An unprecedented combination of metabolic pathways in the cryptic plastid of a non-photosynthetic euglenophyte

Zoltán Füssy[1,2,#], Kristína Záhonová[1,2,3,#], Aleš Tomčala[1,§], Juraj Krajčovič[4], Vyacheslav Yurchenko[3], Miroslav Oborník[1,5], Marek Eliáš[3,*]

[1] Institute of Parasitology, Biology Centre ASCR, 370 05 České Budějovice, Czech Republic

[2] Faculty of Science, Charles University, BIOCEV, 252 50 Vestec, Czech Republic

[3] Life Science Research Centre, Department of Biology and Ecology and Institute of Environmental Technologies, Faculty of Science, University of Ostrava, 701 00 Ostrava, Czech Republic

[4] Department of Biology, Faculty of Natural Sciences, University of ss. Cyril and Methodius in Trnava, 917 01 Trnava, Slovakia

[5] University of South Bohemia, Faculty of Science, 370 05 České Budějovice, Czech Republic

[#] These authors contributed equally to this work.

[§] Present address: University of South Bohemia, Faculty of Fisheries and Protection of Waters, CENAKVA, 370 05 České Budějovice, Czech Republic

[*] Corresponding author:

E-mail: marek.elias@osu.cz (ME)

Abstract

The non-photosyntetic alga *Euglena longa* harbours a cryptic plastid of unknown function. By a combination of bioinformatic and biochemical approaches we found out that this organelle houses a surprising set of metabolic processes. Biosynthesis of isoprenoid precursors and fatty acids is absent and the tetrapyrrole pathway is incomplete, whereas phospholipids and glycolipids are still being produced in the *E. longa* plastid. Unprecedented among non-photosynthetic plastids is the ability of this organelle to make tocopherols and a phylloquinone derivative. The most striking attribute is the presence of a linearized Calvin-Benson (CB) pathway including RuBisCO, together with ferredoxin-$NADP^+$ reductase and the ferredoxin/thioredoxin system. We hypothesize that the linear CB pathway is regulated by the redox status of the *E. longa* cell, in effect functioning as a redox valve bypassing the glycolytic oxidation of glyceraldehyde-3-phosphate to 3-phosphoglycerate. Altogether, the *E. longa* plastid defines a new class of relic plastids.

Key words: Calvin-Benson cycle, Euglenophyceae, evolution, non-photosynthetic plastids, phylloquinone, redox balance

Abbreviations: ACP, acyl carrier protein; CBC, Calvin-Benson cycle; FAS, fatty acid synthesis; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; IPP, isopentenyl pyrophosphate; MEP, 2-C-methyl-d-erythritol 4-phosphate; MGDG/DGDG mono-/digalactosyldiacylglycerol; MVA, mevalonate; OH-PhQ, 5'-monohydroxyphylloquinone; PhQ, phylloquinone; SQDG, sulfoquinovosyldiacylglycerol; SP, signal peptide; TMD, transmembrane domain

Introduction

Plastids are organelles that evolved from a cyanobacterial endosymbiont in the ancestor of Archaeplastida and later found their way into other eukaryotic lineages by secondary or even higher-order endosymbioses (Keeling 2013; McFadden 2014; Ponce-Toledo et al. 2017). Photosynthetic harvesting of solar energy is supposedly the primary metabolic function and evolutionary advantage of plastid acquisition. However, plastids also host a variety of other metabolic pathways, such as biosynthesis of amino and fatty acids, isopentenyl pyrophosphate (IPP) and its derivatives (isoprenoids), and tetrapyrroles (Neuhaus and Emes 2000; Oborník and Green 2005; Van Dingenen et al. 2016). Hence, reversion of photosynthetic lineages to heterotrophy typically does not imply plastid loss, and non-photosynthetic plastids are found in many eukaryotic lineages (Wilson et al. 1996; Sanchez-Puerta et al. 2007; Slamovits and Keeling 2008; Janouškovec et al. 2015; Kamikawa et al. 2017; Hadariová et al. 2018). In these cases, metabolic integration

55 of the plastid presumably resulted in the host biochemistry being dependent on compound(s) supplied by
56 the organelle (Oborník et al. 2009; Lim and McFadden 2010; Janouškovec et al. 2015; Hadariová et al.
57 2018).

