Academic Press.
ISBN: 9780123942791

Academic Press

CHAPTER SIX

Mitochondrial Genomes of Photosynthetic Euglenids and Alveolates

Pavel Flegontov and Julius Lukeš[1]
Biology Centre, Institute of Parasitology, Czech Academy of Sciences and Faculty of Science, University of South Bohemia, České Budějovice (Budweis), Czech Republic
[1]Corresponding author. E-mail: jula@paru.cas.cz

Contents

Abstract

Euglenida belong to the eukaryotic supergroup Excavata, the members of which possess the most varied mitochondrial genomes in terms of their structure and gene content. Heterotrophic protists represent the majority of Excavata, as only the Euglenida contain a green plastid, apparently acquired by secondary endosymbiosis. The sister group of Euglenida, the mostly parasitic Kinetoplastida, have an extremely complex mitochondrial DNA (kinetoplast DNA), which is usually composed of thousands of mutually interlocked DNA circles. Most mRNAs encoded by this genome are rendered translatable only after they undergo intricate editing via insertions and/or deletions of uridines. The mitochondrial DNA of the other sister group, Diplonemida, is unique as its transcripts must be massively *trans*-spliced before translation. None of these complex mechanisms has so far been found in the mitochondrial genome and transcriptome of *Euglena gracilis*, the best studied member of Euglenida. Its mitochondrial DNA exists in the form of numerous differently sized linear fragments. Their

Advances in Botanical Research, Volume 63
ISSN 0065-2296,
http://dx.doi.org/10.1016/B978-0-12-394279-1.00006-5

Mitochondrial Genome Evolution, First Edition, 2012, 127–153

significant fraction is non-coding and full of various repeats, which intersperse fragments of a handful of protein-coding genes. Mostly photosynthetic dinoflagellates and parasitic apicomplexans with a relic plastid constitute a large and diverse group within alveolates. All species of this group share the most reduced mitochondrial genome found, containing just three, and in some cases probably two, protein-coding genes along with highly fragmented rRNA genes, and no tRNA genes. Mitochondrial genomes of dinoflagellates and those of smaller groups within the apicomplexa–dinoflagellata assemblage, perkinsids and chromerids, in all cases have a recombining, highly scrambled sequence, and frequently demonstrate other non-canonical features in structure and expression: fused genes, extensive RNA editing, *trans*-splicing, 5′ oligoU caps, loss of start and stop codons, extensive translational frameshifting. Some of these oddities apparently appeared in several groups independently, probably due to relaxed selective constraints in tiny organellar genomes.

1. MITOCHONDRIAL GENOMES OF EUGLENIDS

1.1. Phylogeny of Euglenida

The Euglenida, a group of protists, have been intensely studied throughout most of the twentieth century. This interest was stimulated by their apparent ecological significance, as well as the ease with which they can be cultivated in a simple and cheap medium. In the era predating molecular biology, the species *Euglena gracilis* was the subject of numerous physiological studies. After the advent of molecular biology, it served as a model protist, because the genome of its green plastid was the second plastid genome to be completely sequenced (Hallick *et al.*, 1993). However, surprisingly little has been known about its mitochondrial and nuclear genomes.

It has been well established that the Euglenida belong to the superkingdom Excavata (Fig. 6.1), which arguably represents the earliest branch of the eukaryotic tree (Cavalier-Smith, 2010). Within this morphologically and genetically extremely diverse group of single-celled eukaryotes, Euglenida, along with their two sister groups Diplonemida and Kinetoplastida, constitute the phylum Euglenozoa. The following morphological features unite these free-living, commensalic and parasitic flagellates: (1) with very few exceptions of aflagellar stages, all cells carry at least a single flagellum equipped with a prominent structure called a paraflagellar rod; (2) a morphologically pronounced flagellar pocket; and (3) a single, usually reticulated, mitochondrion with tubular cristae (Adl *et al.*, 2005). Moreover, the most prominent common molecular features include polycistronic transcription, massive *trans*-splicing and, with very few exceptions, the absence of introns (Lukeš, Hashimi, & Zíková, 2005; Lukeš, Leander, &

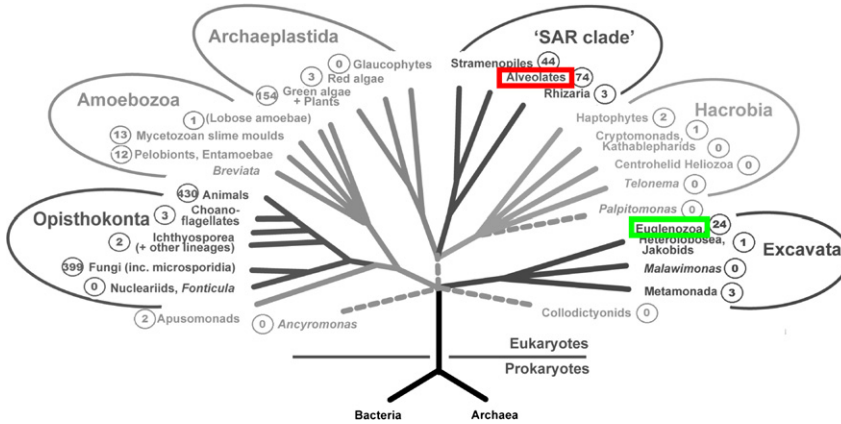

Figure 6.1 Current view of eukaryotic diversity; both protist groups dealt within this chapter are highlighted in colour. The tree is based on Roger and Simpson (2009). The numbers of whole-genome sequencing projects for a given group are shown in circles. See the colour plate.

Keeling, 2009). Although kinetoplastids and diplonemids have a single, often reticulated, mitochondrion in their cells, the situation is slightly more complex in euglenids. Although most species also carry one, usually large, mitochondrion, *Peranema* and likely some other species contain multiple small organelles (Hall, 2005; Roy, Faktorová, Lukeš, & Burger, 2007).

A unique and important feature of euglenids is their acquisition of a green plastid via secondary endosymbiosis (Archibald & Keeling, 2002). Despite earlier claims (Hannaert *et al.*, 2003), the kinetoplastids did not seem to have harboured a plastid in the course of their evolutionary history (Leander, 2004; Simpson, Stevens, & Lukeš J, 2006). The numerous fully sequenced genomes available for the kinetoplastid flagellates belonging to the genera *Trypanosoma* and *Leishmania* (El-Sayed *et al.*, 2005) strongly support this scenario, although these mostly parasitic protists seem to have acquired a handful of plastid-derived genes by horizontal gene transfer (Týč, Long, Jirků,& Lukeš, 2010). No complete genome is so far available for Diplonemida to convincingly address this important question for this least known group, although the most studied species, *Diplonema papillatum*, is currently the subject of a whole-genome initiative (Č. Vlček and G. Burger, personal communication). Current phylogenetic analyses based on numerous nuclear-encoded genes identify diplonemids as the most closely related group to euglenids (Simpson, Gill, Callahan, Litaker, & Roget, 2004), a relationship that is certainly not supported by the structure of their mitochondrial genomes (see below).

1.2. Mitochondrial Genomes of Euglenids

Although the mitochondrial DNA of euglenids is poorly known and does not seem to be very complex (Spencer & Gray, 2011), this certainly does not apply to its sister groups. The Kinetoplastida belong to organisms with the best studied mitochondrial DNA, also termed kinetoplast (k) DNA. The extremely complex kDNA of *Trypanosoma brucei* is composed of well-described maxicircles and minicircles interconnected into a single network, as well as of hundreds of proteins responsible for the maintenance and replication of this network (Lukeš *et al.*, 2005; Stuart, Schnaufer, Ernst, & Panigrahi, 2005). A similar situation holds for the mitochondrial proteome of this causative agent of African sleeping sickness, as up to 1000 proteins have been identified by numerous methods in its single mitochondrion (Panigrahi *et al.*, 2009).

The initial studies on *E. gracilis* on what was likely mitochondrial DNA were performed by Ray and Hanawalt (1965), and further characterization of this DNA occurred in the 1970s. Several authors have shown that the mitochondrial DNA of *E. gracilis* is represented by heterogeneously sized linear molecules (Nass, Schori, Ben-Shaul, & Edelman, 1974), which, however, may have high genomic complexity. Based on hybridization experiments, it was estimated that this complexity may reach up to 70 kb (Crouse, Vandrey, & Stutz, 1974; Talen, Sanders, & Flavell, 1974), at that time a very high level for an organellar genome. Although it was of general interest to learn more about this genome, especially for comparison with the intensely studied mitochondrial genomes of kinetoplastid parasites, a hiatus in any progress in this field lasted for more than two decades.

It was only in 1997 that a more detailed analysis of this long-ignored genome was initiated. First, electron microscopy and hybridization experiments confirmed both the presence of heterogeneous linear molecules, with the peak reaching size around 5 kb, and the overall complexity of this organellar genome of ~70 kb (Yasuhira & Simpson, 1997). The same authors were unable to resolve mitochondrial DNA molecules under the conditions of a pulse-field gel, even after γ-irradiation. This applied only to the mitochondrial DNA, as RNA from the mitochondrion was easily recovered, with four prominent bands likely representing ribosomal (r) RNA (Yasuhira & Simpson, 1997). Another approach to obtaining mitochondrial DNA and RNA from *E. gracilis* was recently adopted by Spencer and Gray (2011), who purified intact mitochondrial vesicles from lyzed cells.

The AT rich mitochondrial DNA from this preparation was comprised of differently sized linear molecules with peaks around 4.0 and 7.5 kb. Several laboratories thus confirmed the composition of *E. gracilis* mitochondrial DNA from variously sized linear molecules, yet these results remain to be reconciled with the fact that this mitochondrial DNA cannot enter the gel under pulse-field conditions. Unusual conformations have been implied, such as two- or three-dimensional networks of DNA molecules or unusually tight associations with proteins (Yasuhira & Simpson, 1997) or even branched molecules (Spencer & Gray, 2011), yet no data are so far available to support any of these claims.

The laboratories of Larry Simpson and Jean-Michel Grienenberger both established the first sequence information from the mitochondrial genome of *E. gracilis*. In both cases, part of the open reading frame (ORF) of cytochrome *c* oxidase subunit 1 (*cox1*) was obtained, revealing several unexpected features (Yasuhira & Simpson, 1997; Tessier, van der Speck, Gualberto, & Grienenberger, 1997). Since it was difficult to clone mitochondrial DNA fragments, *cox1* sequences were obtained using a 5′ and 3′ RACE (rapid amplification of cDNA ends) protocol. Invariably, only fragments of the *cox1* ORF were retrieved, mostly representing the 5′ or 3′ part of this highly conserved gene. A full-size gene was never obtained (Yasuhira & Simpson, 1997).

