# Sequence Evidence for the Presence of Two Tetrapyrrole Pathways in *Euglena gracilis*

Luděk Kořený[1,2], and Miroslav Oborník*[,1,2]

[1]Department of Molecular Biology of Protists, Biology Centre of the Academy of Sciences of the Czech Republic, Institute of Parasitology, České Budějovice, Czech Republic

[2]Department of Molecular Biology, Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic

*Corresponding author: E-mail: obornik@paru.cas.cz.

All newly sequenced data were deposited in GenBank under accession numbers: JF292577–JF292587.

**Accepted:** 22 March 2011

## Abstract

Genes encoding enzymes of the tetrapyrrole biosynthetic pathway were searched within *Euglena gracilis* EST databases and 454 genome reads and their 5′ end regions were sequenced when not available. Phylogenetic analyses and protein localization predictions support the hypothesis concerning the presence of two separated tetrapyrrole pathways in *E. gracilis*. One of these pathways resembles the heme synthesis in primarily heterotrophic eukaryotes and was presumably present in the host cell prior to secondary endosymbiosis with a green alga. The second pathway is similar to the plastid-localized tetrapyrrole syntheses in plants and photosynthetic algae. It appears to be localized to the secondary plastid, presumably derived from an algal endosymbiont and probably serves only for the production of plastidial heme and chlorophyll. Thus, *E. gracilis* represents an evolutionary intermediate in a metabolic transformation of a primary heterotroph to a photoautotroph through secondary endosymbiosis. We propose here that the tetrapyrrole pathway serves as a highly informative marker for the evolution of plastids and plays a crucial role in the loss of plastids.

**Key words:** tetrapyrrole synthesis, heme pathway, *Euglena gracilis*, secondary endosymbiosis, evolution, plastid.

Tetrapyrroles rank among the most important molecules for life on Earth. One of the tetrapyrrole compounds, heme, colors our blood where it functions as an oxygen transporter. However, it has many other cardinal functions in living systems, such as electron transport in the respiration chain, oxidative stress response, oxygen metabolism, detoxification, regulation of gene expression, and many others. Chlorophyll is another essential tetrapyrrole; this molecule enables phototrophic organisms to utilize the energy of light. All tetrapyrroles are synthesized from δ-aminolevulinic acid (ALA) through a route that is conserved in organisms from all three domains of life. However, ALA can be formed in two distinct ways and most organisms use only one of them. Archaea and all bacteria except α-proteobacteria synthesize ALA from glutamate bound to $tRNA^{Glu}$, serving for both protein and tetrapyrrole syntheses (Panek and O'Brian 2002). This molecule is recognized by glutamyl-tRNA reductase (GTR), which catalyzes its reduction to glutamate 1-semialdehyde, a molecule that is further converted into ALA by glutamate 1-semialdehyde aminotransferase (GSA-AT). These two reactions are called the C5 pathway. Another pathway for ALA synthesis was found in α-proteobacteria. They utilize a single enzyme called ALA-synthase (ALAS) to form ALA through the condensation of succinyl-CoA with glycine (the so-called C4 or Shemin pathway). ALA-synthase has probably evolved from GSA-AT because both enzymes share sequence and structural similarities (Schulze et al. 2006). ALAS is also encoded in the genomes of several Actinobacteria but is used exclusively for the synthesis of antibiotics and not for tetrapyrrole biosynthesis (Petříček et al. 2006).

Eukaryotic tetrapyrrole biosynthesis is more complicated. It was largely affected by their complex evolutionary history involving endosymbioses that gave rise to the semiautonomous organelles such as mitochondria and plastids. Most heterotrophic eukaryotes form ALA by the C4 pathway in the mitochondrion. The gene encoding ALA-synthase was likely transferred to the eukaryotic nucleus from the α-proteobacterial ancestor of the mitochondrion by endosymbiotic gene transfer. Alternatively, photosynthetic eukaryotes use the plastid C5 pathway, and both crucial enzymes (GTR and GSA-AT) are encoded by the originally cyanobacterial genes that are now located in the nucleus (Oborník and