58 The most extensively studied relic plastid is undoubtedly the apicoplast of apicomplexan parasites
59 (especially of *Plasmodium falciparum* and *Toxoplasma gondii*). The knowledge of the apicoplast has
60 expanded tremendously since its discovery two decades ago (McFadden and Yeh 2017). The apicoplast is
61 ultimately derived from a red alga (Williamson et al. 1994; Moore et al. 2008; Janouškovec et al. 2010),
62 with an ochrophyte alga being a possible direct donor of the plastids in apicomplexans and their
63 photosynthetic relatives, chromerids (Ševčíková et al. 2015; Füssy and Oborník 2017). Except for the
64 recently characterized corallicolid apicomplexans (Kwong et al. 2019), the apicoplast genome contains only
65 genes related to gene expression, protein turnover, and assembly of FeS clusters, not considering a few short
66 hypothetical open reading frames of unknown function (Wilson et al. 1996; Janouškovec et al. 2015). The
67 essentiality of the apicoplast for parasite survival has attracted much attention, partly because this organelle
68 is a promising target for parasite-specific inhibitors (e.g. Miller et al. 2013; McFadden and Yeh 2017). So
69 far, three plastid pathways (encoded by the nuclear genome) seem to condition the apicoplast retention: non-
70 mevalonate IPP synthesis, haem synthesis, and type II fatty acid synthesis (FASII). IPP biosynthesis in
71 particular is vital for the bloodstream form of *P. falciparum* (Yeh and DeRisi 2011). On the other hand, the
72 mosquito and liver stages of *P. falciparum* are dependent on haem synthesis (Nagaraj et al. 2013), while
73 FASII is indispensable for pellicle formation in *Toxoplasma gondii* (Ke et al. 2014; Martins-Duarte et al.
74 2016).

75 Less is known about the actual metabolic functions of plastids in other non-photosynthetic algal
76 lineages. Many of them have a similar metabolic capacity as the apicoplast (Sanchez-Puerta et al. 2007;
77 Slamovits and Keeling 2008; Fernández Robledo et al. 2011), and some house an even more complex
78 metabolism that includes amino acid biosynthesis and carbohydrate metabolism pathways (Borza et al.
79 2005; Pombert et al. 2014; Smith and Lee 2014). Until recently, IPP synthesis seemed to be a process
80 conserved even in the most reduced relic plastids, such as the genome-lacking plastids of certain alveolates
81 (Matsuzaki et al. 2008; Janouškovec et al. 2015), but non-photosynthetic plastids lacking the characteristic
82 plastidial (MEP or DOXP) pathway of IPP biosynthesis are now known (Kamikawa et al. 2017; Graupner
83 et al. 2018; Dorrell et al. 2019). Thus, there is generally a metabolic reason for a plastid retention, although
84 the cases of plastid dependency differ between lineages.

85 An interesting group to study non-photosynthetic plastids are the euglenophytes. Like their prime
86 representative *Euglena gracilis*, most euglenophytes are mixotrophs containing a complex three-
87 membranes-bound photosynthetic plastid derived from a green alga belonging to Pyramimonadales (Turmel
88 et al. 2009; Leander et al. 2017; Jackson et al. 2018). However, non-photosynthetic mutants of *E. gracilis*,

89 induced by an antibiotic or mutagenic treatment, are often capable of heterotrophic living even after
90 presumed plastid loss (reviewed in Krajčovič et al. 2002; Hadariová et al. 2018). This might be enabled by
91 metabolic independence of the *E. gracilis* cell on the plastid. For instance, *E. gracilis* was shown to possess
92 two parallel haem synthesis pathways, one located in the mitochondrion/cytosol and another in the plastid
93 (Weinstein and Beale 1983). In the long run, this redundancy might predestine one of the pathways for loss,
94 provided that the other can efficiently supply the end-product to all compartments requiring it (Kořený and
95 Oborník 2011; Cihlář et al. 2016). On the other hand, localization of an essential pathway or its part
96 exclusively into the plastid will impose essentiality of the organelle as such.

97 Several lineages of euglenophytes independently became secondarily heterotrophic, but evidence
98 for the presence of a plastid has been provided only for *Euglena longa* (originally called *Astasia longa*), a
99 close relative of *E. gracilis* (Marin et al. 2003; Nudelman et al. 2003). The documentation of the organelle
100 at the cytological level is spurious (Webster et al. 1967; Kivic and Vesk 1974; Hachtel 1996), but the
101 complete plastid genome sequence was reported nearly two decades ago (Gockel and Hachtel 2000). As
102 expected, it lacks all the photosynthesis-related genes, except for *rbcL* encoding the large subunit of
103 ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO). We recently documented the existence of a
104 nucleus-encoded small subunit (RBCS) of the *E. longa* RuBisCO enzyme synthesized as a precursor
105 polyprotein, but its processing into monomers could not be demonstrated (Záhonová et al. 2016). The
106 physiological role of the *E. longa* RuBisCO and the whole plastid remains unknown, although indirect
107 evidence suggests that the plastid is essential for *E. longa* survival (Siemeister et al. 1990a, b; Gockel et al.
108 1994; Hadariová et al. 2017).