This lack of a full-size ORF initially implied the existence of a process similar to RNA editing of kinetoplastid flagellates, due to which parts of the mitochondrial genetic information are encrypted at the DNA level. Since the extensive insertions and/or deletions of uridines in the kDNA transcripts are specified by small abundant and heterogeneous molecules called guide RNAs (Blum, Bakalara, & Simpson, 1990), a search for homologous RNA species was also performed in the *E. gracilis* organelle. Guide RNAs can be tracked down relatively easily due to their ability to be capped *in vitro* by the activity of guanylyltransferase and radioactively labelled GTP (Blum *et al.*, 1990). However, a thorough search proved that these specialized uridine-tailed RNA molecules are not present either in *E. gracilis* (Yasuhira & Simpson, 1997) or in *D. papillatum* (D.A. Maslov, personal communication). Therefore, sufficient evidence is now available to conclude that the uridine insertion/deletion type of RNA editing, pervasive in the sister group of euglenids, the Kinetoplastida, is apparently absent from the *E. gracilis* mitochondrion. This lack of editing in euglenids was used to support a scenario whereby this enigmatic and complex process

emerged relatively late in evolution. Furthermore, trypanosomes and leishmanias are known to use the non-canonical TGA triplet to encode tryptophan in their mitochondria, which is a substantial departure from the universal genetic code, where this triplet is used as one of three stop codons. Throughout the *E. gracilis cox1* gene, only the TGG triplet is invariably used to specify tryptophan (Yasuhira & Simpson, 1997; Tessier *et al.*, 1997). More recent deposition into GenBank of two protein-coding gene sequences from the *E. gracilis* mitochondrial genome, namely cytochrome *c* oxidase subunit 2 (cox2) and NADH dehydrogenase subunit 6 (nad6) (unpublished data), further supports the notion that RNA editing mechanistically similar to the process known from kinetoplastids is most likely lacking in the euglenids.

In combination with the apparent absence of guide RNAs, these data indicate that the structure and transcription of the mitochondrial DNA of euglenids may significantly differ from that of the kinetoplastids. However, in two euglenid species, *Petalomonas cantuscigni* and *P. mediocanellata*, an electron-dense mitochondrial inclusion body was observed that was reminiscent of the kDNA disk of kinetoplastids (Leander, Triemer, & Farmer MA, 2001), although a similar structure was not observed in the mitochondria of other euglenids, such as *Peranema trichophorum* and *Entosiphon sulcatum* (Roy *et al.*, 2007). Moreover, using 4′,6-diamidino-2-phenylindole (DAPI) stain, Roy *et al.* (2007) showed that the inclusion body in *P. cantuscygni* does not contain DNA, so its proposed homology with the kDNA disk (Leander *et al.*, 2001) is unlikely. The abundant mitochondrial DNA of *P. cantuscygni* is distributed throughout the organelle in a network-like pattern, reminiscent of the related diplonemids (see below). The other two euglenids investigated, *Peranema* and *Entosiphon*, contain a significantly smaller amount of DNA in their mitochondria, which seems to be scattered in the form of multiple fluorescent spheres or agglomerates (Roy *et al.*, 2007), similar to structures reported earlier in *E. gracilis* (Hayashi and Ueda, 1989; Hayashi-Isimaru, Ueda, & Nonaka, 1993).

Using buoyant density CsCl–bisbenzimide gradient centrifugation, attempts were made to separate the mitochondrial and nuclear DNA of these three euglenids. Inspection of the AT rich fraction of *P. cantuscygni* DNA by electron microscopy revealed the presence of linear molecules, as well as small and large circles. It was proposed that the various molecules observed could be produced by replication via the rolling circle mechanism. Sequencing of the mitochondrial DNA-enriched fraction had so far detected only a fragment of subunit 6 of the ATPase gene (Roy *et al.*, 2007).

Preliminary small-scale sequence analyses of mitochondrial fractions of *Peranema* and *Entosiphon* failed to identify any protein-coding regions. However, this comparative analysis showed that multi-chromosome mitochondrial genomes are likely widespread in the euglenid flagellates (Roy *et al.*, 2007).

Sequencing of several linear mitochondrial chromosomes of *E. gracilis* revealed a common interesting feature. Regardless of whether the sequenced molecule contained fragments of the small subunit mitoribosomal RNAs (SSU rRNA), *cox1*, *cox2*, or *cox3* coding regions, they were flanked on both the 5′ and 3′ ends by highly conserved repeats, which have been proposed to play a role in replication, recombination and/or transcription (Spencer & Gray, 2011). Small subunit (SSU) and large subunit (LSU) rRNA genes are bipartide, and all four fragments can be capped by the action of guanylyltransferase. Therefore, they are very likely independently transcribed. This situation has been used as an argument to support a scenario of rRNA evolution, according to which ancestral ribosomes rRNA fragments were held together by inter- and intramolecular interactions (Boer & Gray, 1988).

Gene fragments constituting the mitochondrial genome of *E. gracilis* along with respective full-size copies led Spencer and Gray (2011) to postulate a model to explain the emergence of guide RNAs, which provide information for editing of the mitochondrial transcripts in related kinetoplastid flagellates. Initially, by illegitimate recombination via the flanking conserved repeats, a gene fragment becomes located on a small circle. If this circular DNA happens to contain an origin of replication and a promoter, which is opposite to the gene fragment, a small antisense RNA complementary to the fragment of the parental gene will be produced. It is such small RNAs that, in collaboration with several intricate protein complexes, execute the exact insertions/deletions of uridine residues into mitochondrial mRNAs of all kinetoplastids studied so far (Lukeš *et al.*, 2005). Despite its speculative nature, this model explains for the first time the emergence of the unique guide RNA-directed editing machinery and also the complex kDNA network, earlier dubbed an evolutionary improbable structure (Lukeš *et al.*, 2002).

1.3. Mitochondrial Genomes of Diplonemids

The uniquely complex nature of the mitochondrial genomes of euglenids and kinetoplastids, supporting the anything goes postulation for

mitochondrial genomes (Burger, Gray, & Lang, 2003; Gray, Lang, & Burger G, 2004), is surpassed by what has been uncovered so far in the mitochondrion of the third group: the non-photosynthetic diplonemids. In the mitochondrial genome of *D. papillatum*, a model species representing this least common, usually commensalic group, genes are invariably fragmented. Each gene fragment, called a module, is individually located on a circular chromosome belonging to one of two types labelled A and B (Marande, Lukeš, & Burger G, 2005). Thousands of these minicircles of conserved structure are freely dispersed throughout the mitochondrial lumen (Marande *et al.*, 2005). Separate non-overlapping precursor RNAs are assembled into a mature transcript via extensive *trans*-splicing (Marande & Burger, 2007), the mechanism of which remains unknown. However, it is obvious that the mechanism must be highly complex, because, for example, in the case of *cox1*, it is able to *trans*-splice together, in an orderly manner, nine separately transcribed fragments. To complicate matters even further, at least in one case, six uridines are inserted between two modules, implying that splicing and editing machineries may exist next to each other (Marande & Burger, 2007). This en-block insertion of uridines is highly conserved among diplonemids; it was recently encountered in three other members of this group (*D. ambulator*, *Diplonema* sp. and *Rhynchopus euleeides*) (Kiethega, Turcotte, & Burger, 2011). So far, ten protein-coding genes were predicted to be assembled from three to 12 modules ranging in size from 60 to 350 bp. However, with the exception of a 3′ module of the LSU rRNA, the remaining LSU and all SSU rRNA fragments remain elusive. These combined features qualify the mitochondrial genome of *D. papillatum* as truly the most bizarre (Vlček *et al.*, 2011).

1.4. Conclusion

Comparative studies of the mitochondrial genomes of various euglenozoans led to the prediction that in the common evolutionary ancestor of kinetoplastids, diplonemids and euglenids, rampant genome fragmentation produced via the neutral evolutionary pathway dramatically different, unique and highly complex organellar genomes and transcriptomes (Flegontov, Gray, Burger, & Lukeš, 2011). Furthermore, the euglenozoan mitochondrial genomes share several unique features with homologous genomes of dinoflagellates, a totally unrelated group of mostly photosynthetic protists (Keeling *et al.*, 2005). These shared characteristics further

reinforce the theory of cascades of convergent evolution between both ecologically important lineages (Lukeš *et al.*, 2009).

2. MITOCHONDRIAL GENOMES OF PHOTOSYNTHETIC ALVEOLATES

2.1. Phylogeny of Alveolata

The Alveolata includes three well-established and well-known groups: predatory Ciliata, parasitic Apicomplexa with a relic non-photosynthetic plastid in most species, and their sister group Dinoflagellata (Leander & Keeling, 2004), which are either photosynthetic, predatory or parasitic (Hackett, Anderson, Erdner, Bhattacharya, 2004). It is hypothesized that all alveolates (Fig. 6.1), heterokonts, and related groups united in the taxon. Chromalveolata (Adl *et al.*, 2005) got their secondary plastids in a single endosymbiotic event from a red alga (Archibald, 2009; Janouškovec, Horák, Oborník, Lukeš, & Keeling, 2010; Keeling, 2009), with subsequent partial or complete plastid losses in some lineages.

Recently, several new groups have been recognized within the Apicomplexa–Dinoflagellata assemblage. A non-parasitic predatory group, Colpodellida (Berney, Fahrni, & Pawlowski, 2004; Brugerolle, 2002; Kuvardina *et al.*, 2002; Leander, Kuvardina, Aleshin, Mylnikov, & Keeling, 2003; Mylnikov, 2009) was placed within Apicomplexa, previously an exclusively parasitic group. Colpodellids have an apical complex, an eponymous diagnostic feature of Apicomplexa, a complex of organelles used for cell invasion or predation. A typical apical complex is composed of rhoptries and micronemes (extrusive organelles) enveloped by a microtubule-formed conoid (Leander and Keeling, 2003).

