Protist, Vol. xx, xxx–xxx, xx 2012
http://www.elsevier.de/protis
Published online date xxx

Protist

ORIGINAL PAPER

An Analysis of Dinoflagellate Metabolism Using EST Data

Erin R. Butterfield[a], Christopher J. Howe[b], and R. Ellen R. Nisbet[a,1]

[a]Sansom Institute for Health Research, University of South Australia, North Terrace, Adelaide, SA 5000, Australia
[b]Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge, Cambridgeshire, CB2 1QW, United Kingdom

Submitted April 20, 2012; Accepted September 10, 2012
Monitoring Editor: Michael L. Ginger

The dinoflagellates are an important group of eukaryotic, single celled algae. They are the sister group of the Apicomplexa, a group of intracellular parasites and photosynthetic algae including the malaria parasite *Plasmodium*. Many apicomplexan mitochondria have a number of unusual features, including the lack of a pyruvate dehydrogenase and the existence of a branched TCA cycle. Here, we analyse dinoflagellate EST (expressed sequence tag) data to determine whether these features are apicomplexan-specific, or if they are more widespread. We show that dinoflagellates have replaced a key subunit (E1) of pyruvate dehydrogenase with a subunit of bacterial origin and that transcripts encoding many of the proteins that are essential in a conventional ATP synthase/Complex V are absent, as is the case in Apicomplexa. There is a pathway for synthesis of starch or glycogen as a storage carbohydrate. Transcripts encoding isocitrate lyase and malate synthase are present, consistent with ultrastructural reports of a glyoxysome. Finally, evidence for a conventional haem biosynthesis pathway is found, in contrast to the Apicomplexa, *Chromera* and early branching dinoflagellates (*Perkinsus, Oxyrrhis*).

Key words: Dinoflagellate; mitochondria; metabolism; EST analysis; alveolate; glyoxysome.

Introduction

Dinoflagellates are unicellular, flagellate protists found in a diversity of ecosystems. Approximately 50% of taxa are photosynthetic, while the remainder have lost the ability to carry out photosynthesis. Photosynthetic dinoflagellates are important primary producers, forming part of both the marine and fresh water phytoplankton. A few species can produce red tides, toxic algal blooms that can lead to shellfish poisoning. Dinoflagellates also form symbioses with corals and a breakdown of this symbiosis, perhaps in response to elevated ocean temperatures, leads to expulsion of the symbionts and coral bleaching (Sanchez-Puerta et al. 2007).

Some of the most unusual genomic organizations of all species are found in the dinoflagellate algae. The nuclear genome is exceptionally large (Holm-Hansen 1969), with multiple copies of many genes, often arranged in tandem repeats (Bachvaroff and Place 2008). The smallest dinoflagellate genome, of *Symbiodinium*, is estimated to be 3×10^9 bp, similar in scale to the human genome (Lin 2006). Therefore, no genome has yet been fully sequenced and instead, extensive EST

[1]Corresponding author; fax +61 8 830 22389
e-mail ellen.nisbet@unisa.edu.au (R.E.R. Nisbet).

http://dx.doi.org/10.1016/j.protis.2012.09.001

Please cite this article in press as: Butterfield ER, et al. An Analysis of Dinoflagellate Metabolism Using EST Data. Protist (2012), http://dx.doi.org/10.1016/j.protis.2012.09.001

libraries have been created (Bachvaroff et al. 2004; Hackett et al. 2005; Leggat et al. 2007; Slamovits and Keeling 2008).

The chloroplast genome is fragmented and greatly reduced (Koumandou et al. 2004). Recently, significant fragments of the dinoflagellate mitochondrial genome have been sequenced from a number of species including *Alexandrium catenella, Amphidinium carterae, Crypthecodinium cohnii*, *Karlodinium micrum* and *Oxyrrhis marina* (reviewed in Nash et al. 2008). All contain just three protein-encoding genes (*coxI, coxIII, cob*) and fragmentary rRNA-encoding genes, with the majority of genes having been relocated to the nucleus (Jackson et al. 2007; Nash et al. 2007, 2008; Slamovits et al. 2007).

Dinoflagellates are a sister group to the Apicomplexa, a group including a number of obligate parasites which include many important human and animal pathogens such as *Plasmodium* (causative agent of malaria), *Toxoplasma* and *Cryptosporidium* (Walker et al. 2011). Many of the Apicomplexa contain a remnant chloroplast (Dorrell and Smith 2011), but only *Chromera* and closely related species are able to carry out photosynthesis (Moore et al. 2008). The dinoflagellates and Apicomplexa are sister groups to the ciliates, as shown in Figure 1a. Ciliates are a large and diverse group of ciliated protozoa which can be free-living, facultatively parasitic or parasitic. Many species are symbionts. There is no evidence for the presence of a chloroplast in ciliates, although a few genes of possible algal origin have been identified (Archibald 2008). The ciliates, dinoflagellates and Apicomplexa together are referred to as the alveolates.

It has been known for many years that mitochondrial function in *Plasmodium* is unconventional. This has been shown both by biochemical and genome analyses (Bowman et al. 1961; Gardner et al. 2002). In the asexual stage the majority of cellular ATP is synthesized through glycolysis and lactic acid production (Uyemura et al. 2004). There is no evidence for a mitochondrial acetyl-CoA source, as the sole pyruvate dehydrogenase (PDH) is located in the apicoplast (Foth et al. 2005). Despite this, it is clear that the *Plasmodium* mitochondrion is functional (Howe and Purton 2007; Painter et al. 2007). A recent metabolic analysis has shown that the *Plasmodium* TCA cycle is branched, bifurcating at 2-oxoglutarate dehydrogenase, and that malate is produced as a waste product (Olszewski et al. 2010). The mitochondrial electron transport chain is missing a conventional complex I, replacing it with a single-subunit NADH dehydrogenase (Mogi and Kita 2010), and major components

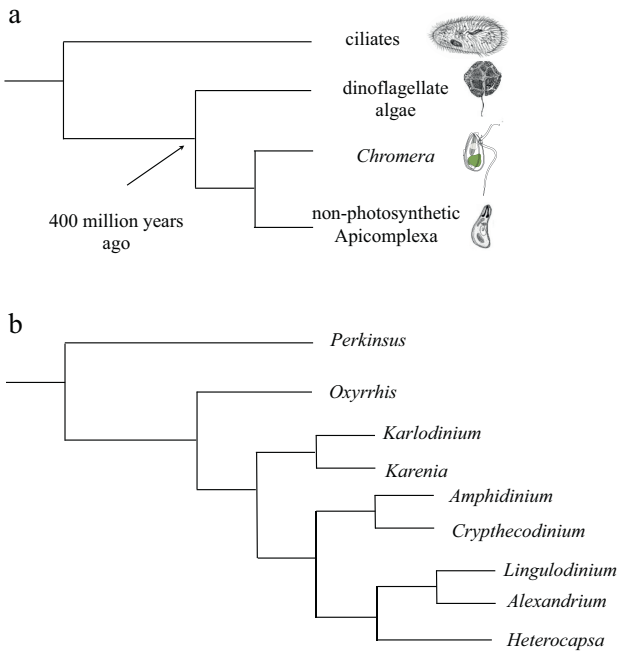


Figure 1. Dinoflagellates, Apicomplexa and ciliate relationships. **a**. Schematic diagram showing relationships within the Alveolata: ciliates (non-photosynthetic), dinoflagellate algae, and the Apicomplexa. *Chromera* is a recently-discovered photosynthetic apicomplexan. A diagram of a model species is shown for each group. **b**. Schematic diagram showing relationships between dinoflagellate species for which EST datasets are available. The diagram is based on two hsp90 and rDNA phylogenies inferred by Hoppenrath and Leander (2010). There is poor phylogenetic support for many dinoflagellate relationships, and so this diagram should be taken as a guide only. However, it is clear that *Perkinsus* and *Oxyrrhis* are early-branching dinoflagellates and the remaining groups are monophyletic.

from complex V (ATP synthase) are also absent. Instead, the electron transport chain may be used for the regeneration of ubiquinone (Painter et al. 2007; van Dooren et al. 2006). Other apicomplexan species have a diversity of mitochondrial metabolic functions, with *Cryptosporidium parvum* containing a mitosome, a derived mitochondrion lacking a genome (Henriquez et al. 2005).

In contrast, the ciliates contain a conventional mitochondrion, undergoing 'textbook style' biochemistry. A proteomic analysis of isolated mitochondria from the ciliate *Tetrahymena thermophila* identified all key components of the TCA cycle (Smith et al. 2007b), and a functional electron transport chain, together with a significant number of proteins (45%) with no known function.

Please cite this article in press as: Butterfield ER, et al. An Analysis of Dinoflagellate Metabolism Using EST Data. Protist (2012), http://dx.doi.org/10.1016/j.protis.2012.09.001

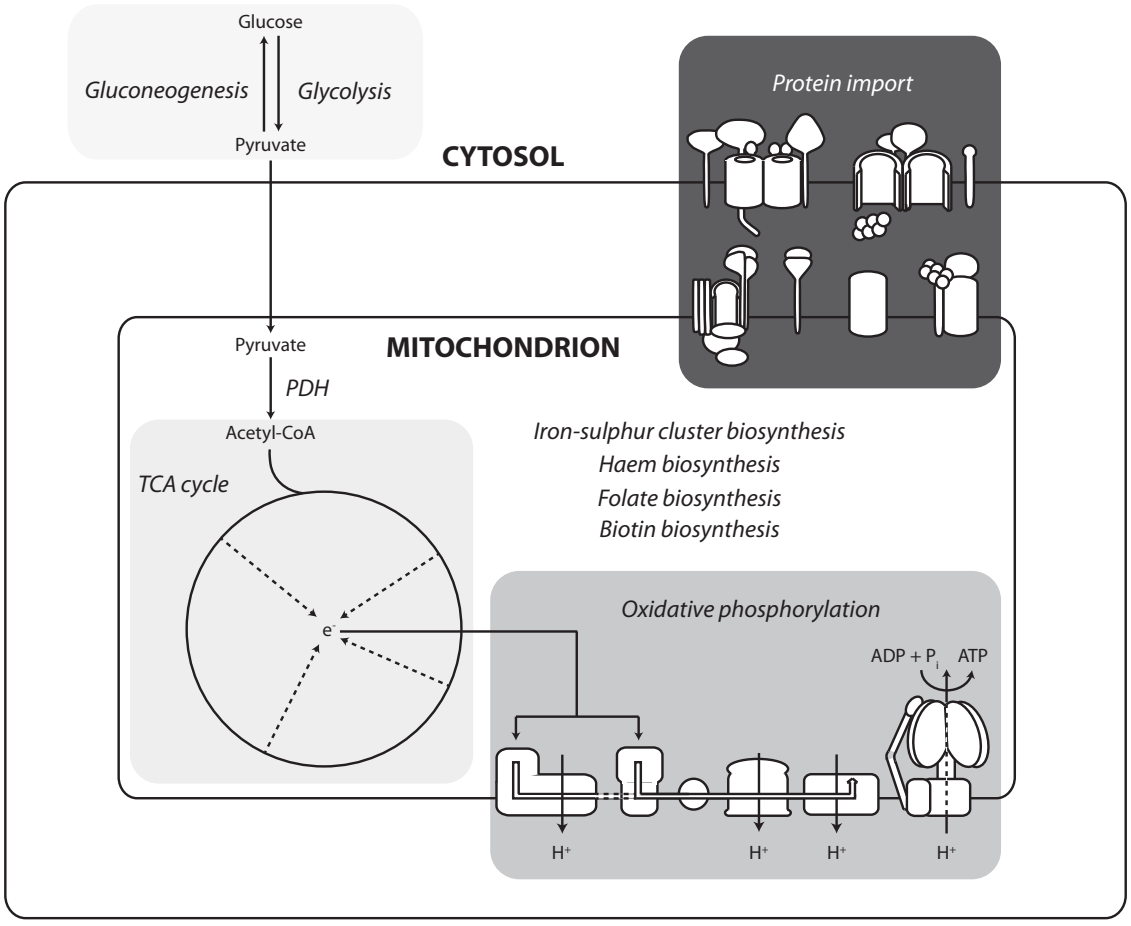


Figure 2. Major biochemical pathways essential to mitochondrial function. Glycolysis (shaded lightest grey) – although not a mitochondrial pathway – provides pyruvate for the mitochondrion. Following its import into the mitochondrion, pyruvate is converted to acetyl-CoA via PDH for use in the TCA cycle (shaded light grey). Reducing equivalents produced during the TCA cycle are used for the process of oxidative phosphorylation (shaded dark grey) where they power the electron transport chain. The proton motive force created in this process is used for the synthesis of ATP by Complex V/ATP synthase and also helps power protein import (shaded darkest grey). These pathways are supported by several different co-factors whose synthesis pathways primarily occur in the mitochondrion. Synthesis pathways include Fe-S cluster biosynthesis, haem biosynthesis, folate biosynthesis and biotin biosynthesis. Key PDH: pyruvate dehydrogenase complex, TCA cycle: tricarboxylic acid cycle.