109 To provide a key resource for investing the biology of *E. longa* and its plastid, we recently generated
110 a transcriptome assembly for this species and demonstrated that it is complete and directly comparable to
111 the transcriptome assemblies reported for *E. gracilis* (Záhonová et al. 2018). Evaluation of a set of high-
112 confidence candidates for plastid-targeted proteins enabled us to conclude that nucleus-encoded plastidial
113 proteins in *E. longa* employ N-terminal targeting presequences of the same two characteristic classes, as
114 known from *E. gracilis*. The *E. longa* transcriptome revealed various unusual features of the plastid
115 biogenesis and maintenance machinery shared with photosynthetic euglenophytes, but also supported the
116 lack of the photosynthesis-related machinery and suggested specific reductions of several plastidial house-
117 keeping functions presumably reflecting the loss of photosynthesis (Záhonová et al. 2018). However, the
118 repertoire of anabolic and catabolic pathways localized to the *E. longa* colourless plastid has not been
119 investigated and is the subject of the present paper.

120 To chart the main paths of the metabolic map of the *E. longa* plastids, we searched for homologs of
121 enzymes underpinning pathways known from plastids of other species. The reconstruction was greatly
122 facilitated by the recent characterization of the *E. gracilis* plastid metabolic network based on a proteomic

123 analysis of the organelle (Novák Vanclová et al. 2019). N-terminal regions of the candidates were evaluated
124 for characteristics of presequences predicting a specific subcellular localization to distinguish those likely
125 representing plastid-targeted proteins from enzymes located to other parts of the cell. Some of the
126 bioinformatics predictions were further tested by biochemical analyses. Our study provides the first
127 comprehensive view of a non-photosynthetic secondary plastid of green-algal origin and shows that the
128 metabolic capacity of the *E. longa* plastid is strikingly different from those of the apicoplast and other relic
129 plastids characterized in sufficient detail.
130
131
132 **Results and Discussion**
133 ***E. longa* plastid lacks the MEP pathway of IPP biosynthesis, yet has kept the production of tocopherol**
134 **and a phylloquinone derivative**
135 Two parallel pathways of IPP biosynthesis exist in *E. gracilis* (Kim et al. 2004): the mevalonate (MVA)
136 pathway localized to the mitochondrion (first three enzymes) and the cytosol (the rest), and the plastid-
137 localized 2-C-methyl-D-erythritol (MEP) pathway, but only enzymes of the MVA pathway were found in
138 *E. longa* (Table S1, Fig. 1A). *E. longa* thus joins the group of recently discovered plastid-bearing eukaryotes
139 lacking the MEP pathway, namely the colourless diatom *Nitzschia* sp. NIES-3581 (Kamikawa et al. 2017)
140 and various colourless chrysophytes (Graupner et al. 2018; Dorrell et al. 2019). In contrast, the plastid-
141 localized MEP pathway in apicomplexans and related alveolates (i.e. Myzozoa) is essential as a source of
142 precursors for the synthesis of all cellular isoprenoids, since the cytosolic MVA pathway was lost from this
143 group (Janouškovec et al. 2015; Waller et al. 2016). The retention of the MEP pathway in the colourless
144 plastids of diverse non-photosynthetic chlorophytes (Figueroa-Martinez et al. 2015) is similarly explained
145 by the loss of the MVA pathway in this group (Lohr et al. 2012).
146 The MEP pathway in *E. gracilis* provides precursors for the synthesis in the plastid of terpenoid
147 compounds connected to photosynthesis, namely carotenoids and plastoquinone (Kim et al. 2004; Novák
148 Vanclová et al. 2019). As expected, the respective enzymes are all missing from *E. longa*. However, we
149 were surprized to find out that the *E. longa* plastid appears still involved in terpenoid metabolism,
150 specifically in its phytol branch. Photosynthetic eukaryotes generally produce three types of phytol
151 derivatives, tocopherols (vitamin E), phylloquinone (PhQ; vitamin K_1) and chlorophyll (its phytyl chain),
152 starting with a common precursor phytyl-PP, which is (directly or indirectly via salvage of phytol liberated
153 by chlorophyll degradation) made by reduction of geranylgeranyl-PP synthesized in the plastid by the MEP
154 pathway (Gutbrod et al. 2019). *E. gracilis* proved to be unusual not only because of lacking the conventional
155 geranylgeranyl-PP reductase (Novák Vanclová et al. 2019), but also for making phytol from precursors
156 provided by the MVA pathway (Disch et al. 1998; Kim et al. 2004). The route of phytol synthesis from the

157 cytosolic isoprenoids is currently unknown, though it was proposed that phytyl-PP is synthesized in the *E.*
158 *gracilis* plastid exclusively by step-wise phosphorylation of phytol by phytol kinase (VTE5) and phytyl
159 phosphate kinase (VTE6), which are enzymes normally employed for recycling phytol from chlorophyll
160 degradation (Novák Vanclová et al. 2019). This scheme is supported by the fact that *E. longa* has retained
161 VTE5 as well as VTE6, both proteins being highly similar to their *E. gracilis* orthologs and exhibiting
162 putative plastid targeting presequences (Fig. S1; Table S1). Since *E. longa* lacks chlorophyll, these two
163 enzymes must have a function independent of phytol recycling.