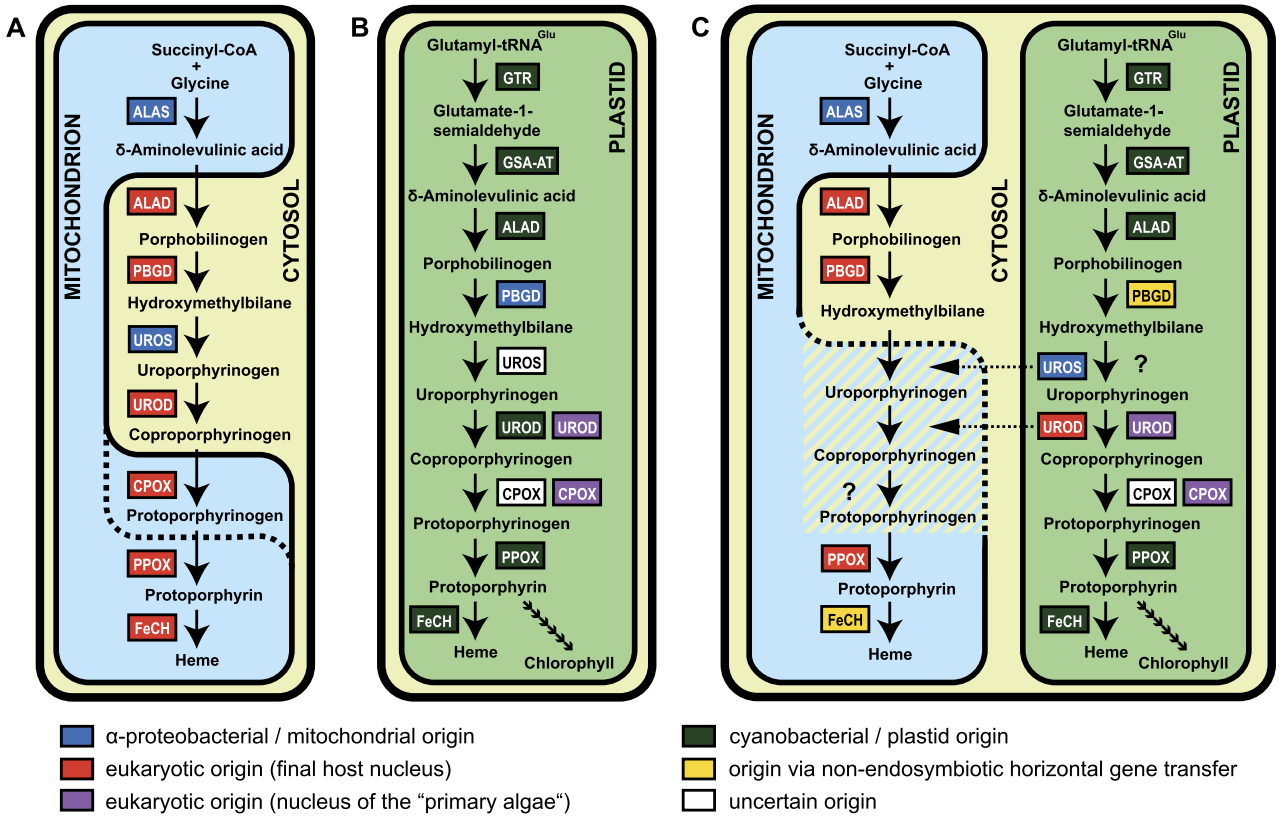


**Fig. 1.**—Origins and subcellular localizations of the tetrapyrrole synthesis enzymes in different eukaryotes. *A*. Heme biosynthesis starting with C4 pathway of primary heterotrophic eukaryotes. The localization of CPOX differs between the animal cell (continuous line) and yeast (dashed line). *B*. Tetrapyrrole biosynthesis pathway that is shared by most of the photosynthetic eukaryotes starting from glutamyl-tRNA$^{Glu}$. *C*. Origins and suggested locations of the two tetrapyrrole pathways in *Euglena gracilis*. Origins of the enzymes are inferred from the phylogenetic trees that are depicted in supplementary figure S1 (Supplementary Material online). The locations of PPOX and FeCH in the plastid of *E. gracilis* are implied only by their origins because the N-terminal sequences are not available for these enzymes. Dashed arrowheads indicate possible dual localizations of UROS and UROD in the plastid and either cytosol or mitochondrion of *E. gracilis*. The question marks stand for enzyme orthologues that are present in either photosynthetic or heterotrophic eukaryotes but were not found in the limited sequence data of *E. gracilis*.