Little is known about dinoflagellate metabolism. We have therefore searched the combined dinoflagellate EST libraries for transcripts encoding proteins involved in key metabolic pathways including glycolysis, gluconeogenesis, oxidative phosphorylation, iron-sulphur cluster biosynthesis, haem biosynthesis, folate biosynthesis, biotin biosynthesis and protein import, since these pathways are critical to mitochondrial function (Fig. 2). We show that several features of mitochondrial biochemistry are different from the ciliates and Apicomplexa. We also show that there is diversity within the dinoflagellates, with the basal dinoflagellate lineages (*Perkinsus* and *Oxyrrhis*, Fig. 1b) containing small, but significant differences.

Results and Discussion

Throughout this section, full details of each transcript (including NCBI identification and species in which the transcript was identified) are given in the Supplementary Table S1.

Please cite this article in press as: Butterfield ER, et al. An Analysis of Dinoflagellate Metabolism Using EST Data. Protist (2012), http://dx.doi.org/10.1016/j.protis.2012.09.001

Synthesis of ATP

One of the key functions of mitochondria is the conservation of energy in the form of ATP. In classical mitochondria, the synthesis of ATP by complex V is the final step in a long series of biochemical reactions that starts in the cytosol with the production of pyruvate by glycolysis. Pyruvate enters the mitochondrion where it is oxidatively decarboxylated to form acetyl-CoA and then oxidised to CO_2 by the TCA cycle, releasing NADH and $FADH_2$. These enter the electron transport chain, resulting in the synthesis of ATP with oxygen as a terminal acceptor, releasing water (Fig. 3).

Glycolysis/Gluconeogenesis and conversion of pyruvate: We searched for sequences encoding key components of the glycolysis pathway in the dinoflagellate EST databases. Although ESTs corresponding to the majority of enzymes were present, there were none encoding hexokinase, an enzyme which encodes an irreversible step in glycolysis. Instead, in *Perkinsus* only, a transcript encoding glucokinase was present, an enzyme which catalyses the same reaction but with a lower affinity for glucose (a protein alignment is shown in Supplementary Table S2). The use of glucokinase at the start of the glycolysis pathway has previously been reported in a number of protist lineages including *Giardia*, *Trichomonas, Trypanosoma* and *Leishmania* (Cáceres et al. 2007; Henze et al. 2001; Wu et al. 2001), although many other eukaryotes contain both enzymes. Ciliates are also missing hexokinase (Smith et al. 2007b), and instead contain a glucokinase (NCBI ref: XP_002287250). Thus, it would appear that the gene encoding a glucokinase has independently been transferred into the eukaryotic lineage several times (Henze et al. 2001), and the presence in dinoflagellates and ciliates adds a further gene transfer event(s) to those already described. No transcript encoding glucose-6-phosphatase was identified. This is perhaps not surprising as glucose-6-phosphatase activity, if combined with a bidirectional glucose transport pathway, would lead to glucose efflux.

There are previous reports of dinoflagellates using starch as a storage polysaccharide (Deschamps et al. 2008). This was confirmed by our EST searches, which revealed transcripts encoding phosphoglucomutase, phosphorylase, starch synthase and UDP- or ADP-glucose pyrophosphorylase, all required for the use of starch or glycogen as a storage carbohydrate.

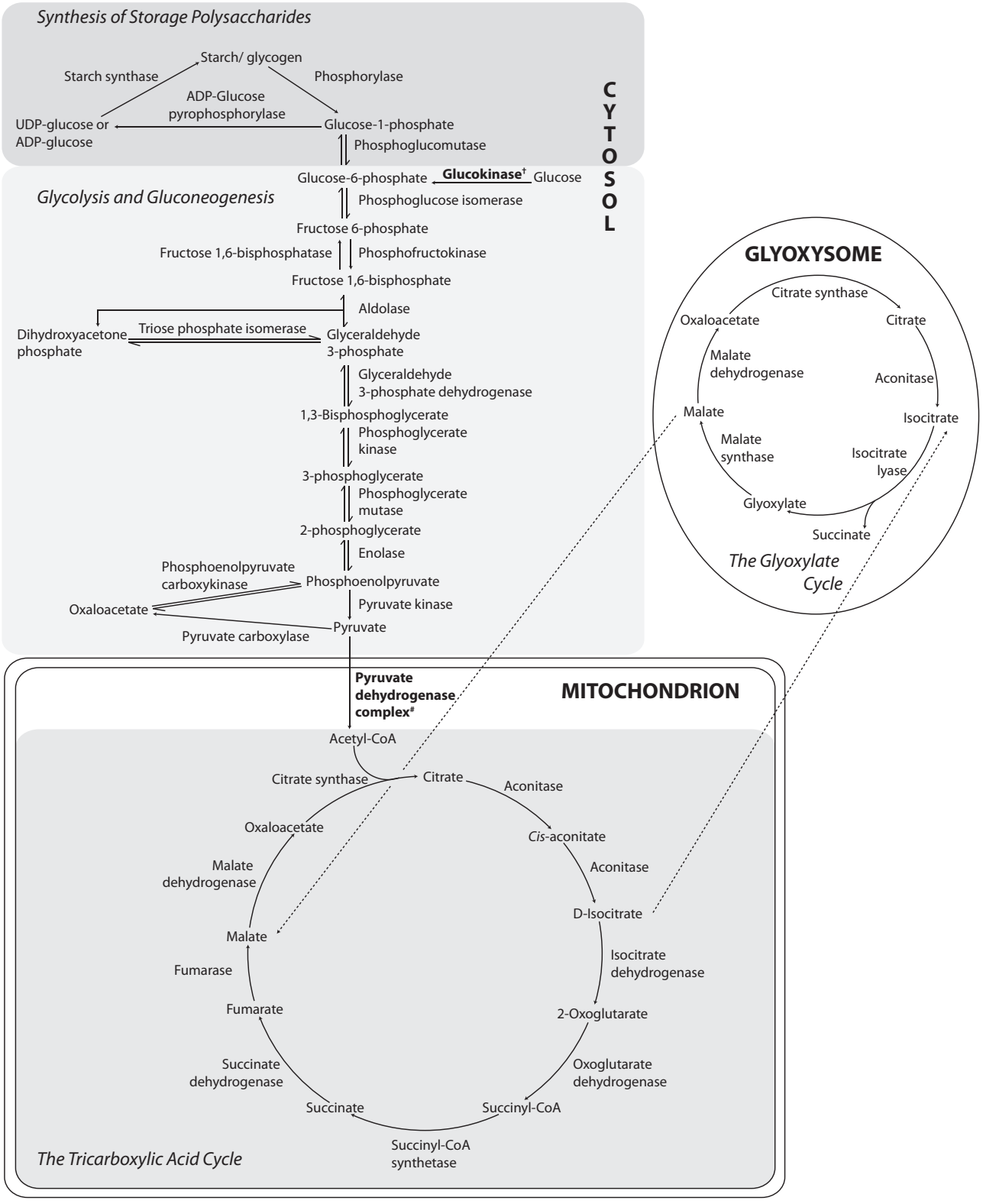
The irreversible conversion of pyruvate to acetyl-CoA is carried out by the pyruvate dehydrogenase complex (PDH) in aerobic mitochondria. Mitochondrial PDH consists of four types of subunits, E1α, E1β, E2 and E3, where E1α and E1β form a heterotetramer ($\alpha_2\beta_2$) and E3 is shared with the 2-oxoglutarate dehydrogenase complex. In some bacteria, such as *Corynebacterium* and many others, the heterotetrameric E1 complex is replaced by a homodimeric E1 (α_2) complex sharing no sequence similarity (Schreiner et al. 2005). Although ciliates contain a conventional mitochondrial pyruvate dehydrogenase (Smith et al. 2007b), Apicomplexa do not, with the only PDH being located in the remnant chloroplast (Foth et al. 2005; Ralph 2005).

In order to determine if dinoflagellates possessed a PDH, the EST databases were searched for sequences encoding subunits of the *Plasmodium* apicoplast PDH, the *Tetrahymena* mitochondrial PDH and the diatom *Thalassiosira pseudonana* mitochondrial PDH (Armbrust et al. 2004; Foth et al. 2005; Smith et al. 2007b). Photosynthetic dinoflagellates contained transcripts encoding E2 and E3 proteins, but we found none encoding the conventional E1 proteins. Instead we were able to identify in photosynthetic dinoflagellates a transcript encoding the homodimeric E1 (α_2) subunit, similar to that found in *Corynebacterium* and related bacteria (Schreiner et al. 2005).

In contrast, *Perkinsus* and *Oxyrrhis* did not contain transcripts encoding E1 (of either form) or E2. Instead, transcripts encoding the enzyme pyruvate:NADP oxidoreductase, more commonly found in anaerobic mitochondria, were identified as previously reported for these organisms and also for *Cryptosporidium* (Hug et al. 2010). In general, anaerobic mitochondria utilize alternatives to PDH. Organisms usually contain either pyruvate:ferredoxin oxidoreductase (PFO), an O_2

Figure 3. Overview of proposed sugar metabolism in dinoflagellates. The dinoflagellates are able to store polysaccharides in the form of starch or glycogen (shaded dark grey) as shown through EST data. The dinoflagellates appear to be able to undergo glycolysis (shaded light grey). Although no hexokinase was identified, a glucokinase is present in *Perkinsus*. The pyruvate dehydrogenase complex contains a bacterial-type E1 subunit. The dinoflagellate TCA cycle (shaded medium grey) appears to be conventional in contrast to *Plasmodium*. EST data also supports the presence of a glyoxysome within the dinoflagellates. Key †: present in *Perkinsus* only, #: bacterial-type E1 subunit.

Please cite this article in press as: Butterfield ER, et al. An Analysis of Dinoflagellate Metabolism Using EST Data. Protist (2012), http://dx.doi.org/10.1016/j.protis.2012.09.001



Please cite this article in press as: Butterfield ER, et al. An Analysis of Dinoflagellate Metabolism Using EST Data. Protist (2012), http://dx.doi.org/10.1016/j.protis.2012.09.001

sensitive enzyme, or the related enzyme pyruvate:NADP oxidoreductase (PNO), a fusion of PFO and a NADPH cytochrome P450 reductase domain (Rotte et al. 2000; Stechmann et al. 2008). No other dinoflagellate species contained transcripts encoding PNO or PFO.