Another new twig on the apicomplexan stem is Chromerida, lacking the complete apical complex (Oborník *et al.*, 2011) and containing a fully functional secondary plastid (Janouškovec, Horák, Oborník, Lukeš, Keeling *et al.*, 2010). Thus, this group is the closest photosynthetic relative of Apicomplexa (Keeling, 2008; Moore *et al.*, 2008; Okamoto & McFadden, 2008). The first species of Chromerida described was an alga, *Chromera velia*, which lives in association with corals, but very likely also has a free-living stage (Moore *et al.*, 2008; Oborník *et al.*, 2011; Weatherby, Murray, Carter, & Šlapeta, 2011). The second species isolated as CCMP3155 and described recently as *Vitrella brassicaformis* (Oborník *et al.*, 2012) forms a distinct lineage unexpectedly distant from *C. velia* (Janouškovec *et al.*, 2010; Oborník *et al.*,

2012). Chromerida and Colpodellida may be more closely related to each other than they are to the crown apicomplexans (Moore *et al.*, 2008).

The dinoflagellate branch has come under study more recently for several reasons, the main one being the paramount ecological significance of these protists. The tentative branching order within this species-rich group is as follows (Bachvaroff, Handy, Place, & Delwiche, 2011; Gómez, López-García, Nowaczyk, & Moreira, 2009, Gómez, Moreira, & López-García, 2010; Hoppenrath & Leander, 2010; Saldarriaga, Taylor, Cavalier-Smith, Menden-Deuer, Keeling, 2004; Skovgaard, Meneses, & Angélico, 2009): (1) Perkinsozoa (Perkinsidae, Perkinsea), intracellular parasites of bivalve molluscs (*Perkinsus*) and protists (*Cryptophagus*, *Parvilucifera*, *Rastrimonas*), probably have a relic plastid (Fernandes Robledo *et al.*, 2011), share spliced leader RNAs with crown dinoflagellates (Joseph *et al.*, 2010; Zhang, Campbell, Sturm, Dungan, & Lin *et al.*, 2011) and an apical complex with crown apicomplexans and colpodellids (Leander and Keeling, 2003); (2) Ellobiopsida (*Ellobiopsis*, *Thalassomyces*, *Ellobiocystis*, *Parallobiopsis*), a group with uncertain position, mostly ectoparasites of crustaceans (Gómez et al., 2009); (3) Marine Alveolate Group I, intracellular parasites of fish eggs (*Ichthyodinium*) and marine protists (*Duboscquella*) (Groisillier, Massana, Valentin, Vaulot, & Guillou *et al.*, 2006; Harada, Ohtsuka, & Horiguchi, 2007; Skovgaard *et al.*, 2009); (4) Syndiniales or Marine Alveolate Group II (*Amoebophrya*, *Hematodinium*, *Syndinium*), endoparasitic dinoflagellates without plastids, found mainly in crustaceans and protists, such as other dinoflagellates (Groisillier *et al.*, 2006; Guillou *et al.*, 2008; Skovgaard, Massana, Balagué, & Saiz, 2005; Stentiford & Shields, 2005); (5) *Oxyrrhis*, a group with uncertain position relative to Syndiniales, a predatory dinoflagellate with some evidence of a relic plastid (Bachvaroff *et al.*, 2011; Jackson, Gornik, & Waller, 2011; Saldarriaga *et al.*, 2004; Slamovits & Keeling, 2011); (6) Noctilucales, an early-branching group of photosynthetic dinoflagellates (Gómez *et al.*, 2010); (7) Dinophyceae (Dinokaryota), core dinoflagellates (Hoppenrath & Leander, 2010; Saldarriaga *et al.*, 2004) with secondary or tertiary plastids (Oborník, Janouškovec, Chrudimský, & Lukeš, 2009). Groups (3)–(7) have been united in the taxon Dinoflagellata (Saldarriaga *et al.*, 2004; Skovgaard *et al.*, 2005).

With the exception of a medically important apicomplexan parasite, *Cryptosporidium*, which has mitosomes (Keithly, Langreth, Buttle, & Mannella, 2005), all groups of the Apicomplexa–Dinoflagellata branch investigated so far contain a conventional mitochondrion. Mitochondrial genomes were studied mainly in core dinoflagellates and core

apicomplexans. Recent studies focused also on *Hematodinium* (Jackson *et al.*, 2011), *Perkinsus* (Masuda, Matsuzaki, & Kita, 2010; Zhang *et al.*, 2011), *Chromera* (Flegontov *et al.*, unpublished data) and *Vitrella* (Janouškovec *et al.*, unpublished data). The published results on dinoflagellates and *Perkinsus* and our unpublished findings on chromerids are discussed in this section.

2.2. Mitochondrial Genomes in Ciliates and Parasitic Apicomplexans

Ciliata, the basal group of Alveolata, have linear-mapping mitochondrial genomes with a normal gene number: two rRNAs, seven tRNAs, 21 protein-coding genes of known function, and 22 ciliate-specific ORFs (Brunk, Lee, Tran, Li, 2003; Burger *et al.*, 2000). Both large subunit (LSU) and small subunit (SSU) rRNAs are split into two separately encoded fragments; many tRNA genes are lost from the mitochondrial genome, with the corresponding tRNAs imported from the cytosol (Rusconi & Cech, 1996). Alternative start codons AU(U/A), (G/U)UG are used in at least 7 of 43 protein-coding genes (Burger *et al.*, 2000; Edqvist, Burger, Gray, 2000). UGA encodes tryptophan, UAG is unassigned, and so the only stop codon remaining in use is UAA (Burger *et al.*, 2000).

Apicomplexa diverged much further from the canonical mitochondrial genome structure. The genome is reduced to just three protein-coding genes, cytochrome oxidase subunits 1 (*cox1*) and 3 (*cox3*), and apocytochrome *b* (*cob*), arranged on circularly permutated linear molecules (Feagin, 1992). *Cox2*, universally present in other mitochondrial genomes, is transferred to the nucleus (Waller & Keeling, 2006). Apicomplexans apparently lack genes coding for subunits of complex I (NADH dehydrogenase) of the respiratory chain (Gardner, Hall, Fung, White, & Berriman, 2002). Ribosomal RNAs are highly fragmented: 23 LSU and SSU fragments were found in *Plasmodium*, but some functionally important parts of the rRNAs are missing in this set of fragments (Feagin, Mericle, Werner, & Morris, 1997; Kairo, Fairlamb, Gobright, & Nene, 1994). Such absences might be accounted for by targeting of small rRNA fragments from the cytoplasm.

A precedent for such an rRNA import exists, as mitochondrial import of cytosolic 5S rRNA has been demonstrated in mammals (Entelis, Kolesnikova, Dogan, Martin, & Tarassov, 2001). Moreover, all tRNAs are imported into the apicomplexan mitochondria from the cytosol (Esseiva, Naguleswaran, Hemphill, & Schneider, 2004). It has been suggested that

tRNA-fMet is imported even from the plastid (Barbrook, Howe, & Purton, 2006a; Howe & Purton, 2007), but evidence for this type of import is still lacking. In *cox1* and *cox3* genes, AUA or AUU are used as start codons, AUG is used in *cob* (Feagin, 1992; Kairo *et al.*, 1994; Rehkopf, Gillespie, Harrell, & Feagin, 2000). Stop codon usage is identical to ciliates; the only codon used is UAA (Rehkopf *et al.*, 2000). All transcripts, including rRNA fragments, are oligoadenylated (Gillespie, Salazar, Rehkopf, & Feagin, 1999; Rehkopf *et al.*, 2000).

2.3. Mitochondrial Genomes in Dinokaryota

Mitochondrial genomes of core dinoflagellates, Dinokaryota, have the same extremely reduced coding capacity of three protein-coding genes and fragmented rRNAs, but have accumulated numerous oddities in genome architecture and expression (Nash, Nisbet, Barbrook, & Howe, 2008; Waller & Jackson, 2009). Therefore, it is safe to assume that dinoflagellate mitochondrial DNA evolved from an apicomplexan-like state in the last common ancestor (LCA) of Apicomplexa and Dinokaryota. Expansion of actively recombining non-coding DNA increased genome size greatly, created truncated pseudogenes and tiny gene fragments, and put remaining full-length genes in dozens of sequence contexts. Recombination-driven rearrangements are common in the mtDNA of land plants as well (Knoop, 2004), however, never to the extent observed in dinoflagellates.

Mitochondrial DNA of dinoflagellates is revealed by Southern blot analysis as a pool of heterogeneous molecules, 6–10 kb and longer (Chaput, Wang, & Morse, 2002; Jackson *et al.*, 2007; Nash *et al.*, 2007; Norman & Gray, 2001). Preliminary pulse-field gel electrophoresis experiments suggest that an upper chromosome size limit for *Amphidinium carterae* is ~30 kb (Nash *et al.*, 2008). In the same species, non-coding DNA content is estimated at 85% (Nash *et al.*, 2007), and it also seems high in other species. Non-coding regions are rich in inverted repeats capable of forming stem-loop structures: ~50–150 bp repeats ~10–30 bp apart in *A. carterae* (Nash *et al.*, 2007), >6 bp repeats no more than 5 bp apart in *Karlodinium micrum* (Jackson *et al.*, 2007) or >9 bp repeats no more than 50 bp apart in *Crypthecodinium cohnii* and *K. micrum* (Jackson *et al.*, 2007; Norman & Gray, 2001). No sequence conservation in the inverted repeats is apparent between species apart from a higher than average GC content. In other organisms, stem-loop structures are thought to play a role in the control of mitochondrial replication (Arunkumar & Nagaraju, 2006), transcript

stability (Kuhn, Tengler, & Binder S, 2001) and genome recombination (Bartoszewski, Katzir, & Havey, 2004), but their role in dinoflagellates remains unknown.

Limited shotgun cloning, polymerase chain reaction (PCR) surveys, and Southern blot analyses in several species (Chaput *et al.*, 2002; Imanian & Keeling, 2007; Jackson *et al.*, 2007; Kamikawa, Nishimura, & Sako, 2009; Nash *et al.*, 2007; Norman & Gray, 2001) all point in one direction: the genome structure in Dinokaryota is chaotic, with full-length genes, truncated genes, small gene fragments, and non-coding DNA mixed in numerous arrangements. Sequence divergence in gene fragments is always negligible, suggesting active ongoing recombination. It appears that gene arrangements may not be completely random; deeper genome sequencing is needed to address this issue. In *A. carterae*, *cox3* and *cob* are usually arranged head-to-head with variable spacers. *Cob* and *cox1* were amplified tail-to-tail with only one spacer. No single DNA molecule containing all three genes was shown by either restriction digestion followed by Southern blot analysis or by PCR (Nash *et al.*, 2007). In *C. cohnii* and *K. micrum*, a pool of small gene fragments contained *cox1* and *cox3* sequences, but *cob* sequences were lacking (Jackson *et al.*, 2007). In *K. micrum*, the following arrangements have been encountered in PCR amplicons: *cox1-cob*, *cox1-cox3*, *cob-cob*, and *cob-cox3* (Jackson *et al.*, 2007). In *Alexandrium catenella cob-cox1*, *cox1-cox1* intergenic spacers of random structure were sequenced. Moreover, some *cox1-cob* spacers contained *cox3* copies considered pseudogenes due to the lack of a conserved region at the 3′ end (Kamikawa *et al.*, 2009).