164 In addition of its role in chlorophyll synthesis, phytyl-PP is used by *E. gracilis* to make tocopherols
165 and a PhQ derivative, 5'-monohydroxyphylloquinone (OH-PhQ; Ziegler et al. 1989; Watanabe et al. 2017;
166 Novák Vanclová et al. 2019). All four enzymes mediating synthesis of α-tocopherol from phytyl-PP and
167 homogentisate were identified in *E. gracilis* and are localized to its plastid (Novák Vanclová et al. 2019).
168 Interestingly, orthologs of all four enzymes are found in *E. longa* as well, all with a typical plastid-targeting
169 presequence or at least with the N-terminal region being highly similar to the *E. gracilis* counterpart (Table
170 S1), consistent with their presumed plastidial localization (Fig. 1A). Homogentisate itself is apparently made
171 outside the plastid, most likely in the mitochondrion, as the enzyme responsible for its synthesis (4-
172 hydroxyphenylpyruvate dioxygenase) is not found in the *E. gracilis* plastid proteome and the respective
173 proteins have a predicted mitochondrial transit peptide in both *E. gracilis* and *E. longa* (Table S1). Our
174 analysis thus predicts that like its photosynthetic cousin, *E. longa* produces α-tocopherol. To test this
175 directly, we analysed extracts from *E. longa* by means of HPLC-MS/MS. For comparison, we employed
176 samples from *E. gracilis* grown at two different conditions (in light and in darkness). Tocopherols were
177 detected in both species (Fig. 1B), with α-tocopherol being the dominant form present in equivalent amounts
178 in all three samples (Fig. 1C). The signals of β- and/or γ-tocopherol (indistinguishable by our method) and
179 of δ-tocopherol suggest that tocopherol cyclase, and possibly also tocopherol O-methyltransferase, of both
180 *Euglena* species can process substrates with or without the 3-methyl group on the benzene ring (Fig. S2).

181 The synthesis of OH-PhQ in *E. gracilis* is understood only partially, with only three enzymes of the
182 pathway previously identified at the molecular level: the large multifunctional protein PHYLLO, apparently
183 localized to the cytosol and catalysing the first four steps leading to o-succinylbenzoate; MenA catalysing
184 phytylation of dihydroxynaphthoate localized in the plastid; and MenG (demethylnaphthoquinone
185 methyltransferase), possessing a typical N-terminal plastid-targeting presequence but not directly confirmed
186 as plastidial by proteomics (Novák Vanclová et al. 2019). Strikingly, *E. longa* expresses homologs of these
187 three *E. gracilis* proteins, although at low levels resulting in a low RNA-seq read coverage and thus
188 incomplete or inaccurate assembly of two of the respective transcripts (the PHYLLO sequence split into
189 several contigs, the MenA sequence with a frameshift corrected by RT-PCR; see Materials and Methods).
190 Nevertheless, N-terminal parts of both sequences were intact and confirmed the same subcellular

191 localization as in *E. gracilis* (Fig. 1A, Table S1). In agreement with these insights, OH-PhQ could be
192 detected in an extract from the *E. longa* culture (Fig. 1D, Fig. S3), although its abundance was smaller
193 compared to that in *E. gracilis* by an order of magnitude (Fig. 1E).

194 These findings document another unexpected function of the *E. longa* plastid, OH-PhQ synthesis,
195 albeit the plastid-associated part of the pathway cannot be presently reconstructed in full detail. One
196 uncertainty concerns the middle steps of the pathway, since like *E. gracilis* (see Novák Vanclová et al.
197 2019), *E. longa* also lack homologs of the conventional enzymes that are responsible for converting *o*-
198 succinylbenzoate to dihydroxynaphthoate and localized (at least in eukaryotes studied in sufficient detail)
199 in the peroxisome (Cenci et al. 2018). In contrast, tentative evidence was presented for the association of
200 the respective enzyme activities (presumably corresponding to enzymes non-orthologous to the
201 conventional ones) with the plastid envelope in *E. gracilis* (Seeger and Bentley 1991), raising a possibility
202 of a similar arrangement in *E. longa*. Secondly, the molecular identity of the putative hydroxylase catalysing
203 the final step of OH-PhQ synthesis is unknown, so its plastidial localization in *E. gracilis* or *E. longa* cannot
204 be ascertained. It is, nevertheless, likely given the fact that OH-PhQ is primarily needed in the plastid, at
205 least in photosynthetic eukaryotes (Ziegler et al. 1989). Thirdly, a previously unknown step – reduction of
206 the naphthoquinone ring – was recently demonstrated as a prerequisite for the reaction catalysed by MenG
207 to proceed in plants and cyanobacteria (Fatihi et al. 2015). The respective reductase is well conserved among
208 diverse plant and algal groups as well as cyanobacteria (Cenci et al. 2018), but we did not identify its close
209 homologs in any of the euglenophytes transcriptome assemblies, suggesting that euglenophytes employ an
210 alternative enzyme that yet needs to be characterized.