Phylogenetic analyses have previously shown that the presence of the PFO domain dates prior to the split of the dinoflagellates and Apicomplexa, as *Perkinsus, Oxyrrhis* and *Cryptosporidium* PFO domains group together to the exclusion of all other PFO sequences (Hug et al. 2010). It therefore seems likely that that the genes encoding a conventional pyruvate dehydrogenase E1 subunit were lost prior to the separation of the dinoflagellate and apicomplexan lineages. These were then replaced by either PNO (in the early-branching dinoflagellate and Apicomplexa lineages) or with a bacterial-type E1 (other dinoflagellate lineages). Thus, the history of the pyruvate dehydrogenase in dinoflagellates is extremely interesting, for it involves two separate gene transfer events, each solving the same issue (the absence of E1 enzyme) in a different way.

TCA cycle: The first step in the TCA cycle is the condensation of oxaloacetate and acetyl-CoA by citrate synthase to form citrate. Transcripts corresponding to citrate synthase were identified in the dinoflagellate EST databases. The second step is the conversion of citrate to isocitrate by aconitase, an Fe-S cluster protein. Most eukaryotes contain two forms of aconitase, one mitochondrial (involved in the TCA cycle) and a cytoplasmic version. *Plasmodium* and *Tetrahymena* contain a single aconitase (similar to the cytoplasmic version) dually targeted to the cytosol and mitochondrion (Smith et al. 2007b; van Dooren et al. 2006). ESTs for a single isoform of aconitase were identified in the dinoflagellates. However, in contrast to the situation with *Plasmodium* and *Tetrahymena*, it was more similar to the mitochondrial isoform. It is possible that the ancestral alveolate contained both types of aconitase and the dinoflagellates subsequently lost the cytosolic form, whereas the ciliates and Apicomplexa lost the mitochondrial form, retargeting the cytoplasmic form.

The third step in the TCA cycle is the decarboxylation of isocitrate to form 2-oxoglutarate by isocitrate dehydrogenase. Transcripts were found encoding an isocitrate dehydrogenase which appeared to be NADP-linked, based on sequence similarity to other isocitrate dehydrogenases, and targeted to the mitochondrion (Danne and Waller 2011). No evidence of transcripts for cytosolic or NAD-linked isoforms was found. The NADP-linked enzyme may function in the reverse direction (isocitrate synthesis) in mitochondria of other organisms, although roles in the direction of isocitrate oxidation have also been postulated (reviewed by Sazanov and Jackson 1994). It is possible that dinoflagellates possess an NAD-linked enzyme with transcripts at low abundance. Alternatively, (and assuming the TCA cycle functions in the usual direction in contrast to the situation in *Plasmodium* (Olszewski et al. 2010)) the NADP-linked enzyme might be used for isocitrate oxidation. The resulting NADPH could be used directly, or the reducing equivalents transferred to NAD by a transhydrogenase, for which ESTs were also found. Biochemical data on the NADPH/NADP ratio in dinoflagellate mitochondria would help to clarify the role of the NADP-linked isocitrate dehydrogenase. ESTs corresponding to an oxoglutarate translocator were identified in two *Alexandrium* species, indicating that 2-oxoglutarate produced by isocitrate oxidation could be exported from the mitochondrion if necessary. The fourth step, the conversion of 2-oxoglutarate to succinyl-CoA by oxoglutarate dehydrogenase is possible, as transcripts encoding this enzyme were also found. However, no transcripts encoding 2-oxoglutarate dehydrogenase were identified in *Perkinsus* (although they were present in *Oxyrrhis*).

Transcripts encoding the three subsequent enzymes from the TCA cycle, succinyl-CoA synthase, succinate dehydrogenase and fumarase were all identified in dinoflagellates. The final step in the TCA cycle is the conversion of malate to oxaloacetate by malate dehydrogenase. Most organisms contain two malate dehydrogenases, one in the cytosol and one in the mitochondrion (together with a third in the chloroplast of photosynthetic organisms). Searches revealed transcripts encoding a malate dehydrogenase in dinoflagellates, which is predicted to be mitochondrial (Danne and Waller 2011). *Plasmodium* contains a single malate dehydrogenase which does not localize to the mitochondrion (van Dooren et al. 2006), with a malate:quinone oxidoreductase enzyme operating instead (Uyemura et al. 2004). A malate:quinone oxidoreductase has previously been identified in dinoflagellates. However, it is likely to be cytosolic so it cannot play the same role as in *Plasmodium* (Nosenko and Bhattacharya 2007).

A glyoxysome: In addition to the mitochondrial TCA cycle, various eukaryotic organisms including plants and filamentous fungi contain a specialized organelle, the glyoxysome, which carries out the glyoxylate pathway allowing the TCA cycle to be partially bypassed. This pathway includes two key enzymes, isocitrate lyase and malate synthase.

Please cite this article in press as: Butterfield ER, et al. An Analysis of Dinoflagellate Metabolism Using EST Data. Protist (2012), http://dx.doi.org/10.1016/j.protis.2012.09.001

Transcripts encoding both enzymes were found in dinoflagellates. Any succinate produced could be converted to succinyl-CoA and thus re-enter the TCA cycle. A dinoflagellate glyoxysome has previously been suggested based on ultrastructural data (Bibby and Dodge 1973).

The 2-oxoglutarate bypass: There was no evidence of sequences encoding the E1 or E2 subunits of 2-oxoglutarate dehydrogenase in *Perkinsus*. However, a bypass of this step has recently been reported for some cyanobacteria (Zhang and Bryant 2011). The first step of the bypass involves 2-oxoglutarate decarboxylase, a TPP-containing enzyme homologous to acetolactate synthase. Searching *Perkinsus* ESTs with the predicted protein sequence from the cyanobacterium *Synechococcus* sp. PCC7002 identified a 2-oxoglutarate decarboxylase homologue. Searching the other dinoflagellate ESTs with the predicted cyanobacterial protein sequence revealed transcripts encoding acetolactate synthase, but no 2-oxoglutarate decarboxylase. Thus it appears that the *Perkinsus* ESTs include a sequence encoding 2-oxoglutarate decarboxylase, but the other dinoflagellate EST collections do not.

The second step in the bypass of 2-oxoglutarate dehydrogenase described by Zhang and Bryant (2011) is succinic semialdehyde dehydrogenase, which gives succinate as a product. Searching with the predicted sequence of this protein from *Synechococcus* sp. PCC7002 gave clear matches to transcripts encoding proteins from *Perkinsus* as well as photosynthetic dinoflagellates. Overall, it therefore seems that *Perkinsus* has the 2-oxoglutarate dehydrogenase bypass that allows the TCA cycle to function in the absence of 2-oxoglutarate dehydrogenase itself. The photosynthetic dinoflagellates, which appear to have a 2-oxoglutarate dehydrogenase (see above), lack the bypass.

Oxidative phosphorylation: NADH and $FADH_2$ from the TCA cycle are used as electron donors to the mitochondrial electron transport chain which drives ATP synthesis, pumps protons and maintains the mitochondrial proton motive force necessary to power the protein import apparatus. Five complexes are involved in oxidative phosphorylation.

Complex I: Complex I, or NADH dehydrogenase, is a large, multi-subunit enzyme which oxidizes NADH to NAD^+ and passes electrons through a series of iron-sulphur proteins to ubiquinone as well as acting as a proton pump. The number of proteins recognized as involved in the complex varies considerably according to species, from 14 or so in bacteria to 43 or so in mammalian cells and 49 in plants (Klodmann et al. 2010). We therefore searched for transcripts encoding the components of this multi-subunit enzyme. Transcripts encoding numerous Complex I subunits were identified, with results shown in Supplementary Table S1. This is a different range of subunits from that found in the diatom algae. It therefore seems likely that Complex I is present, although with significant differences in the polypeptide composition compared with other eukaryotic lineages.

However, some organisms do not contain a conventional Complex I at all. Instead, these organisms (including *Plasmodium*) contain a single subunit NADH dehydrogenase (Mogi and Kita 2010). We therefore searched for transcripts encoding the single subunit NADH dehydrogenase. Transcripts encoding a single subunit NADH dehydrogenase with similarity to that found in *Plasmodium* were identified. In addition, ESTs for the alternative oxidase, which catalyses direct electron transfer from ubiquinone to oxygen (Matus-Ortega et al. 2011), were found in *Perkinsus* and *Oxyrrhis* as well as the photosynthetic dinoflagellates.

Complex II: Complex II, succinate dehydrogenase (SDH), consists of two to four core proteins, plus various accessory proteins depending on species. Of the four core proteins, transcripts were identified which encode subunits SDH-A (succinate dehydrogenase flavoprotein) and SDH-B (succinate dehydrogenase (ubiquinone) iron-sulphur subunit), which are also found in Apicomplexa and plants. There was no evidence for a dinoflagellate SDH-C (succinate dehydrogenase cytochrome b_{560} subunit) or SDH-D (integral membrane protein). SDH-C has been identified in diatoms and animals, although not in Apicomplexa or plants, while SDH-D is apparently absent from plants, Apicomplexa and diatoms.

Ubiquinone: Coenzyme Q (ubiquinone) resides in the inner mitochondrial membrane, linking complexes I and II with complex III. It also participates in pyrimidine biosynthesis (Do et al. 2001; Macedo et al. 2002). A variety of enzymes are involved in coenzyme Q biosynthesis. Although we were able to identify transcripts encoding putative methyltransferases in dinoflagellates, as previously reported with other eukaryotic species, we were not able definitively to identify transcripts encoding enzymes catalysing all proposed steps (Kanehisa and Goto 2000; Kanehisa et al. 2012; van Dooren et al. 2006).

Complex III: The dinoflagellate complex III (the cytochrome bc_1 complex) has previously been shown to be functional (Kamikawa et al. 2009). We were able to identify transcripts encoding

Please cite this article in press as: Butterfield ER, et al. An Analysis of Dinoflagellate Metabolism Using EST Data. Protist (2012), http://dx.doi.org/10.1016/j.protis.2012.09.001

four proteins involved in complex III: a ubiquinol-cytochrome *c* reductase iron-sulphur subunit, cytochrome *b* (encoded on the mitochondrial genome), ubiquinol-cytochrome *c* reductase and the cytochrome c_1 subunit. In addition, there was evidence for a possible QCR6 ubiquinol-cytochrome *c* reductase hinge protein. The remaining proteins involved in complex III in animal mitochondria are not found in plants, diatom algae or Apicomplexa. The sole exception is QCR9, which is found in plants. However, no transcripts encoding this protein were identified in dinoflagellates.

Complex IV: Complex IV, also known as cytochrome *c* oxidase, consists of 14 subunits in animals, compared with just 9 subunits reported from the apicomplexan *Plasmodium*. We identified eight subunits: cox1, cox2, cox2A, cox3, cox5b, cox6b, cox10, cox11, together with possible cox61, cox15 and cox17 (cytochrome *c* oxidase assembly protein subunit 17, a very small protein that has not yet been identified in diatoms). It should be noted that Complex IV assembly protein subunits 10 and 11 (cox10 and cox11) were identified only in *Perkinsus* and not in other dinoflagellate species.

Complex V: The final step in the electron transport chain is the synthesis of ATP by complex V. This is a multi-subunit complex, composed of around 20 or so proteins. ESTs for the key catalytic proteins (α and β) were identified, as well as the membrane-bound c subunit and OSCP, δ, ε and γ. However, ESTs for the majority of the stator, which holds the complex together, were missing. As the same set of genes are missing from all alveolate genomes sequenced to date, it would appear that the alveolates have either a novel or a highly divergent set of proteins making up the stator complex. Recently, a novel complex V structure has been proposed for the ciliate *Tetrahymena thermophila*, including a number of newly described proteins (Nina et al. 2010). However, none of these proteins has been identified in either dinoflagellates or Apicomplexa. It is thus unclear how the complex functions.