Despite the abundance of truncated gene copies, in some studies only full-length transcripts were revealed in expressed sequence tag (EST) datasets (Nash *et al.*, 2007), with RACE (Kamikawa *et al.*, 2009) and Northern blot analysis (Norman & Gray, 2001). On the other hand, apparently non-functional transcripts of pseudogenes or gene fragments were found in other studies: transcripts of *cox1* with insertions, some of them containing *cob* fragments (Imanian & Keeling, 2007); polycistronic transcripts with protein-coding and rRNA gene fragments (Jackson *et al.*, 2007); a long transcript matching apparently non-coding DNA (Jackson *et al.*, 2007); *cox3* transcript truncated at the 3′ end, with fragments of *cox1* and *cob* (Chaput *et al.*, 2002). In the latter species, *Lingulodinium polyedrum* (previous name *Gonyaulax polyedra*), *cob* and *cox3* probes hybridized to a smear of transcripts (Chaput *et al.*, 2002). Orderly transcription with defined promoters upstream of genes is difficult to imagine in such a disordered genomic system where a gene can be flanked by dozens of totally different sequences. Promoters located

within genes would lead to transcription of many truncated gene copies. In our view, the most reasonable assumption based on the data available is that (almost) all mitochondrial DNA in dinoflagellates is transcribed, but quickly degraded and therefore not visible in some experimental setups. Mature transcripts are most likely generated by cleavage at both ends and reach detectable concentrations.

Similar to apicomplexans, all transcripts in the dinoflagellate mitochondria, including rRNA fragments, are oligoadenylated (Chaput *et al.*, 2002; Jackson *et al.*, 2007; Kamikawa, Inagaki, & Sako, 2007, 2009; Nash *et al.*, 2007). Transcripts of *cox3* require *trans*-splicing in all Dinokaryota that have been investigated (Jackson *et al.*, 2007; Nash *et al.*, 2007; Waller & Jackson, 2009). In *K. micrum*, mature *cox3* transcripts are 839 nt in length, but some cDNAs are oligoadenylated at 712 nt. The genome contains a single 712-nt-long ORF immediately followed by a stop codon and *cox3* fragments including positions 718–839. Transcripts of the long and the short fragments are apparently *trans*-spliced taking six As from the oligoA tail of the long fragment (Jackson *et al.*, 2007). The *cox3* transcript in a basally branching species *A. carterae* lacks these oligoA-derived nucleotides (Nash *et al.*, 2007). The splicing mechanism remains totally unknown, and no intron-like sequences have been identified so far (Jackson *et al.*, 2007; Waller & Jackson, 2009).

Another striking feature of the dinoflagellate mitochondrial genetic system is extensive RNA editing. All three protein-coding genes are edited (and some rRNA fragments; see below), but editing was investigated mostly in *cob* and *cox1*. Predominant changes are A-G (~50%); U-C and C-U changes are also common. However, almost all possible changes were detected: G-A, U-G, G-U, G-C, C-G, A-U, U-A, A-C (Gray, 2003; Jackson *et al.*, 2007; Lin, Zhang, Spencer, Norman, & Gray, 2002; Lin, Zhang, & Gray, 2008; Zhang & Lin, 2005; Zhang & Lin, 2008; Zhang, Bhattacharya, Maranda, & Lin, 2008). Moreover, G-C substitutions seem to be unique to dinoflagellates. Such versatility of an editing system is totally unprecedented (Gray, 2003). Editing occurs mostly at the first and second codon positions usually affecting 2–3% of the nucleotide sequence and up to 6% in *cox3* of *K. micrum* (Jackson *et al.*, 2007). Ile-Val and Phe-Leu amino acid changes are most common as a result of editing (Waller & Jackson, 2009). It is particularly intriguing that editing is not progressing in the 5′ to 3′ direction or vice versa (Nash *et al.*, 2007) and that the editing sites are distributed in clusters (Waller & Jackson, 2009). Only full-length mature transcripts are edited in *K. micrum* (Jackson *et al.*, 2007). Sometimes, editing

eliminates in-frame UAG stop codons, unassigned in alveolates, such as two stop codons in *cox1* of *A. carterae* or one stop codon in *cox3* of *K. micrum* (Jackson *et al.*, 2007; Lin *et al.*, 2002; Nash *et al.*, 2007; Zhang & Lin, 2005). New editing sites are constantly evolving, but some sites are apparently highly conserved (Jackson *et al.*, 2007). The mechanism of RNA editing in dinoflagellates and its possible significance remain completely unknown. Apparently, (de)amination should be involved in relatively frequent transitions and base excision replacement in more rare transversions. Editing increases the GC content, and this may be important for the use of nucleus-encoded tRNAs imported into the mitochondrion (Waller & Jackson, 2009). Few sites in *A. carterae* gene fragments match post-edited transcripts suggesting that a guide RNA-like mechanism might be involved (Nash *et al.*, 2007), yet no such sites were found in *K. micrum* (Jackson *et al.*, 2007).

Extensive editing has been reported in ~25 species of Dinokaryota (Lin *et al.*, 2008; Zhang *et al.*, 2008): *Adenoides eludens* (Zhang *et al.*, 2007), *Akashiwo sanguinea* (Zhang *et al.*, 2007; Zhang et al., 2008), *Alexandrium tamarense*, *A. affine*. *A. catenella* (Kamikawa *et al.*, 2009; Zhang and Lin, 2005; Zhang *et al.*, 2005, 2007, 2008), *Ceratium longipes*, *Ceratocorys horrida* (Zhang *et al.*, 2007), *Dinophysis acuminata* (Zhang *et al.*, 2008), *Durinskia baltika* (Imanian & Keeling, 2007), *Gambierdiscus toxicus*, *Gonyaulax cochlea*, *Gymnodinium simplex* (Zhang *et al.*, 2007), *Karenia brevis* (Zhang *et al.*, 2005, 2007, 2008), *Karlodinium micrum* (Jackson *et al.*, 2007; Zhang & Lin, 2005; Zhang *et al.*, 2005, 2007, 2008), *Pfiesteria shumwayae*, *P. piscidida* (Zhang & Lin, 2002; Zhang & Lin, 2005; Zhang *et al.*, 2005, 2007, 2008), *Prorocentrum minimum*, *P. micans* and other *Prorocentrum* spp. (Lin, Zhang, & Jiao, 2006; Zhang & Lin, 2005; Zhang *et al.*, 2005, 2007, 2008), *Protoceratium reticulatum* (Zhang *et al.*, 2007, 2008), *Pyrocystis lunula*, *P. noctiluca* (Zhang *et al.*, 2007), *Pyrodinium bahamense* (Zhang *et al.*, 2005), *Scrippsiella* sp. and *S. sweeneyae* (Zhang *et al.*, 2005, 2007, 2008). Editing seems to be missing in *Noctiluca scintillans* (Zhang & Lin, 2008; Zhang *et al.*, 2007), a member of Noctilucales, and in basally branching species of Dinokaryota, *Heterocapsa triquetra* and *H. rotundata* (Zhang & Lin, 2008; Zhang *et al.*, 2005, 2007, 2008). Moreover, editing is not very extensive in some other basally branching species: *A. carterae*, *A. operculatum* (Nash *et al.*, 2007; Zhang & Lin, 2008; Zhang et al. 2007, 2008), *C. cohnii*, *Symbiodinium* sp. and *S. microadriaticum* (Zhang & Lin, 2005; Zhang *et al.*, 2005, 2007, 2008). These observations suggest that RNA editing appeared and spread within Dinokaryota.

The general consensus is that, in Dinokaryota AUG, start codons are missing in the conserved 5′ regions of transcripts, and it is difficult to

determine which codons act as alternative starts (Nash *et al.*, 2008; Waller & Jackson, 2009). Potential AUG occurs in *cox1* of *C. cohnii*, but these positions are not conserved (Jackson *et al.*, 2007). In *A. catenella*, isoleucine AU(A/U/C) or leucine (U/C)UG codons were found near the start of *cob* and *cox1* transcripts (Kamikawa *et al.*, 2009), and AUG was found near the start of the *cob* transcript. This AUG codon is also conserved in five other *Alexandrium* species and *Gonyaulax* sp. (Kamikawa *et al.*, 2008), and another potential AUG occurs in the *cob* of *K. micrum*, but sequence conservation starts upstream of this codon (Jackson *et al.*, 2007). In summary, it is still not clear whether AUG is used in some cases, and which alternative start codons are utilized.

Reliance on stop codons is also relaxed in Dinokaryota: oligoadenylation in *cox1* and *cob* occurs before UAG and UAA codons in *A. carterae* (Nash *et al.*, 2007), *A. catenella* (Kamikawa *et al.*, 2009), *Pfiesteria piscicida*, *Prorocentrum minimum*, *L. polyedrum* and *Karenia brevis* (Jackson *et al.*, 2007). In *cox3*, a UAA codon is generated by the oligoadenylation process itself in *A. carterae* (Nash *et al.*, 2007) and *K. micrum* (Jackson *et al.*, 2007). Creation of termination codons in such a manner is not unprecedented, as similar oligoadenylation is also known to occur in human mitochondria (Chrzanowska-Lightowlers, Temperley, Smith, Seneca, & Lightowlers *et al.*, 2004), and stop codons are occasionally, but not systematically, absent in plant mitochondria (Raczynska *et al.*, 2006). The mechanism of stop-codon-free termination in dinoflagellates remains to be established. To explain the available data, reliance on tmRNA-like molecules or special ribosome release factors has been implied (Nash *et al.*, 2008; Waller & Jackson, 2009). Furthermore, it was also proposed that oligoA translation producing lysine stretches can be tolerated in *cox1* and *cob*, because positively charged amino acids frequently occur at the C-termini of these proteins in other eukaryotes; however, *cox3* lacks a positively charged C-terminus and requires a defined terminator (Waller & Jackson, 2009).