211 *E. longa* seems to be the first eukaryote with a non-photosynthetic plastid documented to have
212 retained the pathways (or final parts thereof) for tocopherols and OH-PhQ synthesis localized in the
213 organelle. The presence of tocopherols in *E. longa* is, however, not that surprising, as their function is not
214 restricted to photosynthetic tissues in plants and were detected also in bleached (i.e. non-photosynthetic) *E.*
215 *gracilis* mutants (Maeda and DellaPenna 2007; Watanabe et al. 2017). As potent lipophilic antioxidants
216 involved especially in membrane protection from damage caused by lipid peroxidation, tocopherols might
217 be used by *E. longa* as part of its protective mechanisms against reactive oxygen species generated by the
218 action of mitochondria and peroxisomes. The retention of OH-PhQ synthesis in *E. longa* may seem more
219 puzzling, given the fact that the best-established role of (OH-)PhQ in plants and algae is its functioning as
220 an electron carrier within the photosystem I (Ziegler et al. 1989; Brettel 1997). PhQ was additionally
221 proposed to serve as an electron acceptor in the process of disulfide bond formation in thylakoid lumenal
222 proteins required for proper function of photosystem II (Furt et al. 2010; Karamoko et al. 2011). A homolog
223 of the respective oxidoreductase (LTO1) exists in *E. gracilis* (Table S1), but not in the transcriptome data
224 from *E. longa*, consistent with the lack of photosystem II, and perhaps thylakoids entirely in the latter

225 species. Interestingly, PhQ was detected in the plasma membrane in plant tissues and proposed to be
226 involved in photosynthesis-unrelated redox processes at the cell surface (Lochner et al. 2003; Schopfer et
227 al. 2008). PhQ may even be synthesised directly in the plasma membrane, as a recent report from plants has
228 documented the existence of alternative forms of the terminal enzymes of PhQ biosynthesis that result from
229 alternative splicing and are localized to the plasma membrane rather than the plastid (Gu et al. 2018).
230 However, the protein sequences of both MenA and MenG enzymes in *E. longa* carry typical plastid-targeting
231 presequences, so they are unlikely to operate in the plasma membrane. We thus propose that OH-PhQ in *E.*
232 *longa* is involved in a hitherto uncharacterized, photosynthesis-unrelated plastid-resident process.
233