Haem lyase: One key process in the assembly of the multi-subunit electron transport chain is the attachment of haem to apocytochrome *c*. At least three different methods for this are used, Systems I, II and V, with each species containing just one System (Allen 2011; Stevens et al. 2011). System I (Ccm) is found in plant mitochondria, archaea and many Gram negative bacteria. No evidence for this System was found in dinoflagellates. System III consists of a haem lyase (cytchrome *c* haem lyase), and is the system found in fungi and animals, as well as the apicomplexan *Plasmodium.* Although we identified a transcript corresponding to haem lyase from *Perkinsus*, consistent with a previous observation based on gene sequence data (Allen et al. 2008b), no such transcripts were identified for other dinoflagellates. The presence of a System V has been postulated in Euglenozoa for incorporation of haem into cytochromes where the haem-binding motif comprises single cysteines, but the proteins involved have not yet been identified (Allen 2011). Allen et al. (2008b) argue that haem lyase is generally used in organisms that have no mitochondrially-encoded System I components, and *c*-type cytochromes with conventional haem-binding motifs containing two cysteines. Both the latter criteria are seen with dinoflagellates (Nash et al. 2008). Given this, and the indications of haem lyase in *Perkinsus* and Apicomplexa, the most parsimonious explanation is that dinoflagellates also use a haem lyase, represented by transcripts that were too rare to be included in the EST databases. However, we cannot exclude the possibility of a different system acquired by lateral gene transfer.

(Two further mechanisms, Systems II and IV, are normally used for the insertion of haem into chloroplast proteins. We found evidence for sequences encoding components of System IV. We found evidence for sequences encoding System II components in the chloroplast genomes of the dinoflagellates *Durinskia* and *Kryptoperidinium*, which contain essentially intact diatom endosymbionts (Imanian et al. 2010), but not in other dinoflagellates. The significance of this for cytochrome maturation in chloroplasts is unclear.)

Synthesis of Co-enzymes

In addition to providing cellular energy, mitochondria play a key role in the synthesis of various co-enzymes. The biosynthesis of these compounds is complex, and often only one or two steps occur in the mitochondrion (Rébeillé et al. 2007). Products are exported from the mitochondrion for use throughout the cell.

Iron sulphur cluster biosynthesis: The biosynthesis of iron-sulphur (Fe-S) clusters is an essential biochemical pathway that has most likely existed since the origin of life (Seeber 2002). Three different iron sulphur cluster biosynthesis pathways exist in eukaryotic cells - one for each of the chloroplast, mitochondrion and the cytosol (Balk and Pilon 2011). The SUF (sulphur mobilization) pathway is found in chloroplasts, the ISC (iron sulphur cluster) pathway is found in mitochondria and the CIA (cytosolic iron-sulphur) pathway is found in the

Please cite this article in press as: Butterfield ER, et al. An Analysis of Dinoflagellate Metabolism Using EST Data. Protist (2012), http://dx.doi.org/10.1016/j.protis.2012.09.001

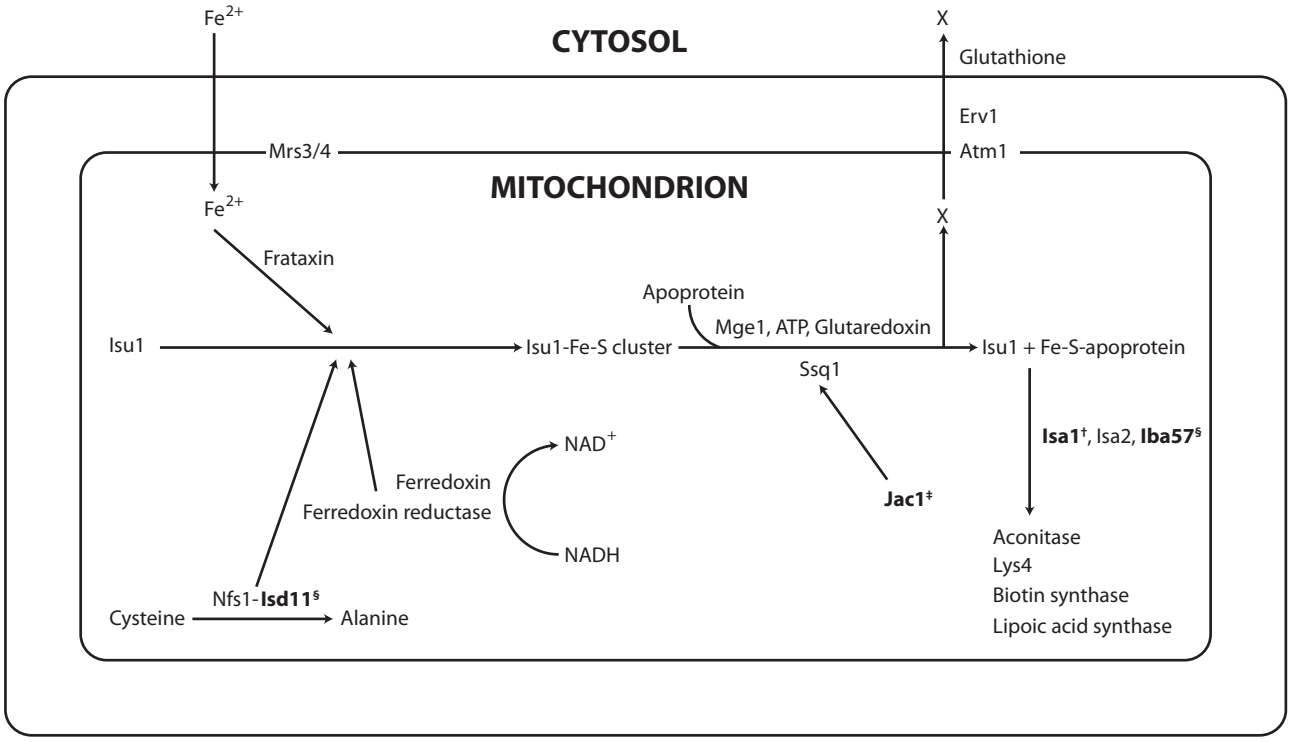


Figure 4. Fe-S cluster biosynthesis. EST data suggest dinoflagellate mitochondria are able to produce Fe-S clusters. Key §: no transcripts identified, †: present in *Perkinsus* only, ‡: present in *Perkinsus* and *Oxyrrhis* only.

cytosol. In each case, the synthesis of Fe-S clusters occurs in two distinct steps: the co-ordination of iron and sulphur onto protein backbones using cysteine linkages followed by the transfer of the cluster onto apoproteins (Fig. 4).

Iron is imported into the mitochondrion using the homologous transporters Mrs3 and Mrs4, and is then donated to the Isu1 scaffold protein via frataxin. Although we were able to identify transcripts encoding mitochondrial carrier proteins in dinoflagellates, we were unable to identify them specifically as Mrs3 or Mrs4. Transcripts encoding frataxin and Isu1 were identified. Nfs1 then forms a complex with the Isd11 protein, and catalyses the release of sulphide from cysteine (Adam et al. 2006; Balk and Pilon 2011). We identified ESTs corresponding to transcripts for Nfs1 but not for Isd11. Although Isd11 is found in all other eukaryotes, and appears to be essential, it is missing from bacteria (Lill and Mühlenhoff 2005; Richards and van der Giezen 2006), which may suggest that the dinoflagellate process is more similar to that present in bacteria, or that the transcript is missing from the dataset.

An electron transport chain is also required for formation of the Fe-S clusters on the Isu1 which consists of a ferredoxin and ferredoxin reductase (Lill and Mühlenhoff 2005). We identified transcripts for ferredoxin and ferredoxin reductase. However, localization studies are required to demonstrate that the ferredoxin reductase is present in the mitochondrion, since the translated transcripts show greater similarity to chloroplast ferredoxin reductase sequences than to mitochondrial sequences (as shown in Supplementary Table S2).

Finally, the Fe-S clusters are transferred to apoproteins. The hsp70-type chaperone protein Ssq1 (in yeast; HscA in plants) is activated by a co-chaperone Jac1 (in yeast; HscB in plants) which releases the cluster from Isu1 (which is then reused). Several further proteins are involved with the attachment of the cluster to the apoprotein, including Isa1, Isa2, Iba57 and glutaredoxin, although their exact role has not been characterized in algae (Balk and Pilon 2011).

Although we were able to identify transcripts encoding hsp70, it was unclear if these were the mitochondrial Ssq1 or a cytosolic hsp70. Transcripts encoding Jac1 were identified only in *Perkinsus* and *Oxyrrhis*. A transcript encoding a putative glutaredoxin was identified, as were transcripts encoding Isa2. However, transcripts encoding Isa1 were only identified in *Perkinsus.* We were unable to identify transcripts for Iba57 in

the dinoflagellates (Balk and Pilon 2011; Lill and Mühlenhoff 2005, 2008).

In addition to the assembly of iron-sulphur clusters for use within the organelle, the mitochondrion is also responsible for the first steps in cytosolic iron-sulphur cluster assembly. This process involves the transporters Atm1 (an ATP-binding cassette transporter) and Erv1 (sulphhydryl oxidase) (Balk and Pilon 2011). We identified transcripts encoding both Atm1 and Erv1 in dinoflagellates.

Haem biosynthesis: It has been reported recently that both *Chromera* (a photosynthetic apicomplexan) and non-photosynthetic Apicomplexa contain similar pathways for tetrapyrrole biosynthesis (Kořený et al. 2011). Kořený et al. showed that aminolaevulinic acid was synthesized from the 'Shemin' pathway, involving succinyl-CoA and glycine as precursors, in the mitochondria of *Chromera* as in the other Apicomplexa. Subsequent steps in the synthesis of protohaem and chlorophyll were inferred as chloroplast located in *Chromera*, with synthesis as far as uroporphyrinogen or coproporphyrinogen located in the remnant chloroplast of *Toxoplasma* and *Plasmodium* respectively. Kořený et al. (2011) also inferred an ALA synthase using succinyl-CoA and glycine in the mitochondrion of *Perkinsus* and *Oxyrrhis* (Fig. 5A). However, we found no evidence of ESTs encoding such an ALA synthase in the photosynthetic dinoflagellates.

The alternative pathway for ALA synthesis, as carried out by most other photosynthetic eukaryotes, uses glutamyl-tRNA and occurs in the chloroplast (Oborník and Green 2005). Surprisingly, we found no clear evidence for ESTs encoding the enzyme for the first committed step, glutamyl-tRNA reductase. It should be noted that there was a match to a single EST from *Lingulodinium*. A GC % comparison and BLAST analysis reveals that this transcript is most likely a bacterial contamination. However, strong matches were seen when searching for ESTs encoding the enzyme for the next step, glutamate semialdehyde aminomutase. Transcripts encoding enzymes for all remaining steps in the biosynthesis pathway until the divergence into haem and chlorophyll branches were identified (Fig. 5B), including ALA dehydratase, porphobilinogen deaminase, uroporphyrinogen synthase, uroporphyrinogen decarboxylase, coproporphyrinogen oxidase, protoporphyrinogen oxidase, ferrochelatase, magnesium chelatase subunit H, magnesium chelatase subunit I (which may also take over the role of the partially homologous subunit D in dinoflagellates, as no EST corresponding specifically to subunit D was found). Given that transcripts encoding all remaining enzymes in the pathway are present, it seems likely that glutamyl-tRNA reductase is present, and is just not represented in the EST datasets.

When these sequences were used to search the protein databases excluding dinoflagellates the best hits were generally to diatom sequences. The protoporphyrinogen oxidase was an exception to this. When the dinoflagellate *Crypthecodinium cohnii* protoporphyrinogen oxidase homologue was used to search the databases excluding dinoflagellates, the best hits were to plant or green algal sequences. When the dinoflagellate *Karenia brevis* sequence was used, the best hit was to the soil bacterium *Paenibacillus terrae*. However, alignments (Supplementary Table S2) were inconclusive due to the high rate of evolution of this protein. The ferrochelatase enzyme was most similar to those found in Apicomplexa and red algae rather than the form found in green plants and animals.