The ribosomal RNA fragmentation pattern in Dinokaryota remains very similar to that of apicomplexans (Nash *et al.*, 2008; Waller & Jackson, 2009). However, few rRNA fragments have been identified in dinoflagellates probably due to the limited depth of genome/transcriptome sequencing: LSUA, LSUD, LSUE, LSUF, LSUG, and LSU RNA2 were found along with their fragmentary copies in *A. catenella* (Kamikawa *et al.*, 2007; Kamikawa et al., 2009); LSUE, LSUG in *C. cohnii* (Jackson *et al.*, 2007); LSUE, LSUF, LSUG in *Heterocapsa triquetra* (Jackson *et al.*, 2007); LSUA, LSUE, LSUG, LSU RNA2, LSU RNA10, SSU RNA8, and unassigned

RNA7 were found along with some unprocessed precursors or alternative variants in *K. micrum* (Jackson *et al.*, 2007). RNA editing of the same type as in protein-coding transcripts was demonstrated for LSUE in *A. catenella* (Kamikawa *et al.*, 2007) and for LSUA, LSUG in *K. micrum* (Jackson *et al.*, 2007). Surprisingly, rRNA genes were not found in 33 kb of shotgun clones and in PCR products of *A. carterae* mitochondrial DNA (Nash *et al.*, 2007). Import of some, but not all, rRNA fragments from the cytosol was proposed for apicomplexans, but remains purely hypothetical (see above). Import of tRNAs from the plastid into the mitochondrion was suggested for *Plasmodium* (Barbrook *et al.*, 2006a), yet again, experimental data to support this claim is lacking. In principle, the same might be true in dinoflagellates because tRNA-fMet is one of only a handful of tRNAs encoded in the dinoflagellate chloroplast genome (Barbrook, Santucci, Plenderleith, Hiller, & Howe, 2006b; Nelson *et al.*, 2007).

The predicted amino acid sequences of dinoflagellate *cox1*, *cox3* and *cob* show substitutions at several functionally important sites that are conserved in most other organisms (Nash *et al.*, 2008). Despite this sequence divergence and all the molecular oddities described for the dinoflagellate mitochondria, respiratory complexes III and IV activity was detected in *A. catenella* (Kamikawa *et al.*, 2009) suggesting that even such a messy genetic system can produce a properly functioning conventional respiratory chain.

2.4. Mitochondrial Genomes in Other Dinoflagellate Groups, Perkinsids, and Chromerids

Some results obtained from the dinoflagellates branching off prior to Dinokaryota, and also from perkinsids and chromerids, are presented. Mitochondrial DNA of *Hematodinium* sp. shares many features with core dinoflagellates (Jackson *et al.*, 2011), although it contains much less inverted repeats, since tightly packed fragments of different genes were found on genomic amplicons. A transcriptome assembly of 454 reads suggests that gene fragments and non-coding DNA are transcribed, but only full-length mature transcripts were detected with Northern blot analysis. This finding is in line with less extensive results available for Dinokaryota (see above), suggesting that everything is transcribed in these genomes (which can be detected by RNAseq) but probably quickly degraded, and hence not detectable in Northern blots. Unlike all Dinokaryota, the *cox3* transcript is not *trans*-spliced.

Typical dinoflagellate-type RNA editing was demonstrated for all three *Hematodinium* genes: A-G, U-C and C-U changes predominate, but A-U,

G-A and C-G also occur. Apparently non-functional fragmentary transcripts are also edited, unlike in *K. micrum* (Jackson *et al.*, 2007). Editing sites are not conserved between *Hematodinium* and Dinokaryota which, together with the absence of editing in *Heterocapsa* and *Noctiluca*, suggests that RNA editing arose independently in Syndiniales and Dinokaryota. The AUG start codon is not used in *Hematodinium* and the AUU triplet apparently does not take over its function. In *cob* and *cox1*, oligoadenylation occurs prior to the UAA stops. Cases of premature oligoadenylation have also been detected. Unlike in Dinokaryota, in this protest, the UAA stop in *cox3* is encoded. The set of rRNA fragments in *Hematodinium* is the largest found to date, especially regarding the SSU fragments: LSUA, LSUD, LSUE, LSUF, LSUG, LSU RNA2, LSU RNA10; SSUA, SSUB, SSUD, SSUF, SSU RNA8; unassigned RNA6 and RNA7. However, this exceptionally high number of fragments may be the consequence of deep transcriptome sequencing (Jackson *et al.*, 2011). Although variable transcripts were observed in 454 reads for SSUA and SSUB, only single bands were detected by Northern blot analysis. The apicomplexan pattern of fragmentation is conserved in *Hematodinium* as well. Although ongoing opportunity for fragmentation exists due to active recombination, further rRNA disassembly would presumably affect its ability to self-associate (Jackson *et al.*, 2011).

The predatory dinoflagellate *Oxyrrhis marina* with an uncertain phylogenetic position has been extensively studied by sequencing its mitochondrial DNA and EST library (Slamovits & Keeling, 2011; Slamovits, Saldarriaga, Larocque, Keeling, 2007). Its mitochondrial DNA encodes the *cox1*, *cox3* and *cob* genes but, remarkably, *cob* and *cox3* were always detected fused in one ORF. The authors showed that the *cob–cox3* fusion transcript is not processed, but fusion at the protein level was not investigated. The situation may be similar to that described in the distantly related *Acanthamoeba castellanii*, namely that a *cox1–cox2* fusion transcript does not necessarily result in a fusion protein (Lonergan & Gray, 1996). Moreover, *cox3* is very divergent in *O. marina*. *cox3* is absent from the mitochondrial genome of ciliates (Brunk *et al.*, 2003; Burger *et al.*, 2000) and there is no evidence of a mitochondrion-targeted homologue in the *Tetrahymena thermophila* nuclear genome (Eisen *et al.*, 2006).

In *Oxyrrhis*, two different genes were never found on the same EST or mitochondrial genomic fragment; many gene copies probably occur in closely-spaced tandem repeats of the same gene. Fragments of *cob* and *cox3*, but not of *cox1*, were found. Inverted repeats typical for core dinoflagellates are few in *Oxyrrhis*, which is reminiscent of the situation in *Hematodinium*.

Genes with different flanking sequences were found, but only mature polyA transcripts of uniform size can be identified like in many Dinokaryota and in *Hematodinium*. No RNA editing was found in *O. marina* and it was noted that the gene sequences correspond more to the post-edited sequence in Dinokaryota (Slamovits *et al.*, 2007; Zhang & Lin, 2008; Zhang *et al.*, 2007, 2008). Another feature unique to this species is 5′ oligoU caps of 8–9 uridines on mitochondrial mRNAs added by an unknown machinery. Not surprisingly, AUG is not used, with AUU being the more likely start codon. Oligoadenylation occurs prior to UAA stop in *cox1*, UAA stop in *cob–cox3* is created by oligoadenylation like in most Dinokaryota. Finally, oligoadenylated LSUE, LSUG, and LSU RNA10 were identified (Slamovits *et al.*, 2007; Zhang & Lin, 2008; Zhang *et al.*, 2007, 2008).

Mitochondrial DNA of *Perkinsus marinus* is incompletely characterized and its genomic architecture remains unknown (Masuda *et al.*, 2010; Zhang *et al.*, 2011). As in dinoflagellates, *cox1* hybridized to a smear of genomic molecules <10 kb in length. The picture is further complicated by short fragments of mitochondrial DNA being found in the nuclear genome. There are no signs of RNA editing, at least in *cob* and *cox1* (data for *cox3* are unavailable) and no canonical start and stop codons were identified (Jackson *et al.*, 2011; Masuda *et al.*, 2010). What is really remarkable about the *Perkinsus* mitochondrion is very extensive translational frameshifting. Frameshifts occur at all AGG (Gly) and CCC (Pro) codons. AGG and CCC usually occur in the motives UAGGY and CCCCUA, respectively, and the codons are read as AGGY or CCCCU, generating +1 or +2 frameshifts. The mechanism responsible for such frequent frameshifting remains unknown: ribosome stalling or special quadruplet- or quintuplet-decoding tRNAs might be involved. There are ten frameshifts in *cox1* (eight AGGY, two CCCCU) and seven in *cob* (six AGGY, one UAGGC). Frameshifts are not unprecedented in the mitochondrial genomes, but only one or a maximum of two occur per gene (Beckenbach, Robson, & Crozier, 2005; Milbury & Gaffney, 2008; Russell & Beckenbach, 2008). High frameshift density in *Perkinsus* must involve a unique and very efficient frameshifting mechanism. Conservation of the *cox1* sequence in *P. chesapeaki*, *P. olseni* and *P. honshuensis* implies the presence of this mechanism in all these species (Zhang *et al.*, 2011).

We are now completing a deep sequencing study of *Chromera velia* mitochondrial DNA with Illumina and 454 technologies, and of transcriptome with Illumina (Flegontov *et al.*, unpublished data). The organellar genome is essentially dinoflagellate-like (i.e. rich in non-coding

DNA, short GC-rich interspersed repeats apparently acting as recombination hot spots), gene fragments and truncated genes are abundant, canonical start and stop codons are probably missing, rRNAs are fragmented, and all transcripts are oligoadenylated. According to the RNAseq data, the entire genome seems to be transcribed at a high rate but, as in *Hematodinium*, only full-length transcripts were visualized by Northern blot analysis. Conserved *cox1* is found in the genome along with many truncated copies and fragments. Non-fragmented *cox3* is hardly recognizable at the sequence level and is invariably fused into one ORF with a 5′ fragment of *cox1*. Preliminary data from the mitochondrial genome of a distantly related chromerid, *Vitrella brassicaformis*, show a more orderly genome (Flegontov *et al.*, unpublished data) with *cob* and *cox1* fused in one ORF.