Green 2005). The remaining steps of the pathway leading to the formation of heme are the same in all organisms; however, pathways in heterotrophic and photosynthetic eukaryotes differ in the origins of the genes and subcellular localizations of their products. In primarily heterotrophic eukaryotes, ALA is exported from the mitochondrion to the cytosol, where the next four to five steps take place. The pathway is terminated in the mitochondrion again, where heme is needed mainly to form respiratory cytochromes (Camadro et al. 1986; Dailey et al. 2005). Except ALAS and uroporphyrinogen synthase (UROS), all other enzymes appear to be encoded by genes of eukaryotic origin (fig. 1*A*; see supplementary fig. S1, Supplementary Material online for phylogenetic trees). Photosynthetic eukaryotes synthesize both heme and chlorophyll by a single pathway that is localized within the plastid. A majority of the intermediates are used for chlorophyll synthesis, and most of the synthesized heme is used for photosynthetic functions as well (photosynthetic cytochromes, synthesis of bilin pigments). Such a massive flow of intermediates to the chloroplast may be the reason why this organelle took over the synthesis of tetrapyrroles for the whole cell. Most of the genes of tetrapyrrole synthesis were transferred from the cyanobacterial endosymbiont to the nucleus of the eukaryotic host, and the original pathway of the exosymbiont disappeared. The same is true for heterokont algae such as diatoms, which went through a secondary endosymbiosis with their plastids derived from an engulfed rhodophyte (Oborník and Green 2005).

Photosynthetic euglenids acquired their plastid by secondary endosymbiosis with a green alga relatively recently (Rogers et al. 2007). *Euglena gracilis* is the only eukaryote known to possess enzymes of both the C5 and C4 pathways. Weinstein and Beale (1983) used $^{14}$C-labeled precursors of both of these alternative pathways and monitored their incorporation into various tetrapyrroles. They found out that protoheme was labeled when using either glycine or glutamate in wild-type cells, whereas the dark-grown or

aplastidic light-grown cells incorporated only glycine to the synthesized tetrapyrroles. Furthermore, heme-a, which is present exclusively in the mitochondrion, was labeled only when using $^{14}$C glycine, whereas chlorophyll was derived only from glutamate. These findings provided quite strong evidence for the presence of two independent tetrapyrrole synthetic pathways that are spatially separated within the cell. These results were further corroborated by Okazaki et al. (1990), who studied the incorporation of [1-$^{13}$C]glutamate and [2-$^{13}$C]glycine into chlorophyll-a. According to NMR spectra, eight carbons of chlorophyll were labeled when using [1-$^{13}$C]glutamate, which corresponded to the conversion of this precursor into chlorophyll via the C5 pathway. On the other hand, [2-$^{13}$C]glycine was only incorporated into the methyl ester carbon of chlorophyll via the one-carbon metabolic pathway but not via the C4 pathway. Contradicting results were reported in two recent articles (Iida et al. 2002; Iida and Kajiwara 2008). They incubated the cells with $^{13}$C glucose or alanine and determined which carbons of chlorophyll were labeled using NMR spectroscopy. Based on their results, they deduced that chlorophyll was derived from both pools of ALA. However, since the substrates used for the labeling are not metabolized to ALA directly but enter various pathways, the discrepancy may be explained by the scrambling of the $^{13}$C label. Shashidara and Smith (1991) localized porphobilinogen deaminases (PBGD) only to the plastids in different cultures of *E. gracilis* (wild-type light-grown, wild-type dark-grown, and the mutant strain unable to synthesize chlorophyll) and in the naturally heterotrophic *E. longa*. Based on this finding, they concluded that there is only a single tetrapyrrole pathway in *E. gracilis*.