Dihydroorotate dehydrogenase: Dihydroorotate dehydrogenase is involved in pyrimidine biosynthesis, as well as being linked to the electron transport chain where it reduces coenzyme Q to ubiquinone (Krungkrai 1995). An EST encoding this enzyme was found in the dinoflagellates.

Folate biosynthesis: The biosynthesis of folate in plants and algae involves enzymes located in the chloroplast, mitochondrion and the cytosol (Fig. 6) (Rébeillé et al. 2007; Smith et al. 2007a). Initially, pABA is synthesized in the chloroplast, and pterin is synthesized in the cytosol. Both are imported to the mitochondrion. Here, the pterin is activated and linked to the pABA by a bifunctional enzyme, HPPK-DHPS (6-hydroxymethyldihydroprotein pyrophosphokinase – dihydropteroate synthase) forming dihydropteroate. Glutamate is then attached to dihydropteroate by the enzyme DHFS (dihydrofolate synthetase). Transcripts encoding a potential HPPK-DHPS were identified in dinoflagellates, but no obvious DHFS transcript was identified. In plants, the next step is catalyzed by DHFR-TS (dihydrofolate reductase-thymidylate synthase), a bifunctional enzyme. Transcripts encoding a well-conserved DHFR-TS were found in dinoflagellates. At this point, the products (H_4F) can either be exported to the cytoplasm or be retained in the mitochondrion. In both the cytoplasm and the mitochondrion the next step is polyglutamation. This is carried out by folylpolyglutamate synthetase (FPGS). No transcript encoding FPGS was identified in the dinoflagellate EST libraries, perhaps

Please cite this article in press as: Butterfield ER, et al. An Analysis of Dinoflagellate Metabolism Using EST Data. Protist (2012), http://dx.doi.org/10.1016/j.protis.2012.09.001

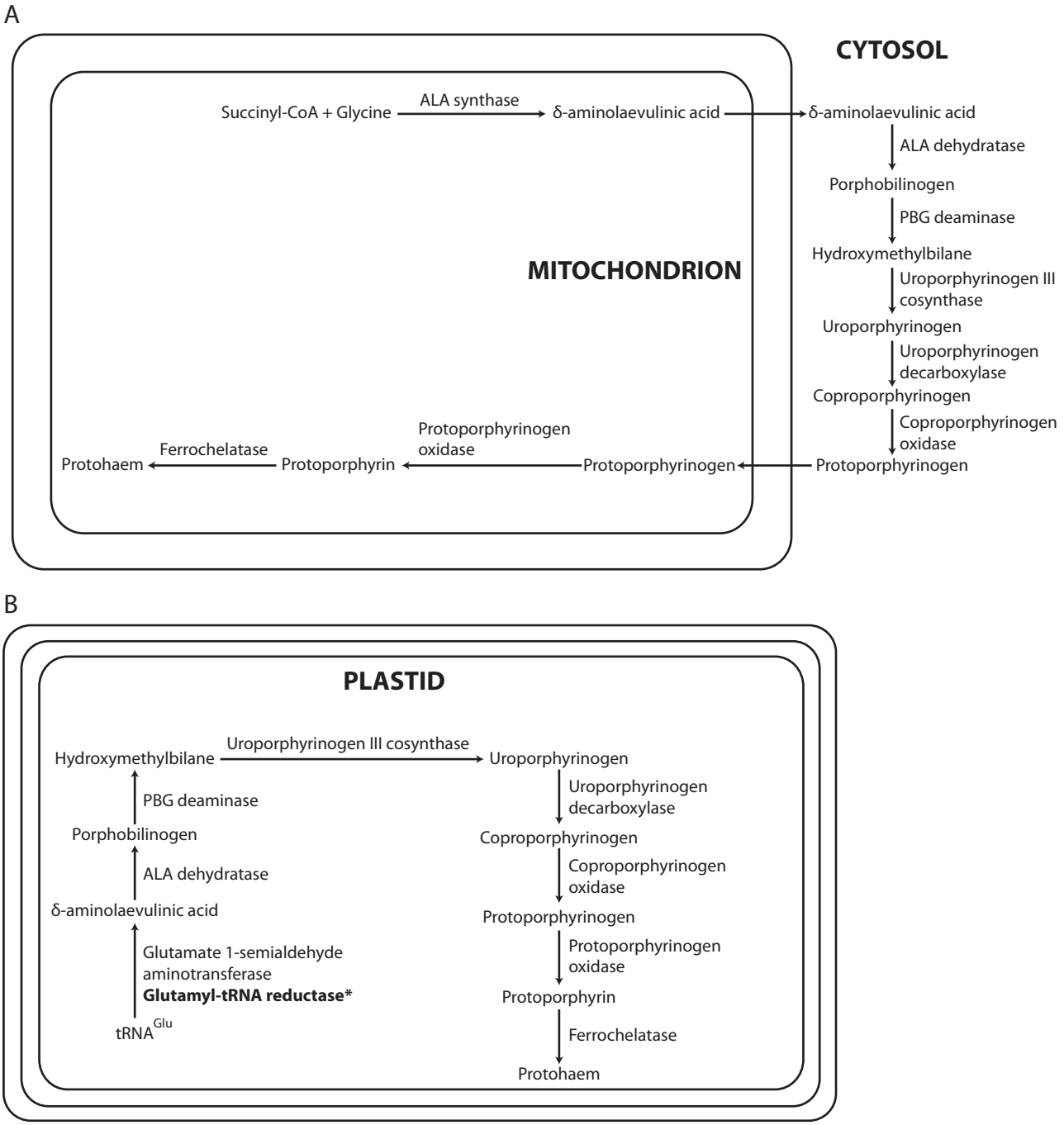


Figure 5. Haem biosynthesis pathways. **A**. Haem biosynthesis pathway used by *Perkinsus marinus* and *Oxyrrhis marina* (Kořený et al. 2011). **B**. Photosynthetic dinoflagellate haem biosynthesis pathway. Key *: possible EST contamination.

because such mRNA transcripts are at a very low abundance.

Biotin biosynthesis: Biotin is an essential co-factor in several enzymes. There is considerable diversity across the eukaryotes, with many species being auxotrophic for biotin. This diversity includes the dinoflagellates; a recent survey showed that 7/27 species examined were auxotrophic (Croft et al. 2006). Biosynthesis of biotin is a four-step process, of which the final two steps occur in the mitochondrion (Fig. 7) (Pinon et al. 2005; Rébeillé et al. 2007). The first step occurs in the cytosol. This

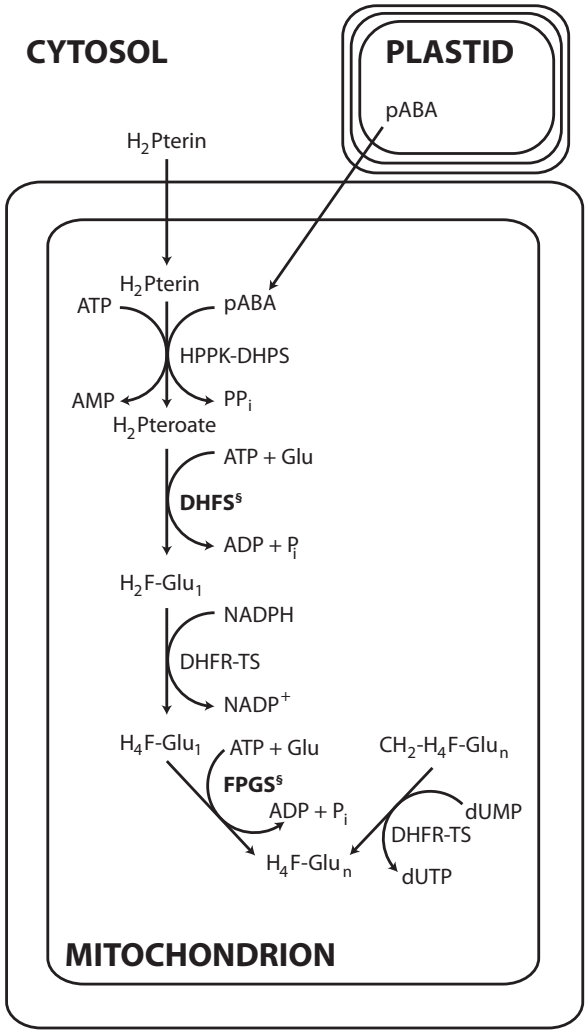


Figure 6. Folate biosynthesis. EST evidence was inconclusive as to the presence of the folate biosynthesis pathway within the dinoflagellates due to the inability to identify DHFS and FPGS. Key §: no transcripts identified, pABA: para-aminobenzoate, H_2Pterin: 6-hydroxymethyldihydropterin, HPPK-DHPS: 6-hydroxymethyldihydroprotein pyrophosphokinase – dihydropteroate synthase, PP_i: pyrophosphate, H_2Pteroate: dihydropteroate, DHFS: dihydrofolate synthase, P_i: phosphate, $H_2F\text{-}Glu_1$: monoglutamated form of dihydrofolate, DHFR-TS: dihydrofolate reductase – thymidylate synthase, $H_4F\text{-}Glu_1$: monoglutamated form of tetrahydrofolate, FPGS: folylpolyglutamate synthetase, $H_4F\text{-}Glu_n$: polyglutamated form of tetrahydrofolate, $CH_2F\text{-}H_4F\text{-}Glu_n$: polyglutamated form of methylene-tetrahydrofolate.

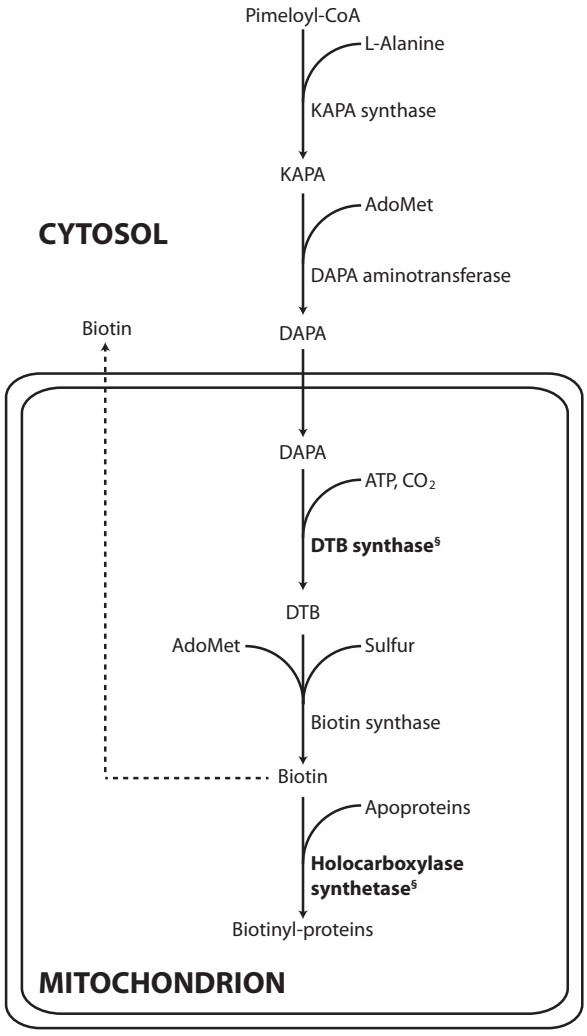


Figure 7. Biotin biosynthesis. The dinoflagellates appear to be able to synthesize biotin. Key §: No transcripts identified, KAPA synthase: 7-keto-8-aminopelargonic acid synthase, KAPA: 7-keto-8-aminopelargonic acid, AdoMet: S-adenosylmethionine, DAPA synthase: 7,8-diaminopelargonic acid, DAPA: 7,8-diaminopelargonic acid, DTB synthase: dethiobiotin synthase, DTB: dethiobiotin.

is the condensation of pimeloyl-CoA and L-alanine, carried out by KAPA synthase, which is encoded by the *bioF* gene. Transcripts encoding this enzyme were identified in *Alexandrium tamarense*, but not in *Oxyrrhis*. The absence from *Oxyrrhis* is not surprising, since it is auxotrophic for biotin (Croft et al. 2006). The second step, the conversion of KAPA to DAPA (7,8 diaminopelargonic acid) is carried out by DAPA aminotransferase (encoded by *bioA*). A putative DAPA aminotransferase was identified.