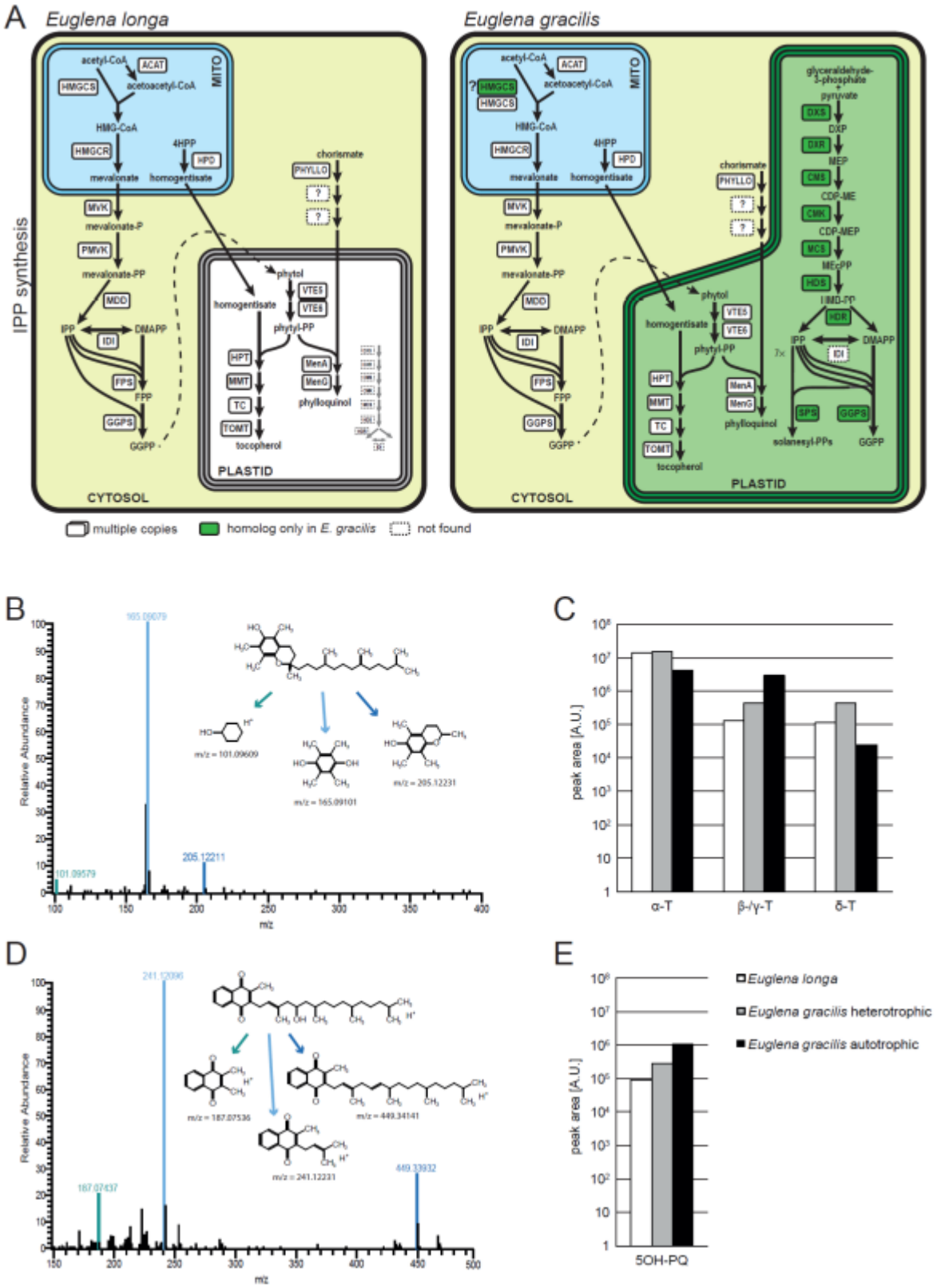


Figure 1: IPP and terpenoid-quinone biosynthesis in *E. longa* and its phototrophic relative *E. gracilis*.

236 **A:** Schematic comparison of the localization and evolutionary origin of enzymes (see colour-coding
237 graphical legend below the "cells"). Abbreviations, IPP synthesis: ACAT – acetyl-CoA acetyltransferase,
238 CDP-ME – 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol, CDP-MEP – 2-phospho-CDP-ME, CMK –
239 CDP-ME kinase, CMS – CDP-ME synthase, DMAPP – dimethylallyl diphosphate, DXP – 1-deoxy-D-
240 xylulose 5-phosphate, DXR – DXP reductase, DXS – DXP synthase, FPP – farnesyl siphosphate synthase,
241 GGPS – geranylgeranyl-diphosphate synthase, HDR – HMB-PP reductase, HDS – HMB-PP synthase,
242 HMB-PP – 4-hydroxy-3-methylbut-2-en-1-yl diphosphate, HMG-CoA – 3-hydroxy-3-methylglutaryl-CoA,
243 HMGCR – HMG-CoA reductase, HMGCS – HMG-CoA synthase, IDI – isopentenyl-diphosphate delta-
244 isomerase, MCS – MEcPP synthase, MDD – mevalonate-diphosphate decarboxylase, MEcPP – 2-C-methyl-
245 D-erythritol 2,4-cyclodiphosphate, MEP – 2-C-methyl-D-erythritol 4-phosphate, MVK – mevalonate
246 kinase, PMVK – phosphomevalonate kinase, PPS – unspecified polyprenyl-diphosphate synthase, ? –
247 unclear substrate; Terpenoid-quinone synthesis: 4HPP – 4-hydroxyphenylpyruvate, HPD –
248 hydroxyphenylpyruvate dioxygenase, HPT – homogentisate phytyltransferase, MMT – MPBQ/MPSQ
249 methyltransferase, TAT – tyrosine aminotransferase, TC – tocopherol cyclase, TMT – tocopherol-O-
250 methyltransferase, VTE5 – phytyl kinase, VTE6 – phytyl-phosphate kinase. **B:** MS/MS spectrum record of
251 *E. longa* α-tocopherol and the proposed fragmentation pattern in positive ionization mode (inset).
252 Monoisotopic masses of particular fragments were obtained by simulation in Xcalibur software. **C:**
253 Semiquantitative comparison of tocopherol species in *E. longa*, heterotrophically (dark) grown *E. gracilis*
254 and autotrophic *E. gracilis*. **D and E:** MS/MS spectrum record of *E. longa* 5-hydroxyphylloquinone and the
255 proposed fragmentation pattern in positive ionization mode (inset); semiquantitative comparison of 5-
256 hydroxyphylloquinone.