Because many expressed sequence tags and 454 genome reads from *E. gracilis* became recently available, we searched for the genes encoding enzymes of tetrapyrrole synthesis hoping to shed some light on their origins and intracellular locations. We amplified the 5′ ends of the incomplete gene sequences for targeting predictions and performed phylogenetic analysis on each gene of this pathway. We have found genes encoding enzymes of both alternative pathways of ALA synthesis (C4 and C5 pathways), but more importantly, most of the subsequent steps were found to be encoded by more than a single gene. According to our phylogenetic analyses (see fig. 1 and supplementary fig. S1, Supplementary Material online), the gene encoding ALA-synthase was derived from the mitochondrion and clusters with heterotrophic eukaryotes, whereas the genes for the C5 pathway appear to originate from a cyanobacterial predecessor of plastids as in other photoeukaryotes. The enzymes of the two subsequent steps are each encoded by two different genes. One of the genes for ALA-dehydratase (ALAD) clusters with photosynthetic eukaryotes and cyanobacteria, whereas the second ALAD and one of the PBGD are related to their counterparts from heterotrophic eukaryotes. Although the monophyly of the clade comprising PBGDs of heterotrophic eukaryotes and *E. gracilis* is not supported well, its separation from PBGDs of photosynthetic eukaryotes is obvious (supplementary fig. S1*E*, Supplementary Material online). The second PBGD clusters with γ-proteobacteria. Only one gene coding for UROS was identified and is related to its homologues from heterotrophic eukaryotes and α-proteobacteria, implying its mitochondrial origin. Two genes coding for uroporphyrinogen decarboxylase (UROD) are both related to the homologues from heterotrophic eukaryotes and are probably a result of gene duplication. The third UROD seems to be of eukaryotic origin as well but specifically clusters with enzymes from photosynthetic eukaryotes, indicating ancestry in the nucleus of endosymbiotic alga. The same is true for the four slightly different genes for coproporphyrinogen oxidase (CPOX) that probably underwent three subsequent gene duplications. The fifth CPOX is related to the enzymes of photosynthetic eukaryotes as well, but this clade has a different and unclear origin. CPOX related to heterotrophic eukaryotes was not identified within the available sequence data. The last two enzymes of heme biosynthesis are encoded each by two genes of different origin. One of the protoporphyrinogen oxidases (PPOXs) and one FeCH exhibit an origin in the cyanobacterial predecessor of plastids. Moreover, the plastid-like FeCH of *E. gracilis* contains conserved C-terminal CAB domain (Sobotka et al. 2010), including a chlorophyll-binding motif, which is present in ferrochelatases of most photosynthetic algae and cyanobacteria but not in the second ferrochelatase of *E. gracilis* and heterotrophs. The second PPOX belongs to a clade consisting of heterotrophic eukaryotes and probably represents the original eukaryotic gene. The second FeCH forms a sister lineage to FeCH from *Naegleria gruberi*, another protist from the supergroup Excavata. Both these genes are placed within the group Bacteroidetes, which indicates their origins via nonendosymbiotic horizontal gene transfer. It appears that FeCH was transferred to the nuclei of different eukaryotic lineages from different bacteria several times independently (supplementary fig. S1*J*, Supplementary Material online). Another gene-transfer event took place in trypanosomatids that are more closely related to *Euglena* than *Naegleria*. This incident probably happened more recently after the pathway was completely lost in Kinetoplastida (Kořený et al. 2010).

According to in silico predictions, the proteins of the tetrapyrrole pathway transferred from the algal endosymbiont possess two transmembrane helices at the N-terminal regions (supplementary fig. S2, Supplementary Material online), which can serve as evidence for the targeting of the protein into the *Euglena* plastid (Durnford and Gray 2006). The first helix represents the signal sequence, whereas the second is the so-called stop signal. The same feature is also present in the N-terminus of the "γ-proteobacterial"

PBGD, for which plastid localization was experimentally shown (Shashidara and Smith 1991). Interestingly, two enzymes representing the original mitochondria–cytosolic pathway of the host (UROS and one UROD) are putatively targeted to the plastid as well. They possess long N-terminal presequences including two hydrophobic patches resembling the signal sequence and the stop signal and the region between these patches has similar biochemical properties as described for the transit peptide (Durnford and Gray 2006). However, dual localization of these enzymes cannot be ruled out. There are several methionines found in the N-terminal presequence of UROS, which may serve as alternative start codons. It should be also mentioned that the N-terminal presequence is available only for one of the two URODs related to heterotrophic eukaryotes. The second one may have a different N-terminus because both duplicates share only 48.5% of 200 alignable amino acid residues. ALAS, an enzyme of C4 pathway, is likely localized in the mitochondrion as in heterotrophic eukaryotes. Although predictions of mitochondrial targeting of this enzyme are not convincing (supplementary table S1, Supplementary Material online), it possesses N-terminal presequence that is comparable in length to other eukaryotic ALA-synthases and has no transmembrane helices (supplementary fig. S3, Supplementary Material online). On the other hand, no N-terminal presequences are present in eukaryotic ALAD and PBGD (supplementary fig. S3, Supplementary Material online). These enzymes are thus probably cytosolic, as in animals or fungi. The PPOX related to the genes from heterotrophic eukaryotes lacks a significant N-terminal presequence as well. However, the same is true for PPOX of heterotrophic eukaryotes, including those where this enzyme was shown to be localized in the mitochondrion. The *Naegleria*-like FeCH is very likely localized in the mitochondrion (supplementary table S1, Supplementary Material online), and similarly to ALAS, it has a comparable N-terminal presequence to FeCH from heterotrophs, with an experimentally confirmed mitochondrial localization. In humans, PPOX closely interacts with FeCH through the inner mitochondrial membrane (Ferreira et al. 1988). Complex formation between these enzymes was also suggested based on the crystal structures (Koch et al. 2004) and later experimental confirmation in a cyanobacterium (Masoumi et al. 2008). The colocalization of the last two steps of the heme pathway in the mitochondrion seems to be conserved among different lineages of heterotrophic eukaryotes. Although in the animal cell, the last three steps are placed in the mitochondrion, in yeast and apicomplexan parasites, CPOX is cytosolic, and only the last two steps take place in the mitochondrion (Camadro et al. 1986; Dailey et al. 2005; van Dooren et al. 2006; Wu 2006; Nagaraj et al. 2009; Nagaraj, Arumugam, et al. 2010; Nagaraj, Prasad, et al. 2010). Based on the data mentioned above, we envisage that both heterotroph-like PPOX and FeCH are localized in the mitochondrion of *E. gracilis*. The suggested origins and localizations of heme-synthesis enzymes in *E. gracilis* are depicted in figure 1.