Please cite this article in press as: Butterfield ER, et al. An Analysis of Dinoflagellate Metabolism Using EST Data. Protist (2012), http://dx.doi.org/10.1016/j.protis.2012.09.001

No transcripts were identified encoding the third enzyme, DTB synthase (encoded by *bioD*). This is perhaps not surprising, as no *bioD* gene has been identified in any algal species, *Plasmodium*, or the higher plant *Arabidopsis thaliana*. It is assumed that this step is carried out by an as yet unidentified enzyme, present across all species missing *bioD* (Croft et al. 2006). The final step is carried out by biotin synthase *(bioB)*, and transcripts encoding this protein were found in photosynthetic dinoflagellates. Biotin is used in the biotinylation of proteins by the enzyme holocarboxylase synthetase, also known as biotin protein ligase. This enzyme is found in the mitochondrion, chloroplast and cytosol, allowing proteins throughout the cell to be biotinylated (Rébeillé et al. 2007). No transcript encoding holocarboxylase synthetase was identified though it seems unlikely that this enzyme is missing from the cell. Perhaps transcripts are only present at low levels and are thus not found in the EST datasets.

Protein Import

As only three proteins are encoded on the dinoflagellate mitochondrial genome, the vast majority of proteins must be imported. Many proteins are involved in the protein translocation process, and the number utilized depends on how far into the mitochondrion the imported protein is required to travel (Fig. 8). Although much is known about mitochondrial protein import in model species, there is significant diversity across the protists, and protein sequences are divergent (e.g. Rada et al. 2011). Imported proteins initially pass through the translocase of the outer membrane (TOM) complex. The TOM complex is composed of multiple subunits. Of these, only transcripts encoding a possible Tom40 and a possible Tom70 were identified in the dinoflagellate EST dataset. It has recently been shown that Tom40 and Tom70 are the only TOM complex proteins found in Microsporidia (Heinz and Lithgow 2012), suggesting that this may be the minimum set of TOM proteins required for successful protein import.

Once proteins have crossed the outer membrane, subsequent sorting occurs. Proteins destined for the matrix cross the inner membrane using the Tim23 complex. We were able to identify transcripts for a reduced set of subunits (Tim50 (possible), Tim23, Tim17 (*Oxyrrhis* only), Tim44, Hsp70, Mge1 and Pam18). We were unable to identify transcripts for Tim21, Pam16 (*Oxyrrhis* only) and Pam17. There is no identified Tim21 homologue in *Plasmodium falciparum* (van Dooren et al. 2006) or *Tetrahymena thermophila* (Smith et al. 2007b). Pam17 is also absent from plants and humans (Carrie et al. 2010b), and Pam16 is absent from many species. Therefore, the absence of these proteins in photosynthetic dinoflagellates is not inconsistent with functional protein import (Heinz and Lithgow 2012).

Proteins which are destined for the inner membrane are inserted by one of two complexes. The Tim22 complex identifies proteins with an internal targeting sequence for the inner mitochondrial membrane. This complex consists of Tim18, Tim22 and Tim54. Although ESTs encoding Tim22 and a possible Tim18 were present, Tim54 appeared to be absent. Tim54 has also not been identified in *Arabidopsis thaliana* (Lister et al. 2004) which would suggest that this protein is not essential. The remaining proteins with an N-terminal targeting sequence are inserted into the mitochondrial membrane by the Oxa1 protein, transcripts for which were present.

Those proteins which ultimately end up in the outer membrane are guided by the tiny Tims to the SAM complex (Sorting and Assembly Machinery, also called the TOB (topogenesis of outer membrane β-barrel proteins) complex) (Dolezal et al. 2006; Dukanovic and Rapaport 2011; Wiedemann et al. 2004). Of the tiny Tims, only transcripts encoding Tim10 were identified, together with a possible Tim8 in photosynthetic dinoflagellates. The numbers of tiny Tims across the eukaryotes varies, with most species containing homohexamers (Tim8 and 13 or Tim9 and 10), although *Cryptosporidium* has a single tiny Tim and microsporidia have lost the complex entirely (Heinz and Lithgow 2012). Tiny Tims are imported into the mitochondrion using MIA (mitochondrial intermembrane space assembly machinery) which is composed of the Mia40 and Erv1 proteins (Carrie et al. 2010a). We identified transcripts for Erv1 but were unable to identify transcripts for Mia40 (Allen et al. 2008a). Assuming the tiny Tims are indeed located in the intermembrane space, it is unclear how this is achieved.

Once proteins are imported through the outer mitochondrial membrane, beta-barrel proteins destined to the outer mitochondrial membrane are inserted by the SAM complex. This complex is composed of three proteins Sam50/Tob55/Omp85, Sam37/Mas37/Tom37 and Sam35/Tob38/Tom38 together with Mdm10 and Mim1 (Dukanovic and Rapaport 2011). However, Sam50 is the only protein conserved across all eukaryotic species (with the exception of *Giardia*) (Heinz and Lithgow 2012). We were unable to identify transcripts for any of these proteins in the photosynthetic dinoflagellates.

Please cite this article in press as: Butterfield ER, et al. An Analysis of Dinoflagellate Metabolism Using EST Data. Protist (2012), http://dx.doi.org/10.1016/j.protis.2012.09.001

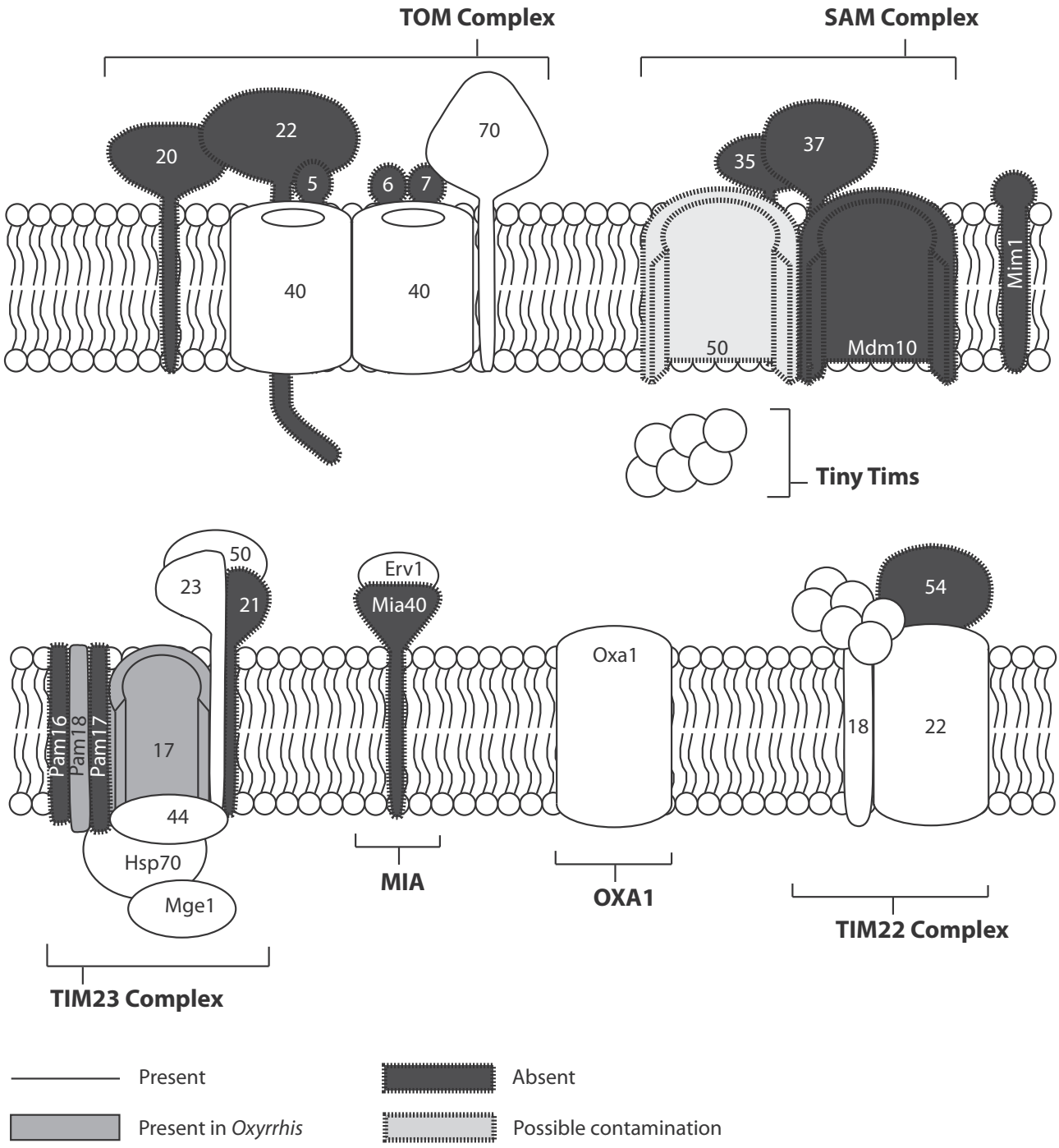


Figure 8. Protein import in the dinoflagellates. EST data suggest that the dinoflagellates contain the core subunits of all import pathways with the exception of the SAM complex (the subunit Sam50 may be bacterial contamination of the EST library). Adapted from Schleiff and Becker (2011). Key TOM complex: translocase of the outer membrane complex, SAM complex: sorting and assembly machinery (also called TOB complex: topogenesis of outer membrane β-barrel proteins), TIM23 complex: translocase of the inner membrane 23 complex, MIA: mitochondrial intermembrane space assembly machinery, TIM22 complex: translocase of the inner membrane 22 complex.

Please cite this article in press as: Butterfield ER, et al. An Analysis of Dinoflagellate Metabolism Using EST Data. Protist (2012), http://dx.doi.org/10.1016/j.protis.2012.09.001

We did identify a transcript for YaeT, an *Escherichia coli* protein homologous to Sam50 involved in insertion of proteins in the outer membrane (Paschen et al. 2005; Stegmeier and Andersen 2006; Werner and Misra 2005). However, due to the identical protein alignment with *E. coli* (as shown in Supplementary Table S2) and low GC % we cannot rule out the possibility of contamination during the creation of the EST library. It is therefore unclear how β-barrel proteins are inserted into the outer mitochondrial membrane.

A recent paper has shown that the early-branching dinoflagellate *Perkinsus* has a greater range of proteins involved in mitochondrial protein import than photosynthetic dinoflagellates, with the identification of Pam16 and all four tiny Tims, suggesting that these proteins have been lost in photosynthetic dinoflagellates (Alcock et al. 2012).

Once proteins have been imported into the mitochondrion, processing occurs to remove the targeting peptide. Transcripts encoding both the mitochondrial processing peptidase (MPP, both subunits) and the mitochondrial intermediate peptidase (MIP) are present. However, no transcript was identified which encoded the inner membrane protease (IMP), responsible for a second cleavage after MPP (Gakh et al. 2002).

Conclusion

Since the partial sequencing of the dinoflagellate mitochondrial genome in 2007, (Jackson et al. 2007; Nash et al. 2007; Slamovits et al. 2007) it has been clear that the majority of mitochondrial proteins are encoded in the nucleus and imported into the organelle. Here, we have examined the EST datasets from a variety of dinoflagellate algae and show that although some features of mitochondrial biochemistry are conventional, there are numerous unexpected features.