257

258 ***E. longa* plastid plays a limited role in the metabolism of nitrogen-containing compounds**

259 Some of the apparent oddities of the *E. longa* plastid do not stem from the loss of photosynthesis in this
260 species but reflect unusual features of the plastid in euglenophytes in general. These particularly concern
261 plastid functions in the metabolism of nitrogen-containing compounds. Plastid is commonly involved in
262 nitrogen assimilation due to housing nitrite reductase (Giordano and Raven 2014; Sanz-Luque et al. 2015),
263 but it was established a long time ago that *E. gracilis* cannot assimilate nitrate or nitrite (Oda et al. 1979;
264 Kitaoka et al. 1989) and, accordingly, no nitrite reductase can be identified in the transcriptome data from
265 this species and *E. longa*. The plastids of both *Euglena* species apparently also lack the enzymes working
266 immediately downstream, i.e. glutamine synthetase and glutamine oxoglutarate aminotransferase (the
267 GS/GOGAT system common in plastids of other groups; Fernandez and Galvan 2008; Dagenais-
268 Bellefeuille and Morse 2013), indicating that the plastids rely on the import of organic nitrogen from other
269 parts of the cell, like recently proposed for the plastid in chromerids (Füssy et al. 2019).

270 One of the most surprising insights of the recent proteomics-aided analysis of the *E. gracilis* plastid
271 metabolism was the paucity of pathways concerning amino acids (Novák Vanclová et al. 2019). *E. longa* is
272 apparently even more extreme in this regard, because it lacks counterparts of the (predicted or proteomically
273 verified) plastid-targeted forms of serine biosynthesis enzymes found previously in *E. gracilis*, i.e.
274 phosphoglycerate dehydrogenase and phosphoserine phosphatase. Thus, we could localize only two
275 elements of amino acid biosynthesis pathways to the *E. longa* plastid (Fig. S4): serine/glycine
276 hydroxymethyltransferase, whose obvious role is to provide the one-carbon moiety for formylmethionyl-
277 tRNA synthesis required for the plastidial translation; and one of the multiple isoforms of cysteine synthase
278 A, which (like in *E. gracilis*) apparently relies on O-acetyl-L-serine synthesized outside the plastid, due to
279 the absence of a discernible plastid-localized serine O-acetyltransferase (see Novák Vanclová et al. 2019,
280 and Table S2). This is not due to a general reduction of amino acid metabolism in *E. longa* or the
281 incompleteness of the sequence data, as its transcriptome assembly includes homologs of enzymes required
282 for the synthesis of all 20 proteinogenic amino acids, but the respective proteins have predicted localization
283 in compartments other than the plastid (Table S2).
284 Amino acids in the plastid are not only substrates of protein synthesis, but also serve as precursors
285 or nitrogen donors for the synthesis of various other compounds (Moffatt and Ashihara 2002; Gerdes et al.
286 2012). One such pathway, described in detail in the subsequent section, leads to tetrapyrroles. In contrast,
287 the spectrum of reactions related to the metabolism of other nitrogen-containing cofactors or their precursors
288 (B vitamins) is very limited in the plastids of both *Euglena* species. We identified only one such candidate
289 in *E. longa* – the reaction of vitamin B6 salvage catalysed by pyridoxamine 5'-phosphate oxidase, whereas
290 *E. gracilis* additionally expresses two plastid-targeted isoforms of pyridoxine 4-dehydrogenase (Table S3).
291 *De novo* synthesis or salvage of purines and pyrimidines is also absent from the plastid of both *Euglena*
292 species, except for one step present in *E. gracilis* but apparently not *E. longa*: the former exhibits two forms
293 of CTP synthase, one presumably cytosolic and another with a plastid-targeting presequence and found in
294 the plastid proteome reported by Novák Vanclová et al. 2019, whereas *E. longa* expresses an ortholog of
295 only the cytosolic version (Table S3). The lack of a separate plastidial CTP source in *E. longa* may reflect
296 the presumably lower magnitude of RNA synthesis (and possibly also less extensive phospholipid synthesis
297 requiring CTP in the reaction catalysed by CDP-diacylglycerol pyrophosphatase) in its plastid. Finally, *E.*
298 *longa* possesses an ortholog of an enzyme involved in the synthesis of polyamines (spermidine synthase)
299 found in the plastid proteome of *E. gracilis* (Novák Vanclová et al. 2019), but (in contrast to the *E. gracilis*
300 protein) its N-terminus does not fit the characteristics of a plastid-targeting presequence, suggesting that its
301 subcellular localization may be outside of the plastid (Fig. S4). However, we found out that both *E. longa*
302 and *E. gracilis* have another version of this enzyme with an obvious plastid-targeting-like presequence in
303 both species (Table S3), so we cannot rule out the possibility that polyamines are produced in the *E. longa*