Results of our analyses clearly indicate that *E. gracilis* not only possesses genes for both alternative pathways of ALA synthesis but also support the presence of two independent pathways from ALA to heme as well, which is in accordance with the results of previous biochemical studies. We refute the conclusion that there is only one pathway from ALA to heme, which was based on the localization of PBGD solely to the plastid (Shashidara and Smith 1991). Not only we have found a gene for a second PBGD that seems to be localized in the cytosol, but two different enzymes with PBGD activity were purified in the past. One was found in isolated chloroplasts and the second in the cytosol of *E. gracilis* (Rossetti et al. 1986, 1989; Juknat et al. 1989). Because the two genes for PBGD in *E. gracilis* display distinct evolutionary histories, it is not surprising that both enzymes were not identified by the single antibody used by Shashidara and Smith (1991). There are several lines of evidence supporting the presence of two independent pathways for tetrapyrrole synthesis in *E. gracilis*. Apart from the direct biochemical evidence (Weinstein and Beale 1983; Okazaki et al. 1990) and the identification of genes for both pathways in this study, there are other findings indicating that chlorophyll and the nonplastidial heme are synthesized from different pools of ALA. It was found that dark-grown wild-type cells and the aplastidic mutant of *E. gracilis* are not only incapable of chlorophyll synthesis but are not able to form ALA through the C5 pathway (Mayer et al. 1987; Mayer and Beale 1991). Conversely, when the C5 pathway was blocked with the inhibitor of GSA-AT, *E. gracilis* was not able to form chlorophyll (Corriveau and Beale 1986). In another study, a specific point mutation in the chloroplast $tRNA^{Glu}$ that uncouples protein synthesis and the C5 pathway was identified in *E. gracilis* (Stange-Thomann et al. 1994). Chloroplast protein synthesis is not affected in this mutant, but $tRNA^{Glu}$ is not recognized by GTR, the first enzyme of the C5 pathway. Because this mutant was still able to live heterotrophically under aerobic conditions and without an external source of heme, it is clear that biosynthesis of this tetrapyrrole is not affected. However, when a very sensitive assay and highly concentrated extract of *E. gracilis* were used for chlorophyll detection, not even a minimal amount of chlorophyll was detected in this mutant (Russell and Draffan 1978).

Taking all available evidence together, we conclude that *E. gracilis* synthesizes tetrapyrroles via two independent pathways that have distinct origins and serve different cellular compartments. One pathway comes from the exosymbiont and resembles the heme synthesis of heterotrophic eukaryotes, whereas the other one is mostly derived from endosymbiotic alga and produces chlorophyll and heme for the plastid. *Euglena gracilis* with two tetrapyrrole pathways may represent an analogy to the suggested state in the ancestral lineages of chromalveolates, before the plastid