Although glycolytic and gluconeogenic pathways are present, as is the TCA cycle, the pyruvate dehydrogenase complex is unconventional. It would appear that the classical E1 subunit has been replaced in photosynthetic dinoflagellates by a homodimeric E1 protein as found in some bacteria (as shown in Supplementary Table S2). Meanwhile, early branching lineages of dinoflagellates replaced PDH with PNO, usually found in anaerobic mitochondria. Given that Apicomplexa lack a mitochondrial PDH, it would appear that this gene loss/replacement event occurred early on in the alveolate lineage. A second lateral gene transfer event appears to have occurred within the haem biosynthesis pathway, where the protoporphyrinogen oxidase enzyme appears to have been replaced.

A surprising discovery was the apparent absence of 2-oxoglutarate dehydrogenase from *Perkinsus*, associated with the acquisition of a bypass pathway. This bypass pathway, recently identified in cyanobacteria (Zhang and Bryant 2011) appears to have been obtained by *Perkinsus* through lateral gene transfer. The acquisition of genes from other species is typical of dinoflagellates, so is perhaps not unexpected (Nosenko and Bhattacharya 2007; Waller et al. 2006). It is unclear whether the acquisition of these new genes was prior to the loss of the original, dinoflagellate gene, or whether the lateral gene transfer event preceded the loss of the original endogenous gene.

The discovery of a glyoxysome pathway in dinoflagellates confirmed electron micrograph reports from the 1970s (Bibby and Dodge 1973). This pathway is normally found in plants, allowing the TCA cycle to be partially bypassed. Searches showed that the pathway was not present in Apicomplexa. The most likely explanation is that the pathway was present in the ancestor of dinoflagellates and Apicomplexa, and lost in the latter lineage. It will be interesting to see if it is present in the photosynthetic apicomplexans such as *Chromera* An examination of the photosynthetic Apicomplexa *Chromera* and *Vitrella* genomes ought to reveal which scenario is more likely.

The variation in mitochondrial biochemistry within the Apicomplexa (Olszewski et al. 2010), the highly divergent ATP synthase in ciliates (Nina et al. 2010) and the data presented here show that metabolic biochemistry in the alveolates is extremely diverse. It would appear that that these organisms are very good at adapting to many conditions, and the evolution of the mitochondrion plays a significant role in this adaptation. Far from being relatively homogeneous, eukaryotic mitochondria are extremely diverse, and the dinoflagellate mitochondria are no exception.

Methods

Identification of ESTs for proteins putatively targeted to the mitochondria: Translated dinoflagellate EST (expressed sequence tag) libraries were searched using specific protein sequences using the NCBI tBLASTn tool. Translated dinoflagellate EST sequences were obtained using the NCBI tblastn software by selecting the EST dataset and setting organism option to 'dinoflagellate'. Species included *Alexandrium minutum* (Yang et al. 2010), *Amphidinium carterae* (Bachvaroff and Place 2008), *Crypthecodinium cohnii* (Sanchez-Puerta et al.

Please cite this article in press as: Butterfield ER, et al. An Analysis of Dinoflagellate Metabolism Using EST Data. Protist (2012), http://dx.doi.org/10.1016/j.protis.2012.09.001

2007), *Heterocapsa triquetra* (Patron et al. 2005*, Karenia brevis* (Yoon et al. 2005)*, Karlodinium micrum* (Patron et al. 2006)*, Lingulodinium polyedrum* (Tanikawa et al. 2004) and *Symbiodinium.* In addition, '*Perkinsus*' (Joseph et al. 2010) and *'Oxyrrhis'* (Slamovits and Keeling 2008) were also searched where necessary. Full details of each species, together with the relative size of each EST dataset are given in Supplementary Table S3. Dinoflagellate sequences were then used to search back against the whole NCBI protein database to confirm identity where necessary. Returned sequences were placed within their metabolic families in order to determine the presence or absence of each metabolic pathway. Although EST sequences did not allow the identification of targeting sequences, which are in any case not well defined for dinoflagellate mitochondria, a mitochondrial location was generally inferred where this was consistent with the location of equivalent proteins in other organisms.

Acknowledgements

We thank Dr Giselle Walker (University of Otago, New Zealand) for line drawings of alveolate species. We thank Dr Edmund Nash who started this project whilst holding a Biotechnology and Biological Sciences Research Council studentship. ERB is funded by an Australian Postgraduate Award, CJH by a University of South Australia Distinguished Researcher Award and RERN by a University of South Australia Divisional Research Grant.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.protis.2012.09.001.

References

Adam AC, Bornhovd C, Prokisch H, Neupert W, Hell K (2006) The Nfs1 interacting protein Isd11 has an essential role in Fe/S cluster biogenesis in mitochondria. EMBO J **25**:174–183

Alcock F, Webb CT, Dolezal P, Hewitt V, Shingu-Vasquez M, Likic VA, Traven A, Lithgow T (2012) A small Tim homohexamer in the relic mitochondrion of *Cryptosporidium*. Mol Biol Evol **29**:113–122

Allen JW (2011) Cytochrome c biogenesis in mitochondria - Systems III and V. FEBS J **278**:4198–4216

Allen JW, Ferguson SJ, Ginger ML (2008a) Distinctive biochemistry in the trypanosome mitochondrial intermembrane space suggests a model for stepwise evolution of the MIA pathway for import of cysteine-rich proteins. FEBS Letters **582**:2817–2825

Allen JW, Jackson AP, Rigden DJ, Willis AC, Ferguson SJ, Ginger ML (2008b) Order within a mosaic distribution of mitochondrial c-type cytochrome biogenesis systems? FEBS J **275**:2385–2402

Archibald JM (2008) Plastid evolution: remnant algal genes in ciliates. Curr Biol **18**:R663–R665

Armbrust EV, Berges JA, Bowler C. et al. (45 co-authors) (2004) The genome of the diatom *Thalassiosira pseudonana*: Ecology, evolution, and metabolism. Science **306**:79–86

Bachvaroff TR, Place AR (2008) From stop to start: tandem gene arrangement, copy number and *trans*-splicing sites in the dinoflagellate *Amphidinium carterae*. PLoS ONE **3**:e2929

Bachvaroff TR, Concepcion GT, Rogers CR, Herman EM, Delwiche CF (2004) Dinoflagellate expressed sequence tag data indicate massive transfer of chloroplast genes to the nuclear genome. Protist **155**:65–78

Balk J, Pilon M (2011) Ancient and essential: the assembly of iron-sulfur clusters in plants. Trends Plant Sci **16**:218–226

Bibby B, Dodge J (1973) The ultrastructure and cytochemistry of microbodies in dinoflagellates. Planta **112**:7–16

Bowman I, Grant P, Kermack W, Ogston D (1961) The metabolism of *Plasmodium berghei*, the malaria parasite of rodents. 2. An effect of mepacrine on the metabolism of glucose by the parasite separated from its host cell. Biochem J **78**:472–478

Cáceres AJ, Quiñones W, Gualdrón M, Cordeiro A, Avilán L, Michels PAM, Concepción JL (2007) Molecular and biochemical characterization of novel glucokinases from *Trypanosoma cruzi* and *Leishmania* spp. Mol Biochem Parasit **156**:235–245

Carrie C, Murcha M, Whelan J (2010b) An *in silico* analysis of the mitochondrial protein import apparatus of plants. BMC Plant Biol **10**:249

Carrie C, Giraud E, Duncan O, Xu L, Wang Y, Huang S, Clifton R, Murcha M, Filipovska A, Rackham O (2010a) Conserved and novel functions for *Arabidopsis thaliana* MIA40 in assembly of proteins in mitochondria and peroxisomes. J Biol Chem **285**:36138–36148

Croft MT, Warren MJ, Smith AG (2006) Algae need their vitamins. Eukaryot Cell **5**:1175–1183

Danne JC, Waller RF (2011) Analysis of dinoflagellate mitochondrial protein sorting signals indicates a highly stable protein targeting system across eukaryotic diversity. J Mol Biol **408**:643–653

Deschamps P, Guillebeault D, Devassine J, Dauvillée D, Haebel S, Steup M, Buléon A, Putaux JL, Slomianny MC, Colleoni C (2008) The heterotrophic dinoflagellate *Crypthecodinium cohnii* defines a model genetic system to investigate cytoplasmic starch synthesis. Eukaryot Cell **7**:872–880

Do TQ, Hsu AY, Jonassen T, Lee PT, Clarke CF (2001) A defect in Coenzyme Q biosynthesis is responsible for the respiratory deficiency in *Saccharomyces cerevisiae* abc1 mutants. J Biol Chem **276**:18161–18168

Dolezal P, Likic V, Tachezy J, Lithgow T (2006) Evolution of the molecular machines for protein import into mitochondria. Science **313**:314–318

Dorrell RG, Smith AG (2011) Do red and green make brown? Perspectives on plastid acquisitions within chromalveolates. Eukaryot Cell **10**:856–868

Dukanovic J, Rapaport D (2011) Multiple pathways in the integration of proteins into the mitochondrial outer membrane. BBA-Biomembranes **1808**:971–980

Foth BJ, Stimmler LM, Handman E, Crabb BS, Hodder AN, McFadden GI (2005) The malaria parasite *Plasmodium falciparum* has only one pyruvate dehydrogenase complex, which is located in the apicoplast. Mol Microbiol **55**:39–53

Gakh O, Cavadini P, Isaya G (2002) Mitochondrial processing peptidases. BBA-Mol Cell Res **1592**:63–77

Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S, et al., 45 co-authors (2002) Genome sequence of the human malaria parasite *Plasmodium falciparum*. Nature **419**:498–511

Hackett JD, Scheetz TE, Yoon HS, Soares MB, Bonaldo MF, Casavant TL, Bhattacharya D (2005) Insights into a dinoflagellate genome through expressed sequence tag analysis. BMC Genomics **29**:80

Heinz E, Lithgow T (2012) Back to basics: A revealing secondary reduction of the mitochondrial protein import pathway in diverse intracellular parasites. BBA- Mol Cell Res, http://dx.doi.org/10.1016/j.bbamcr.2012.02.006

Henriquez FL, Richards TA, Roberts F, McLeod R, Roberts CW (2005) The unusual mitochondrial compartment of *Cryptosporidium parvum*. Trends Parasitol **21**:68–74

Henze K, Horner DS, Suguri S, Moore DV, Sánchez LB, Müller M, Embley TM (2001) Unique phylogenetic relationships of glucokinase and glucosephosphate isomerase of the amitochondriate eukaryotes *Giardia intestinalis*, *Spironucleus barkhanus* and *Trichomonas vaginalis*. Gene **281**: 123–131

Holm-Hansen O (1969) Algae: amounts of DNA and organic carbon in single cells. Science **163**:87–88

Hoppenrath M, Leander BS (2010) Dinoflagellate phylogeny as inferred from heat shock protein 90 and ribosomal gene sequences. PLoS One **5**:e13220

Howe CJ, 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

Hug LA, Stechmann A, Roger AJ (2010) Phylogenetic distributions and histories of proteins involved in anaerobic pyruvate metabolism in eukaryotes. Mol Biol Evol **27**:311–324

Imanian B, Pombert JF, Keeling PK (2010) The complete plastid genomes of the two 'dinotoms' *Durinskia baltica* and *Kryptoperidinium foliaceum*. PLoS One **5**:e10711

Jackson CJ, Norman JE, Schnare MN, Gray MW, Keeling PJ, Waller RF (2007) Broad genomic and transcriptional analysis reveals a highly derived genome in dinoflagellate mitochondria. BMC Biol **5**:41

Joseph SJ, Fernández-Robledo JA, Gardner MJ, El-Sayed NM, Kuo CH, Schott EJ, Wang H, Kissinger JC, Vasta GR (2010) The alveolate *Perkinsus marinus*: biological insights from EST gene discovery. BMC Genomics **11**:228