2.5. Conclusions

The evolutionary picture that emerges from this review is that of convergent evolution of radically reduced mitochondrial genomes. Tendencies for the loss of start and stop codons, rRNA fragmentation and tRNA loss appear in the ciliates, the most basal alveolate group. Ribosomal RNA fragmentation apparently reaches the maximum permissible level in the last common ancestor of the Apicomplexa–Dinoflagellata group (ADLCA) and remains constant in different lineages. Suspected cases of outright mitochondrial rRNA relocation to the nucleus are an appealing possibility but remain unverified. Oligoadenylation of all transcripts including rRNA fragments became established in the ADLCA. The most significant event that very likely happened in the ADLCA was a drastic reduction of the coding capacity: complete loss of tRNAs and most protein-coding genes except for *cox1*, *cox3* and *cob*. We suggest that this genome reduction, relaxing of selective constraints, was a prerequisite for the development of other striking molecular features independently in different lineages of the Apicomplexa–Dinoflagellata group. Loss of canonical start and stop codons occurred apparently separately in the perkinsid–dinoflagellate branch and in *Chromera* (core apicomplexans retain AUG in one gene and UAA stops in all genes). Mitochondrial genome scrambling follows the same phylogenetic pattern since unscrambling of a messy genome in the ADLCA leading to an orderly genome in core apicomplexans is difficult to imagine. RNA editing of the same type appeared probably separately in *Hematodinium* (Syndiniales) and within Dinokaryota (see the earlier discussion). Gene fusions appeared several times: *cox1–cox3* in *Chromera*, *cob–cox1* in *Vitrella*, *cob–cox3* in

Oxyrrhis. Molecular features unique to some groups include extensive frameshifting in *Perkinsus* and oligoU 5′ caps in *Oxyrrhis*.

These evolutionary pathways, in our view, are best explained by the constructive neutral evolution framework (Flegontov *et al.*, 2011; Lukeš, Archibald, Keeling, Doolittle, & Gray, 2011). Pre-existing enzymatic activities can be recruited for RNA editing, DNA recombination activities for genome scrambling, special tRNAs for frameshifting as long as their effects are tolerated, without any immediate selective benefits. Highly reduced genomes should be most tolerant of any kind of evolutionary tinkering. Hence, the abundance of molecular oddities in mitochondrial genomes of the Apicomplexa–Dinoflagellata group is not that surprising after all.

ACKNOWLEDGEMENTS

This work was supported by the Grant Agency of the Czech Republic (204/09/1667 and P305/11/2179) and the Praemium Academiae award to J.L., who is a Fellow of the Canadian Institute for Advanced Research.

REFERENCES

Adl, S. M., Simpson, A. G. B., Farmer, M. A., Andersen, R. A., Anderson, O. R., Barta, J. R., et al. (2005). The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *Journal of Eukaryotic Microbiology, 52*, 399–451.

Archibald, J. M. (2009). The puzzle of plastid evolution. *Current Biology, 19*, R81–R88.

Arunkumar, K. P., & Nagaraju, J. (2006). Unusually long palindromes are abundant in mitochondrial control regions of insects and nematodes. *PLoS ONE, 1*. e110.

Archibald, J. M., & Keeling, P. J. (2002). Recycled plastids: a "green movement" in eukaryotic evolution. *Trends in Genetics, 18*, 577–584.

Bachvaroff, T. R., Handy, S. M., Place, A. R., & Delwiche, C. F. (2011). Alveolate phylogeny inferred using concatenated ribosomal proteins. *Journal of Eukaryotic Microbiology, 58*, 223–233.

Barbrook, A. C., Howe, C. J., & Purton, S. (2006a). Why are plastid genomes retained in non-photosynthetic organisms? *Trends in Plant Science, 11*, 101–108.

Barbrook, A. C., Santucci, N., Plenderleith, L. J., Hiller, R. G., & Howe, C. J. (2006b). Comparative analysis of dinoflagellate chloroplast genomes reveals rRNA and tRNA genes. *BMC Genomics, 7*, 297.

Bartoszewski, G., Katzir, N., & Havey, M. J. (2004). Organization of repetitive DNAs and the genomic regions carrying ribosomal RNA, cob, and atp9 genes in the cucurbit mitochondrial genomes. *Theoretical and Applied Genetics, 108*, 982–992.

Beckenbach, A. T., Robson, S. K., & Crozier, R. H. (2005). Single nucleotide +1 frameshifts in an apparently functional mitochondrial cytochrome b gene in ants of the genus *Polyrhachis*. *Journal of Molecular Evolution, 60*, 141–152.

Berney, C., Fahrni, J., & Pawlowski, J. (2004). How many novel eukaryotic 'kingdoms'? Pitfalls and limitations of environmental DNA surveys. *BMC Biology, 2*, 13.

Blum, B., Bakalara, N., & Simpson, L. (1990). A model for RNA editing in kinetoplastid mitochondria: guide RNA molecules transcribed from maxicircle DNA provide the edited information. *Cell, 60*, 189–198.

Boer, P. H., & Gray, M. W. (1988). Scrambled ribosomal RNA gene pieces in *Chlamydomonas reinhardtii* mitochondrial DNA. *Cell, 55*, 399–411.

Brugerolle, G. (2002). *Colpodella vorax*: ultrastructure, predation, life-cycle, mitosis, and phylogenetic relationships. *European Journal of Protistology, 38*, 113–125.

Brunk, C. F., Lee, L. C., Tran, A. B., & Li, J. (2003). Complete sequence of the mitochondrial genome of *Tetrahymena thermophila* and comparative methods for identifying highly divergent genes. *Nucleic Acids Research, 31*, 1673–1682.

Burger, G., Zhu, Y., Littlejohn, T. G., Greenwood, S. J., Schnare, M. N., Lang, B. F., et al. (2000). Complete sequence of the mitochondrial genome of *Tetrahymena pyriformis* and comparison with *Paramecium aurelia* mitochondrial DNA. *Journal of Molecular Biology, 297*, 365–380.

Burger, G., Gray, M. W., & Lang, B. F. (2003). Mitochondrial genomes – anything goes. *Trends in Genetics, 19*, 709–716.

Cavalier-Smith, T. (2010). Kingdoms protozoa and chromista and the eozoan root of the eukaryotic tree. *Biology Letters, 6*, 342–345.

Chaput, H., Wang, Y., & Morse, D. (2002). Polyadenylated transcripts containing random gene fragments are expressed in dinoflagellate mitochondria. *Protist, 153*, 111–122.

Chrzanowska-Lightowlers, Z. M., Temperley, R. J., Smith, P. M., Seneca, S. H., & Lightowlers, R. N. (2004). Functional polypeptides can be synthesized from human mitochondrial transcripts lacking termination codons. *Biochemical Journal, 377*, 725–731.

Crouse, E. J., Vandrey, J. P., & Stutz, E. (1974). Hybridization studies with RNA and DNA isolated from *Euglena gracilis* chloroplasts and mitochondria. *FEBS Letters, 42*, 262–266.

Edqvist, J., Burger, G., & Gray, M. W. (2000). Expression of mitochondrial protein-coding genes in *Tetrahymena pyriformis*. *Journal of Molecular Biology, 297*, 381–393.

Eisen, J. A., Coyne, R. S., Wu, M., Wu, D., Thiagarajan, M., Wortman, J. R., et al. (2006). Macronuclear genome sequence of the ciliate *Tetrahymena thermophila*, a model eukaryote. *PLoS Biology, 4*, e286.

El-Sayed, N. M., et al. (2005). Comparative genomics of trypanosomatid parasitic protozoa (and 44 co-authors). *Science, 309*, 404–409.

Entelis, N. S., Kolesnikova, O. A., Dogan, S., Martin, R. P., & Tarassov, I. A. (2001). 5S rRNA and tRNA import into human mitochondria. Comparison of *in vitro* requirements. *Journal of Biological Chemistry, 276*, 45642–45653.

Esseiva, A. C., Naguleswaran, A., Hemphill, A., & Schneider, A. (2004). Mitochondrial tRNA import in *Toxoplasma gondii*. *Journal of Biological Chemistry, 279*, 42363–42368.

Feagin, J. E. (1992). The 6-kb element of *Plasmodium falciparum* encodes mitochondrial cytochrome genes. *Molecular and Biochemical Parasitology, 52*, 145–148.

Feagin, J. E., Mericle, B. L., Werner, E., & Morris, M. (1997). Identification of additional rRNA fragments encoded by the *Plasmodium falciparum* 6 kb element. *Nucleic Acids Research, 25*, 438–446.

Fernández Robledo, J. A., Caler, E., Matsuzaki, M., Keeling, P. J., Shanmugam, D., Roos, D. S., et al. (2011). The search for the missing link: a relic plastid in *Perkinsus*? *International Journal of Parasitology, 41*, 1217–1229.

Flegontov, P., Gray, M. W., Burger, G., & Lukeš, J. (2011). Gene fragmentation: a key to mitochondrial genome evolution in Euglenozoa? *Current Genetics, 57*, 225–232.

Gardner, M. J., Hall, N., Fung, E., White, O., Berriman, M., et al. (2002). Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature, 419*, 498–511.

Gillespie, D. E., Salazar, N. A., Rehkopf, D. H., & Feagin, J. E. (1999). The fragmented mitochondrial ribosomal RNAs of *Plasmodium falciparum* have short A tails. *Nucleic Acids Research, 27*, 2416–2422.

Gómez, F., López-García, P., Nowaczyk, A., & Moreira, D. (2009). The crustacean parasites *Ellobiopsis* Caullery, 1910 and *Thalassomyces* Niezabitowski, 1913 form a monophyletic divergent clade within the Alveolata. *Systematic Parasitology, 74*, 65–74.

Gómez, F., Moreira, D., & López-García, P. (2010). Molecular phylogeny of noctilucoid dinoflagellates (Noctilucales, Dinophyceae). *Protist, 161*, 466–478.

Gray, M. W. (2003). Diversity and evolution of mitochondrial RNA editing systems. *IUBMB Life, 55*, 227–233.

Gray, M. W., Lang, B. F., & Burger, G. (2004). Mitochondria of protists. *Annual Review of Genetics, 38*, 477–524.

Hall, R. P. (2005). Reaction of certain cytoplasmic inclusions to vital dyes and their relation to mitochondria and golgi apparatus in the flagellate *Peranema trichophorum*. *Journal of Morphology, 48*, 105–121.

Groisillier, A., Massana, R., Valentin, K., Vaulot, D., & Guillou, L. (2006). Genetic diversity and habitats of two enigmatic marine alveolate lineages. *Aquatic Microbial Ecology, 42*, 277–291.

Guillou, L., Viprey, M., Chambouvet, A., Welsh, R. M., Kirkham, A. R., Massana, R., et al. (2008). Widespread occurrence and genetic diversity of marine parasitoids belonging to Syndiniales (Alveolata). *Environmental Microbiology, 10*, 3349–3365.