was lost in some of them or took over tetrapyrrole synthesis and become indispensable in others (Oborník et al. 2009). Heterotrophic chromalveolates with no remnants of a plastid, such as ciliates and oomycetes, possess a heme pathway that is similar to that of primary heterotrophs. But the plastid-containing chromalveolates synthesize all tetrapyrroles via a pathway that is derived from the plastid, which makes this organelle essential even after loss of photosynthesis. Apicomplexan parasites can serve as an example. Part of their heme synthetic pathway is localized in a plastid that is no longer photosynthetic but is essential for the cell (Sato et al. 2004). However, apicomplexans, which do not synthesize their own heme such as *Cryptosporidium* or gregarines, do not possess any detectable plastids (Mather and Vaidya 2008; Templeton et al. 2010). *Euglena gracilis* may easily and irreversibly lose the photosynthetic capability of the chloroplast when grown in dark at higher temperatures or after antibiotic treatment (Cook 1974; Nicolas 1981). It appears that *Euglena* is in the stage of symbiosis when it can still completely lose the plastid. There are several mutants in which neither the plastid membranes nor the chloroplast DNA were detected (Osafune and Schiff 1983; Ohki et al. 1984; Ikeda and Takata 2002). *Euglena gracilis* thus probably represents an intermediate state in the transformation of a heterotroph into a photoautotroph through secondary endosymbiosis and shows an arrangement of tetrapyrrole biosynthesis presumably present in the evolutionary history of other secondary algae. The presence of both heme pathways in a single cell may represent the breaking point in the evolution of secondary endosymbiosis, to which a plastid can be completely lost.

## Materials and Methods

### Sequence Retrieval and Phylogenetic Analyses

Sequences of the genes of the heme biosynthesis pathway were searched in the available sequence databases: PEPdb (http://amoebidia.bcm.umontreal.ca/pepdb), GenBank and *E. gracilis* genome data were obtained from the MC Field laboratory sequence database at: http://web.me.com/mfield/Euglena_gracilis (see supplementary table S2, Supplementary Material online for sequence IDs). Where only small pieces of genes available, we amplified larger gene segments from the cDNA of *E. gracilis* strain Z (kindly provided by Professor Krajčovič, Bratislava). The 5′ regions were amplified using the FirstChoice RLM-RACE kit (Ambion) or the spliced-leader specific primer (5′-ACACTTTCTGAGTGTCTATTTTTTTTCG-3′). All newly sequenced data were deposited in GenBank under accession numbers: JF292577-JF292587.

For each enzyme, appropriate homologues were identified by Blast, aligned using MAFFT (Katoh and Toh 2008) and manually edited using BioEdit (Hall 1999). Phylogenetic trees were computed using PhyloBayes 3.2f under the CAT-GTR evolutionary model (Lartillot and Philippe 2004). For each analysis, two independent chains were run for at least 50,000 steps. The convergence of the chains was assessed based on bpcomp analysis (PhyloBayes 3.2f). For the 50% majority consensus trees, each 10th tree was sampled and the first 10% of the trees were discarded. The bootstrap support values were calculated in RAxML 7.0.3 (Stamatakis 2006) using the PROTGAMMALG model after 1,000 replications.

### Putative Localizations of Enzymes of the Tetrapyrrole Synthesis in *E. gracilis*

Transmembrane helices indicating the plastid localization of the enzymes were determined using the TMHMM program (http://www.cbs.dtu.dk/services/TMHMM-2.0/). The presence of mitochondrial transit peptides (mTP) was tested by several prediction tools: TargetP (http://www.cbs.dtu.dk/services/TargetP/), Predotar (http://urgi.versailles.inra.fr/predotar/), MitoProt II (http://ihg.gsf.de/ihg/mitoprot.html), Protein Prowler (http://pprowler.imb.uq.edu.au/), SLP-Local (http://sunflower.kuicr.kyoto-u.ac.jp/~smatsuda/slplocal.html), and SLPFA (http://sunflower.kuicr.kyoto-u.ac.jp/~tamura/slpfa.html).

## Supplementary material

Supplementary figures S1–S3 and tables S1–S2 are available at *Genome Biology and Evolution online* (http://gbe.oxfordjournals.org/).

## Acknowledgments

This work was supported by the Czech Science Foundation (206/08/1423 and 206/09/H026), Academy of Sciences of the Czech Republic (z60220518), and Ministry of Education of the Czech Republic (6007665801).

## Literature Cited

Camadro JM, Chambon H, Jolles J, Labbe P. 1986. Purification and properties of coproporphyrinogen oxidase from the yeast *Saccharomyces cerevisiae*. Eur J Biochem. 156:579–587.

Cook JR. 1974. Irreversible plastid loss in *Euglena gracilis* under physiological conditions. Plant Physiol. 53:284–290.