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

Kanehisa M, Goto S (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res **28**:27–30

Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M (2012) KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Res **40**:D109–D114

Klodmann J, Sunderhaus S, Nimtz M, Jänsch L, Braun HP (2010) Internal architecture of mitochondrial complex I from *Arabidopsis thaliana*. Plant Cell **22**:797–810

Kořený L, Sobotka R, Janouškovec J, Keeling PJ, Oborník M (2011) Tetrapyrrole synthesis of photosynthetic chromerids is likely homologous to the unusual pathway of apicomplexan parasites. Plant Cell **23**:3454–3462

Koumandou VL, Nisbet RER, Barbrook AC, Howe CJ (2004) Dinoflagellate chloroplasts–where have all the genes gone? Trends Genet **20**:261–267

Krungkrai J (1995) Purification, characterization and localization of mitochondrial dihydroorotate dehydrogenase in *Plasmodium falciparum*, human malaria parasite. BBA-Gen Subjects **1243**:351–360

Leggat W, Hoegh-Guldberg O, Dove S, Yellowlees D (2007) Analysis of an EST library from the dinoflagellate (*Symbiodinium* sp.) symbiont of reef-building corals. J Phycol **43**:1010–1021

Lill R, Mühlenhoff U (2005) Iron-sulfur-protein biogenesis in eukaryotes. Trends Biochem Sci **30**:133–141

Lill R, Mühlenhoff U (2008) Maturation of iron-sulfur proteins in eukaryotes: mechanisms, connected processes, and diseases. Annu Rev Biochem **77**:669–700

Lin S (2006) The smallest dinoflagellate genome is yet to be found: a comment on Lajeunesse et al '*Symbiodinium* (Pyrrhophyta) genome sizes (DNA content) are smallest among dinoflagellates'. J Phycol **42**:746–748

Lister R, Chew O, Lee MN, Heazlewood JL, Clifton R, Parker KL, Millar AH, Whelan J (2004) A transcriptomic and proteomic characterization of the *Arabidopsis* mitochondrial protein import apparatus and its response to mitochondrial dysfunction. Plant Physiol **134**:777–789

Macedo CS, Uhrig ML, Kimura EA, Katzin AM (2002) Characterization of the isoprenoid chain of coenzyme Q in *Plasmodium falciparum*. FEMS Microbiol Lett **207**:13–20

Matus-Ortega MG, Salmerón-Santiago KG, Flores-Herrera O, Guerra-Sánchez G, Martínez F, Rendón JL, Pardo JP (2011) The alternative NADH dehydrogenase is present in mitochondria of some animal taxa. Comp Biochem Physiol Part D Genomics Proteomics **6**:256–263

Mogi T, Kita K (2010) Diversity in mitochondrial metabolic pathways in parasitic protists *Plasmodium* and *Cryptosporidium*. Parasitol Int **59**:305–312

Moore RB, Oborník M, Janouškovec J, Chrudimský T, Vancová M, Green DH, Wright SW, Davies NW, Bolch CJS, Heimann K (2008) A photosynthetic alveolate closely related to apicomplexan parasites. Nature **451**:959–963

Nash EA, Barbrook AC, Edwards-Stuart RK, Bernhardt K, Howe CJ, Nisbet RER (2007) Organization of the mitochondrial genome in the dinoflagellate *Amphidinium carterae*. Mol Biol Evol **24**:1528–1536

Please cite this article in press as: Butterfield ER, et al. An Analysis of Dinoflagellate Metabolism Using EST Data. Protist (2012), http://dx.doi.org/10.1016/j.protis.2012.09.001

Nash EA, Nisbet RER, Barbrook AC, Howe CJ (2008) Dinoflagellates: a mitochondrial genome all at sea. Trends Genet **24**:328–335

Nina PB, Dudkina NV, Kane LA, van Eyk JE, Boekema EJ, Mather MW, Vaidya AB (2010) Highly divergent mitochondrial ATP synthase complexes in *Tetrahymena thermophila*. PLoS Biology **8**:e1000418

Nosenko T, Bhattacharya D (2007) Horizontal gene transfer in chromalveolates. BMC Evol Biol **7**:173

Oborník M, Green BR (2005) Mosaic origin of the heme biosynthesis pathway in photosynthetic eukaryotes. Mol Biol Evol **22**:2343–2353

Olszewski KL, Mather MW, Morrisey JM, Garcia BA, Vaidya AB, Rabinowitz JD, Llinás M (2010) Branched tricarboxylic acid metabolism in *Plasmodium falciparum*. Nature **466**:774–778

Painter HJ, Morrisey JM, Mather MW, Vaidya AB (2007) Specific role of mitochondrial electron transport in blood-stage *Plasmodium falciparum*. Nature **446**:88–91

Paschen SA, Neupert W, Rapaport D (2005) Biogenesis of beta-barrel membrane proteins of mitochondria. Trends Biochem Sci **30**:575–582

Patron NJ, Waller RF, Archibald JM, Keeling PJ (2005) Complex protein targeting to dinoflagellate plastids. J Mol Biol **348**:1015–1024

Patron NJ, Waller RF, Keeling PJ (2006) A tertiary plastid uses genes from two endosymbionts. J Mol Biol **357**:1373–1382

Pinon V, Ravanel S, Douce R, Alban C (2005) Biotin synthesis in plants. The first committed step of the pathway is catalyzed by a cytosolic 7-keto-8-aminopelargonic acid synthase. Plant Physiol **139**:1666–1676

Rada P, Doležal P, Jedelský PL, Bursac D, Perry AJ, Šedinová M, Smíšková K, Novotný M, Beltrán NC, Hrdý I, Lithgow T, Tachezy J (2011) The core components of organelle biogenesis and membrane transport in the hydrogenosomes of *Trichomonas vaginalis*. PLoS One **6**:e24428

Ralph SA (2005) Strange organelles – *Plasmodium mitochondria* lack a pyruvate dehydrogenase complex. Mol Microbiol **55**:1–4

Rébeillé F, Alban C, Bourguignon J, Ravanel S, Douce R (2007) The role of plant mitochondria in the biosynthesis of coenzymes. Photosynth Res **92**:149–162

Richards TA, van der Giezen M (2006) Evolution of the Isd11-IscS complex reveals a single alpha-proteobacterial endosymbiosis for all eukaryotes. Mol Biol Evol **23**:1341–1344

Rotte C, Henze K, Müller M, Martin W (2000) Origins of hydrogenosomes and mitochondria: Commentary. Current Opin Microbiol **3**:481–486

Sanchez-Puerta M, Lippmeier JC, Apt KE, Delwiche CF (2007) Plastid genes in a non-photosynthetic dinoflagellate. Protist **158**:105–117

Sazanov LA, Jackson JB (1994) Proton-translocating transhydrogenase and NAD- and NADP-linked isocitrate dehydrogenases operate in a substrate cycle which contributes to fine regulation of the tricarboxylic acid cycle activity in mitochondria. FEBS Lett **344**:109–116

Schleiff E, Becker T (2011) Common ground for protein translocation: access control for mitochondria and chloroplasts. Nat Rev Mol Cel Bio **12**:48–59

Schreiner ME, Fiur D, Holátko J, Pátek M, Eikmanns BJ (2005) E1 enzyme of the pyruvate dehydrogenase complex in *Corynebacterium glutamicum*: molecular analysis of the gene and phylogenetic aspects. J Bacteriol **187**:6005–6018

Seeber F (2002) Biogenesis of iron-sulphur clusters in amitochondriate and apicomplexan protists. Int J Parasitol **32**:1207–1217

Slamovits CH, Keeling PJ (2008) Plastid-derived genes in the non-photosynthetic alveolate *Oxyrrhis marina*. Mol Biol Evol **25**:1297–1306

Slamovits CH, Saldarriaga JF, Larocque A, Keeling PJ (2007) The highly reduced and fragmented mitochondrial genome of the early-branching dinoflagellate *Oxyrrhis marina* shares characteristics with both apicomplexan and dinoflagellate mitochondrial genomes. J Mol Biol **372**:356–368

Smith AG, Croft MT, Moulin M, Webb ME (2007a) Plants need their vitamins too. Curr Opin Plant Biol **10**:266–275

Smith DGS, Gawryluk RMR, Spencer DF, Pearlman RE, Siu K, Gray MW (2007b) Exploring the mitochondrial proteome of the ciliate protozoon *Tetrahymena thermophila*: direct analysis by tandem mass spectrometry. J Mol Biol **374**:837–863

Stechmann A, Hamblin K, Pérez-Brocal V, Gaston D, Richmond GS, van der Giezen M, Clark CG, Roger AJ (2008) Organelles in *Blastocystis* that blur the distinction between mitochondria and hydrogenosomes. Curr Biol **18**:580–585

Stegmeier JF, Andersen C (2006) Characterization of pores formed by YaeT (Omp85) from *Escherichia coli*. J Biochem **140**:275–283

Stevens JM, Zhang Y, Muthuvel G, Sam KA, Allen JW, Ferguson SJ (2011) The mitochondrial cytochrome c N-terminal region is critical for maturation by holocytochrome c synthase. FEBS Lett **585**:1891–1896

Tanikawa N, Akimoto H, Ogoh K, Chun W, Ohmiya Y (2004) Expressed sequence tag analysis of the dinoflagellate *Lingulodinium polyedrum* during dark phase. Photochem Photobiol **80**:31–35

Uyemura SA, Luo S, Vieira M, Moreno SNJ, Docampo R (2004) Oxidative phosphorylation and rotenone-insensitive malate- and NADH-quinone oxidoreductases in *Plasmodium yoelii yoelii* mitochondria *in situ*. J Biol Chem **279**:385–393

van Dooren GG, Stimmler LM, McFadden GI (2006) Metabolic maps and functions of the *Plasmodium* mitochondrion. FEMS Microbiol Rev **30**:596–630

Walker G, Dorrell RA, Schlacht A, Dacks JB (2011) Eukaryotic systematics: a 2011 user's guide for cell biologists and parasitologists. Parasitology **138**:1638–1663

Waller RF, Slamovits CH, Keeling PJ (2006) Lateral gene transfer of a multigene region from cyanobacteria to dinoflagellates resulting in a novel plastid-targeted fusion protein. Mol Biol Evol **23**:1437–1443

Please cite this article in press as: Butterfield ER, et al. An Analysis of Dinoflagellate Metabolism Using EST Data. Protist (2012), http://dx.doi.org/10.1016/j.protis.2012.09.001

Werner J, Misra R (2005) YaeT (Omp85) affects the assembly of lipid-dependent and lipid-independent outer membrane proteins of *Escherichia coli*. Mol Microbiol **57**:1450–1459

Wiedemann N, Frazier AE, Pfanner N (2004) The protein import machinery of mitochondria. J Biol Chem **279**:14473–14476

Wu G, Henze K, Müller M (2001) Evolutionary relationships of the glucokinase from the amitochondriate protist, *Trichomonas vaginalis*. Gene **264**:265–271

Yang I, John U, Beszteri S, Glöckner G, Krock B, Goesmann A, Cembella AD (2010) Comparative gene expression in toxic versus non-toxic strains of the marine dinoflagellate *Alexandrium minutum*. BMC Genomics **11**:248

Yoon HS, Hackett JD, Van Dolah FM, Nosenko T, Lidie KL, Bhattacharya D (2005) Tertiary endosymbiosis driven genome evolution in dinoflagellate algae. Mol Biol Evol **22**:1299–1308

Zhang S, Bryant DA (2011) The tricarboxylic acid cycle in cyanobacteria. Science **334**:1551–1553

Please cite this article in press as: Butterfield ER, et al. An Analysis of Dinoflagellate Metabolism Using EST Data. Protist (2012), http://dx.doi.org/10.1016/j.protis.2012.09.001