Hackett, J., Anderson, D., Erdner, D., & Bhattacharya, D. (2004). Dinoflagellates: a remarkable evolutionary experiment. *American Journal of Botany, 91*, 1523–1534.

Hallick, R. B., Hong, L., Drager, R. G., Favreau, M. R., Monfort, A., Orsat, B., et al. (1993). Complete sequence of *Euglena gracilis* chloroplast DNA. *Nucleic Acids Research, 21*, 3537–3544.

Hannaert, V., Saavedra, E., Duffieux, F., Szikora, J.-P., Rigden, D. J., Michels, P. A. M., et al. (2003). Plant-like traits associated with metabolism of *Trypanosoma* parasites. *Proceedings of the National Academy of Sciences of the United States of America, 100*, 1067–1071.

Harada, A., Ohtsuka, S., & Horiguchi, T. (2007). Species of the parasitic genus *Duboscquella* are members of the enigmatic marine alveolate group I. *Protist, 158*, 337–347.

Hayashi, Y., & Ueda, K. (1989). The shape of the mitochondria and the number of mitochondrial nucleoids during the cell cycle of *Euglena gracilis*. *Journal of Cell Science, 93*, 565–570.

Hayashi-Isimaru, Y., Ueda, K., & Nonaka, M. (1993). Detection of DNA in the nucleoids of chloroplasts and mitochondria in *Euglena gracilis* by immunoelectron microscopy. *Journal of Cell Science, 105*, 1159–1164.

Hoppenrath, M., & Leander, B. S. (2010). Dinoflagellate phylogeny as inferred from heat shock protein 90 and ribosomal gene sequences. *PLoS ONE, 5*. e13220.

Howe, C. J., & Purton, S. (2007). The little genome of apicomplexan plastids: its raison d'etre and a possible explanation for the 'delayed death' phenomenon. *Protist, 158*, 121–133.

Imanian, B., & Keeling, P. J. (2007). The dinoflagellates *Durinskia baltica* and *Kryptoperidinium foliaceum* retain functionally overlapping mitochondria from two evolutionarily distinct lineages. *BMC Evolutionary Biology, 7*, 172.

Jackson, C. J., Norman, J. E., Schnare, M. N., Gray, M. W., Keeling, P. J., & Waller, R. F. (2007). Broad genomic and transcriptional analysis reveals a highly derived genome in dinoflagellate mitochondria. *BMC Biology, 5*, 41.

Jackson, C. J., Gornik, S. G., & Waller, R. F. (2011). The mitochondrial genome and transcriptome of the basal dinoflagellate *Hematodinium* sp.: character evolution within the highly derived mitochondrial genomes of dinoflagellates. *Genome Biology and Evolution* 59–72. http://dx.doi.org/10.1093/gbe/evr122.

Janouškovec, J., Horák, A., Oborník, M., Lukeš, J., & Keeling, P. J. (2010). A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids. *Proceedings of the National Academy of Sciences of the United States of America, 107*, 10949–10954.

Joseph, S. J., Fernández-Robledo, J. A., Gardner, M. J., El-Sayed, N. M., Kuo, C.-H., Schott, E. J., et al. (2010). The alveolate *Perkinsus marinus*: biological insights from EST gene discovery. *BMC Genomics, 11*, 228.

Kairo, A., Fairlamb, A. H., Gobright, E., & Nene, V. (1994). A 7.1 kb linear DNA molecule of *Theileria parva* has scrambled rDNA sequences and open reading frames for mitochondrially encoded proteins. *EMBO Journal, 13*, 898–905.

Kamikawa, R., Inagaki, Y., & Sako, Y. (2007). Fragmentation of mitochondrial large subunit rRNA in the dinoflagellate *Alexandrium catenella* and the evolution of rRNA structure in alveolate mitochondria. *Protist, 158*, 239–245.

Kamikawa, R., Hosoi-Tanabe, S., Yoshimatsu, S., Oyama, K., Masuda, I., & Sako, Y. (2008). Development of a novel molecular marker on the mitochondrial genome of a toxic dinoflagellates, *Alexandrium* spp., and its application in single-cell PCR. *Journal of Applied Phycology, 20*, 153–159.

Kamikawa, R., Nishimura, H., & Sako, Y. (2009). Analysis of the mitochondrial genome, transcripts, and electron transport activity in the dinoflagellate *Alexandrium catenella* (Gonyaulacales, Dinophyceae). *Phycology Research, 57*, 1–11.

Keeling, P. J., Burger, G., Durnford, D. G., Lang, B. F., Lee, R. W., Pearlman, R. E., et al. (2005). The tree of eukaryotes. *Trends in Ecology & Evolution, 20*, 670–676.

Keeling, P. J. (2008). Bridge over troublesome plastids. *Nature, 451*, 896–897.

Keeling, P. J. (2009). Chromalveolates and the evolution of plastids by secondary endosymbiosis. *Journal of Eukaryotic Microbiology, 56*, 1–8.

Keithly, J. S., Langreth, S. G., Buttle, K. F., & Mannella, C. A. (2005). Electron tomographic and ultrastructural analysis of the *Cryptosporidium parvum* relict mitochondrion, its associated membranes, and organelles. *Journal of Eukaryotic Microbiology, 52*, 132–140.

Kiethega, G. N., Turcotte, M., & Burger, G. (2011). Evolutionarily conserved cox1 trans-splicing without cis-motifs. *Molecular Biology and Evolution, 28*, 2425–2428.

Knoop, V. (2004). The mitochondrial DNA of land plants: peculiarities in phylogenetic perspective. *Current Genetics, 46*, 123–139.

Kuhn, J., Tengler, U., & Binder, S. (2001). Transcript lifetime is balanced between stabilizing stem-loop structures and degradation-promoting polyadenylation in plant mitochondria. *Molecular and Cellular Biology, 21*, 731–742.

Kuvardina, O. N., Leander, B. S., Aleshin, V. V., Mylnikov, A. P., Keeling, P. J., & Simdyanov, T. G. (2002). The phylogeny of colpodellids (Alveolata) using small subunit rRNA gene sequences suggests they are the free-living sister group to apicomplexans. *Journal of Eukaryotic Microbiology, 49*, 498–504.

Leander, B. S., Triemer, R. E., & Farmer, M. A. (2001). Character evolution in heterotrophic euglenids. *European Journal of Protistology, 37*, 337–356.

Leander, B. S., & Keeling, P. J. (2003). Morphostasis in alveolate evolution. *Trends in Ecology & Evolution, 18*, 395–402.

Leander, B. S. (2004). Did trypanosomatid parasites have photosynthetic ancestors? *Trends in Microbiology, 12*, 251–258.

Leander, B. S., & Keeling, P. J. (2004). Early evolutionary history of dinoflagellates and apicomplexans (Alveolata) as inferred from hsp90 and actin phylogenies. *Journal of Phycology, 40*, 341–350.

Leander, B. S., Kuvardina, O. N., Aleshin, V. V., Mylnikov, A. P., & Keeling, P. J. (2003). Molecular phylogeny and surface morphology of *Colpodella edax* (Alveolata): insights into the phagotrophic ancestry of apicomplexans. *Journal of Eukaryotic Microbiology, 50*, 334–340.

Lin, S., Zhang, H., Spencer, D. F., Norman, J. E., & Gray, M. W. (2002). Widespread and extensive editing of mitochondrial mRNAS in dinoflagellates. *Journal of Molecular Biology, 320*, 727–739.

Lin, S., Zhang, H., & Jiao, N. (2006). Potential utility of mitochondrial cytochrome b and its mRNA editing in resolving closely related dinoflagellates: a case study of *Prorocentrum* (Dinophyceae). *Journal of Phycology, 42*, 646–654.

Lin, S., Zhang, H., & Gray, M. W. (2008). RNA editing in dinoflagellates and its implications for the evolutionary history of the editing machinery. In H. C. Smith (Ed.), *RNA and DNA editing: Molecular mechanisms and their integration into biological systems* (pp. 280–309). Hoboken, NJ: John Wiley and Sons, Inc.

Lonergan, K. M., & Gray, M. W. (1996). Expression of a continuous open reading frame encoding subunits 1 and 2 of cytochrome c oxidase in the mitochondrial DNA of *Acanthamoeba castellanii*. *Journal of Molecular Biology, 257*, 1019–1030.

Lowe, C. D., Keeling, P. J., Martin, L. E., Slamovits, C. H., Watts, P. C., & Montagnes, D. J. S. (2011). Who is *Oxyrrhis marina*? Morphological and phylogenetic studies on an unusual dinoflagellate. *Journal of Plankton Research, 33*, 555–567.

Lukeš, J., Guilbride, D. L., Votýpka, J., Zíková, A., Benne, R., & Englund, P. T. (2002). Kinetoplast DNA network: evolution of an improbable structure. *Eukaryotic Cell, 1*, 495–502.

Lukeš, J., Hashimi, H., & Zíková, A. (2005). Unexplained complexity of the mitochondrial genome and transcriptome in kinetoplastid flagellates. *Current Genetics, 48*, 277–299.

Lukeš, J., Leander, B. S., & Keeling, P. J. (2009). Cascades of convergent evolution: the corresponding evolutionary histories of euglenozoans and dinoflagellates. *Proceedings of the National Academy of Sciences of the United States of America, 106*, 9963–9970.

Lukeš, J., Archibald, J. M., Keeling, P. J., Doolittle, W. F., & Gray, M. W. (2011). How a neutral evolutionary ratchet can build cellular complexity. *IUBMB Life, 63*, 528–537.

Marande, W., Lukeš, J., & Burger, G. (2005). Unique mitochondrial genome structure in diplonemids, the sister group of kinetoplastids. *Eukaryotic Cell, 4*, 1137–1146.

Marande, W., & Burger, G. (2007). Mitochondrial DNA as a genomic jigsaw puzzle. *Science, 318*, 415.

Masuda, I., Matsuzaki, M., & Kita, K. (2010). Extensive frameshift at all AGG and CCC codons in the mitochondrial cytochrome c oxidase subunit 1 gene of *Perkinsus marinus* (Alveolata; Dinoflagellata). *Nucleic Acids Research, 38*, 6186–6194.

Milbury, C. A., & Gaffney, P. M. (2008). Complete mitochondrial sequence of the eastern oyster *Crassostrea virginica*. *Marine Biotechnology, 7*, 697–712.