Corriveau JL, Beale SI. 1986. Influence of gabaculine on growth, chlorophyll synthesis, and δ-aminolevulinic acid synthase activity in *Euglena gracilis*. Plant Sci. 45:9–17.

Dailey TA, Woodruff JH, Dailey HA. 2005. Examination of mitochondrial protein targeting of haem synthetic enzymes: in vivo identification of three functional haem-responsive motifs in 5-aminolaevulinate synthase. Biochem J. 386:381–386.

Durnford DG, Gray MW. 2006. Analysis of *Euglena gracilis* plastid-targeted proteins reveals different classes of transit sequences. Eukaryot Cell. 5:2079–2091.

Ferreira GC, Andrew TL, Karr SW, Dailey HA. 1988. Organization of the terminal two enzymes of the heme biosynthetic pathway. Orientation of protoporphyrinogen oxidase and evidence for a membrane complex. J Biol Chem. 263:3835–3839.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser. 41:95–98.
Iida K, Kajiwara M. 2008. Carbon source dependence of the ratio of δ-aminolevulinic acid biosynthesis via the C5 and shemin pathways in *Euglena gracilis* (Euglenophyceae). J Phycol. 44: 292–298.
Iida K, Mimura I, Kajiwara M. 2002. Evaluation of two biosynthetic pathways to δ-aminolevulinic acid in *Euglena gracilis*. Eur J Biochem. 269:291–297.
Ikeda S, Takata N. 2002. Deoxyribonuclease II purified from *Euglena gracilis* SM-ZK, a chloroplast-lacking mutant: comparison with porcine spleen deoxyribonuclease II. Comp Biochem Phys B Biochem Mol Biol. 131:519–525.
Juknat AA, Dornemann D, Senger H. 1989. Different porphobilinogenases in cytoplasm and isolated-chloroplasts from light-grown *Euglena gracilis* Z. Z Naturforsch C. 44:81–84.
Katoh K, Toh H. 2008. Recent developments in the MAFFT multiple sequence alignment program. Brief Bioinform. 9:286–298.
Koch M, et al. 2004. Crystal structure of protoporphyrinogen IX oxidase: a key enzyme in haem and chlorophyll biosynthesis. EMBO J. 23:1720–1728.
Kořený L, Lukeš J, Oborník M. 2010. Evolution of the haem synthetic pathway in kinetoplastid flagellates: an essential pathway that is not essential after all? Int J Parasitol. 40:149–156.
Lartillot N, Philippe H. 2004. A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. Mol Biol Evol. 21:1095–1109.
Masoumi A, et al. 2008. Complex formation between protoporphyrinogen IX oxidase and ferrochelatase during haem biosynthesis in *Thermosynechococcus elongatus*. Microbiology. 154:3707–3714.
Mather W, Vaidya AB. 2008. Mitochondria in malaria and related parasites: ancient, diverse and streamlined. J Bioenerg Biomembr. 40:425–433.
Mayer SM, Beale SI. 1991. δ-aminolevulinic acid biosynthesis from glutamate in *Euglena gracilis*. Plant Physiol. 97:1094–1102.
Mayer SM, Beale SI, Weinstein JD. 1987. Enzymatic conversion of glutamate to δ-aminolevulinic acid in soluble extracts of *Euglena gracilis*. J Biol Chem. 262:12541–12549.
Nagaraj VA, Arumugam R, Prasad D, Rangarajan PN, Padmanaban G. 2010. Protoporphyrinogen IX oxidase from *Plasmodium falciparum* is anaerobic and is localized to the mitochondrion. Mol Biochem Parasitol. 174:44–52.
Nagaraj VA, Prasad D, Arumugam R, Rangarajan PN, Padmanaban G. 2010. Characterization of coproporphyrinogen III oxidase in *Plasmodium falciparum* cytosol. Parasitol Int. 59:121–127.
Nagaraj VA, Prasad D, Rangarajan PN, Padmanaban G. 2009. Mitochondrial localization of functional ferrochelatase from *Plasmodium falciparum*. Mol Biochem Parasitol. 168:109–112.
Nicolas P. 1981. Sensitivity of *Euglena gracilis* to chloroplast-inhibiting antibiotics, and properties of antibiotic-resistant mutants. Plant Sci Lett. 22:309–316.
Oborník M, Green BR. 2005. Mosaic origin of the heme biosynthesis pathway in photosynthetic eukaryotes. Mol Biol Evol. 22:2343–2353.