Moore, R. B., Oborník, M., Janouškovec, J., Chrudimský, T., Vancová, M., Green, D. H., et al. (2008). A photosynthetic alveolate closely related to apicomplexan parasites. *Nature, 451*, 959–963.

Mylnikov, A. P. (2009). Ultrastructure and phylogeny of colpodellids (Colpodellida, Alveolata). *The Biological Bulletin, 36*, 582–590.

Nash, E. A., Barbrook, A. C., Edwards-Stuart, R. K., Bernhardt, K., Howe, C. J., & Nisbet, E. R. (2007). Organization of the mitochondrial genome in the dinoflagellate *Amphidinium carterae*. *Molecular Biology and Evolution, 24*, 1528–1536.

Nash, E. A., Nisbet, R. E., Barbrook, A. C., & Howe, C. J. (2008). Dinoflagellates: a mitochondrial genome all at sea. *Trends in Genetics, 24*, 328–335.

Nass, M. M. K., Schori, L., Ben-Shaul, Y., & Edelman, M. (1974). Size and configuration of mitochondrial DNA in *Euglena gracilis*. *Biochimica et Biophysica Acta, 374*, 283–291.

Nelson, M. J., Dang, Y., Filek, E., Zhang, Z., Yu, V. W., Ishida, K., et al. (2007). Identification and transcription of transfer RNA genes in dinoflagellate plastid mini-circles. *Gene, 392*, 291–298.

Norman, J. E., & Gray, M. W. (1997). The cytochrome oxidase subunit 1 gene (cox1) from the dinoflagellate, *Crypthecodinium cohnii*. *FEBS Letters, 413*, 333–338.

Norman, J. E., & Gray, M. W. (2001). A complex organization of the gene encoding cytochrome oxidase subunit 1 in the mitochondrial genome of the dinoflagellate,

Crypthecodinium cohnii: homologous recombination generates two different cox1 open reading frames. *Journal of Molecular Evolution, 53*, 351–363.

Oborník, M., Janouškovec, J., Chrudimský, T., & Lukeš, J. (2009). Evolution of the apicoplast and its hosts: from heterotrophy to autotrophy and back again. *International Journal of Parasitology, 39*, 1–12.

Oborník, M., Vancová, M., Lai, D. H., Janouškovec, J., Keeling, P. J., & Lukeš, J. (2011). Morphology and ultrastructure of multiple life cycle stages of the photosynthetic relative of apicomplexa, *Chromera velia. Protist, 162*, 115–130.

Oborník, M., Modrý, D., Lukeš, M., Černotíková-Stříbrná, E., Cihlář, J., Tesařová, M., et al. (2012). Morphology, ultrastructure and life cycle of *Vitrella brassicaformis* n. sp., n. gen., a novel chromerid from the Great Barrier Reef. *Protist., 163*, 306–323.

Okamoto, N., & McFadden, G. I. (2008). The mother of all parasites. *Future Microbiology, 3*, 391–395.

Panigrahi, A. K., Ogata, Y., Zikova, A., Anupama, A., Dalley, R. A., Acestor, N., et al. (2009). A comprehensive analysis of *Trypanosoma brucei* mitochondrial proteome. *Proteomics, 9*, 434–450.

Raczynska, K. D., Le Ret, M., Rurek, M., Bonnard, G., Augustyniak, H., & Gualberto, J. M. (2006). Plant mitochondrial genes can be expressed from mRNAs lacking stop codons. *FEBS Letters, 580*, 5641–5646.

Ray, D. S., & Hanawalt, P. C. (1965). Satellite DNA components in *Euglena gracilis* cells lacking chloroplasts. *Journal of Molecular Biology, 11*, 760–768.

Rehkopf, D. H., Gillespie, D. E., Harrell, M. I., & Feagin, J. E. (2000). Transcriptional mapping and RNA processing of the *Plasmodium falciparum* mitochondrial mRNAs. *Molecular and Biochemical Parasitology, 105*, 91–103.

Roger, A. J., & Simpson, A. G. B. (2009). Revisiting the root of the eukaryotic tree. *Current Biology, 19*, R165–R167.

Roy, J., Faktorová, D., Lukeš, J., & Burger, G. (2007). Unusual mitochondrial genome structures throughout the Euglenozoa. *Protist, 158*, 385–396.

Rusconi, C. P., & Cech, T. R. (1996). Mitochondrial import of only one of three nuclear-encoded glutamine tRNAs in *Tetrahymena thermophila. EMBO Journal, 15*, 3286–3295.

Russell, R. D., & Beckenbach, A. T. (2008). Recoding of translation in turtle mitochondrial genomes: programmed frameshift mutations and evidence of a modified genetic code. *Journal of Molecular Evolution, 67*, 682–695.

Saldarriaga, J. F., Taylor, F. J. R. M., Cavalier-Smith, T., Menden-Deuer, S., & Keeling, P. J. (2004). Molecular data and the evolutionary history of dinoflagellates. *European Journal of Protistology, 40*, 85–111.

Simspon, A. G. B., Gill, E. E., Callahan, H. A., Litaker, R. W., & Roget, A. G. (2004). Early evolution within kinetoplastids (Euglenozoa), and the late emergence of trypanosomatids. *Protist, 155*, 407–422.

Simpson, A. G. B., Stevens, J. R., & Lukeš, J. (2006). The evolution and diversity of kinetoplastid flagellates. *Trends in Parasitology, 22*, 168–174.

Skovgaard, A., Massana, R., Balagué, V., & Saiz, E. (2005). Phylogenetic position of the copepod-infesting parasite *Syndinium turbo* (Dinoflagellata, Syndinea). *Protist, 156*, 413–423.

Skovgaard, A., Meneses, I., & Angélico, M. M. (2009). Identifying the lethal fish egg parasite *Ichthyodinium chabelardi* as a member of Marine Alveolate Group I. *Environmental Microbiology, 11*, 2030–2041.

Slamovits, C. H., Saldarriaga, J. F., Larocque, A., & Keeling, P. J. (2007). The highly reduced and fragmented mitochondrial genome of the early-branching dinoflagellate *Oxyrrhis marina* shares characteristics with both apicomplexan and dinoflagellate mitochondrial genomes. *Journal of Molecular Biology, 372*, 356–368.

Slamovits, C. H., & Keeling, P. J. (2011). Contributions of *Oxyrrhis marina* to molecular biology, genomics and organelle evolution of dinoflagellates. *Journal of Plankton Research, 33*, 591–602.

Spencer, D. F., & Gray, M. W. (2011). Ribosomal RNA genes in *Euglena gracilis* mitochondrial DNA: fragmented genes in a seemingly fragmented genome. *Molecular Genetics & Genomics, 285*, 19–31.

Stentiford, G. D., & Shields, J. D. (2005). A review of the parasitic dinoflagellates *Hematodinium* species and *Hematodinium*-like infections in marine crustaceans. *Diseases of Aquatic Organisms, 66*, 47–70.

Stuart, K. D., Schnaufer, A., Ernst, N. L., & Panigrahi, A. K. (2005). Complex management: RNA editing in trypanosomes. *Trends in Biochemical Sciences, 30*, 97–105.

Talen, J. L., Sanders, J. P. M., & Flavell, R. A. (1974). Genetic complexity of mitochondrial DNA from *Euglena gracilis*. *Biochimica et Biophysica Acta, 374*, 129–135.

Tessier, L. H., van der Speck, H., Gualberto, J. M., & Grienenberger, J. M. (1997). The cox1 gene from *Euglena gracilis*: a protist mitochondrial gene without introns and genetic code modifications. *Current Genetics, 31*, 208–213.

Týč, J., Long, S., Jirků, M., & Lukeš, J. (2010). YCF45 protein, usually associated with plastids, is targeted into the mitochondrion of *Trypanosoma brucei*. *Molecular and Biochemical Parasitology, 173*, 43–47.

Vlček, C., Marande, W., Teijeiro, S., Lukeš, J., & Burger, G. (2011). Systematically fragmented genes in a multipartite mitochondrial genome. *Nucleic Acids Research, 39*, 979–988.

Waller, R. F., & Jackson, C. J. (2009). Dinoflagellate mitochondrial genomes: stretching the rules of molecular biology. *Bioessays, 31*, 237–245.

Waller, R. F., & Keeling, P. J. (2006). Alveolate and chlorophycean mitochondrial cox2 genes split twice independently. *Gene, 383*, 33–37.

Weatherby, K., Murray, S., Carter, D., & Šlapeta, J. (2011). Surface and flagella morphology of the motile form of *Chromera velia* revealed by field-emission scanning electron microscopy. *Protist, 162*, 142–153.

Yasuhira, S., & Simpson, L. (1996). Guide RNAs and guide RNA genes in the cryptobiid kinetoplastid protozoan, *Trypanoplasma borreli*. *RNA, 2*, 1153–1160.

Yasuhira, S., & Simpson, L. (1997). Phylogenetic affinity of mitochondria of *Euglena gracilis* and kinetoplastids using cytochrome oxidase I and hsp60. *Journal of Molecular Evolution, 44*, 341–347.

Zhang, H., Bhattacharya, D., Maranda, L., & Lin, S. (2008). Mitochondrial *cob* and *cox1* genes and editing of the corresponding mRNAs in *Dinophysis acuminata* from Narragansett Bay, with special reference to the phylogenetic position of the genus *Dinophysis*. *Applied and Environmental Microbiology, 74*, 1546–1554.

Zhang, H., & Lin, S. (2002). Detection and quantification of *Pfiesteria piscicida* by using the mitochondrial cytochrome *b* gene. *Applied and Environmental Microbiology, 68*, 989–994.

Zhang, H., & Lin, S. (2005). Mitochondrial cytochrome b mRNA editing in dinoflagellates: possible ecological and evolutionary associations? *Journal of Eukaryotic Microbiology, 52*, 538–545.

Zhang, H., & Lin, S. (2008). mRNA editing and spliced-leader RNA trans-splicing groups *Oxyrrhis, Noctiluca*, *Heterocapsa*, and *Amphidinium* as basal lineages of dinoflagellates. *Journal of Phycology, 44*, 703–711.

Zhang, H., Campbell, D. A., Sturm, N. R., Dungan, C. F., & Lin, S. (2011). Spliced leader RNAs, mitochondrial gene frameshifts and multi-protein phylogeny expand support for the genus *Perkinsus* as a unique group of alveolates. *PLoS ONE, 6*. e19933.