Oborník M, Janouškovec J, Chrudimský T, Lukeš J. 2009. Evolution of the apicoplast and its hosts: from heterotrophy to autotrophy and back again. Int J Parasitol. 39:1–12.
Ohki Y, Hasegawa K, Musashi A, Tsubo Y. 1984. Loss of chloroplastic DNA in a *Euglena* mutant during growth in darkness. Arch Microbiol. 139:147–150.
Okazaki T, Kurumaya K, Sagae Y, Kajiwara M. 1990. Studies on the biosynthesis of corrinoids and porphyrinoids. IV. Biosynthesis of chlorophyll in *Euglena gracilis*. Chem Pharm Bull. 38:3303–3307.
Osafune T, Schiff JA. 1983. W10BSmL, a mutant of Euglena gracilis var. bacillaris lacking plastids. Exp Cell Res. 148:530–536.
Panek H, O'Brian MR. 2002. A whole genome view of prokaryotic haem biosynthesis. Microbiology. 148:2273–2282.
Petříček M, Petříčková K, Havlíček L, Felsberg J. 2006. Occurrence of two 5-aminolevulinate biosynthetic pathways in *Streptomyces nodosus* subsp. *asukaensis* is linked with the production of asukamycin. J Bacteriol. 188:5113–5123.
Rogers MB, Gilson PR, Su V, McFadden GI, Keeling PJ. 2007. The complete chloroplast genome of the chlorarachniophyte *Bigelowiella natans*: evidence for independent origins of chlorarachniophyte and euglenid secondary endosymbionts. Mol Biol Evol. 24:54–62.
Rossetti MV, Juknat AA, Battle AMD. 1989. Soluble and particulate porphobilinogen-deaminase from dark-grown *Euglena gracilis*. Z Naturforsch C. 44:578–580.
Rossetti MV, et al. 1986. Porphyrin biosynthesis in *Euglena gracilis*-V. Soluble and particulate PBG-ASE. Comp Biochem Phys B. 85:451–458.
Russell GK, Draffan AG. 1978. Light-induced enzyme formation in a chlorophyll-less mutant of *Euglena gracilis*. Plant Physiol. 62: 678–682.
Sato S, Clough B, Coates L, Wilson RJM. 2004. Enzymes for heme biosynthesis are found in both the mitochondrion and plastid of the malaria parasite *Plasmodium falciparum*. Protist. 155: 117–125.
Schulze JO, Schubert WD, Moser J, Jahn D, Heinz DW. 2006. Evolutionary relationship between initial enzymes of tetrapyrrole biosynthesis. J Mol Biol. 358:1212–1220.
Shashidara LS, Smith AG. 1991. Expression and subcellular location of the tetrapyrrole synthesis enzyme porphobilinogen deaminase in light-grown *Euglena gracilis* and three nonchlorophyllous cell lines. Proc Natl Acad Sci U S A. 88:63–67.
Sobotka R, Tichý M, Wilde A, Hunter CN. 2010. Functional assignments for the C-terminal domains of the ferrochelatase from *Synechocystis* PCC 6803: the CAB domain plays a regulatory role and region II is essential for catalysis. Plant Physiol. doi:10.1104/pp.110.167528
Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 22:2688–2690.
Stange-Thomann N, Thomann H-U, Lloyd AJ, Lyman H, Söll D. 1994. A point mutation in Euglena gracilis chloroplast tRNA(Glu) uncouples protein and chlorophyll biosynthesis. Proc Natl Acad Sci U S A. 91:7947–7951.
Templeton TJ, et al. 2010. A genome-sequence survey for *Ascogregarina taiwanensis* supports evolutionary affiliation but metabolic diversity between a gregarine and *Cryptosporidium*. Mol Biol Evol. 27:235–248.
van Dooren GG, Stimmler LM, McFadden GI. 2006. Metabolic maps and functions of the *Plasmodium* mitochondrion. FEMS Microbiol Rev. 30:596–630.
Weinstein JD, Beale SI. 1983. Separate physiological roles and subcellular compartments for two tetrapyrrole biosynthetic pathways in *Euglena gracilis*. J Biol Chem. 258:6799–6807.
Wu B. 2006. Heme biosynthetic pathway in apicomplexan parasites [dissertation]. University of Pennsylvania. Available from ProQuest. Publication number 3246256.

**Associate editor:** Geoff McFadden