304 plastid after all.
305
306 **A residual tetrapyrrole biosynthesis pathway of unclear function is retained in the *E. longa* plastid**
307 Most plastid-bearing eukaryotes synthesize protoporphyrin IX, the common precursor of haem and
308 chlorophyll, via a single pathway wholly or mostly localized to the plastid (Oborník and Green 2005; Cihlář
309 et al. 2016; Füssy and Oborník 2017). *E. gracilis* is one of the known exceptions, because it possesses two
310 independent protoporphyrin synthesis pathways, a mitochondrial-cytosolic and a plastid one (Fig. 2A;
311 Weinstein and Beale 1983; Kořený and Oborník 2011; Lakey and Triemer 2016). Whereas all enzymes of
312 the plastid pathway were identified at the sequence level, previous studies left some gaps in the enzyme
313 assignment to the different steps of the mitochondrial-cytosolic pathway (Kořený and Oborník 2011; Lakey
314 and Triemer 2016). The new sequence data from *E. gracilis* has now enabled us to identify all three missing
315 enzymes of the mitochondrial-cytosolic pathway; specifically novel, apparently cytosolic, isoforms of
316 uroporphyrinogen-III synthase (UROS) and uroporphyrinogen decarboxylase (UROD), and two variants of
317 oxygen-independent coproporphyrinogen oxidase (CPOXi), one cytosolic and one perhaps localized to the
318 mitochondrion. In addition, we found a third UROS homolog that carries a plastid-targeting presequence
319 and constitutes a divergent sister group to the previously known euglenophyte plastidial UROS isoform
320 (Fig. S5; Table S4). Indeed, it was detected in the experimentally determined plastid proteome, together
321 with the previously known plastidial UROS isoform (protein IDs 16898 and 15143 in Novák Vanclová et
322 al. 2019), but the functional significance of the plastidial UROS duplication remains unclear.
323 Our transcriptome data from *E. longa* revealed orthologs of all *E. gracilis* enzymes of the
324 mitochondrial-cytosolic pathway with the same predicted subcellular localization, barring a single enzyme
325 with an incomplete N-terminal sequence precluding confident prediction (Fig. 2A; Table S4). In contrast,
326 orthologs of only six of the *E. gracilis* plastid-targeted enzymes could be identified in the transcriptome of
327 *E. longa*: aminolevulinic acid dehydratase (ALAD), the two UROS isoforms, one of the UROD isoforms,
328 one isoform of CPOXi, and protoporphyrinogen oxidase (PPOX). The first three proteins have putative
329 plastid-targeting presequences, much like their *E. gracilis* orthologs. However, the respective CPOXi
330 protein has a strong mitochondrial targeting presequence, whereas the UROD protein is devoid of any
331 presequence and the presequence of the PPOX protein is markedly shorter and lacks the characteristics of a
332 plastid-targeting signal (Fig. 2B; Table S4). This suggests that only ALAD and the two UROS isoforms are
333 targeted to the plastid in *E. longa*, whereas the other enzymes have been retargeted to the cytosol or
334 mitochondrion in the evolution of the *E. longa* lineage. Strikingly, a putative *E. longa* homolog of the *E.*
335 *gracilis* plastid-localized porphobilinogen deaminase (PBGD), an enzyme for the reaction between those
336 catalysed by ALAD and UROS, was previously detected at both the RNA (by a Northern blot) and protein
337 level (by an immunoblot; Shashidhara and Smith 1991), but we could not identify the respective transcript

338 in the *E. longa* transcriptome data even when the raw reads were searched.

339 Intact isolated *E. gracilis* plastids are capable of chlorophyll synthesis from externally supplied 5-
340 aminolevulinic acid (ALA; Gomez-Silva et al. 1985), so it is possible that the *E. longa* plastid can
341 analogously import ALA from the cytoplasm, thus providing a substrate for the plastid-localized ALAD
342 isoform (Fig. 2A). The surprising absence of the previously documented plastid-localized PBGD in our
343 transcriptome data may be potentially accounted for by the lack of expression of the respective enzyme at
344 culture conditions employed by us. Hence, the pathway may theoretically proceed up to uroporphyrinogen
345 III (the product of the UROS enzyme) in the *E. longa* plastid, but the absence of enzymes for further
346 processing of this compound towards haem and/or chlorophyll (consistent with the lack of photosystems
347 and the cytochrome b_6f complex) indicates that it is consumed by another process.

348 Uroporphyrinogen III is indeed at the beginning of a separate branch converting it in three steps to
349 sirohaem (Dailey et al. 2017), which serves as a cofactor of several enzymes, including nitrite reductase and
350 sulfite reductase (Tripathy et al. 2010). The former enzyme is missing from the *Euglena* spp. (see above),
351 but sulfite reductase is present in both *E. longa* and *E. gracilis* (Table S4) and was detected in the plastid
352 proteome of the latter species (Novák Vanclová et al. 2019). Hence, sirohaem synthesis in the plastid to
353 support the function of sulfite reductase is an attractive hypothesis that might explain the retention of some
354 of the tetrapyrrole synthesis enzymes in the *E. longa* plastid. However, analyses of the transcriptome data
355 alone do not provide any clear answer as to whether this hypothesis is valid. Of the various alternative
356 enzymes mediating sirohaem synthesis in various taxa (Tripathy et al. 2010), only two have confidently
357 identified homologs in both *Euglena* spp. (Table S4). One of them is the multifunctional protein CysG,
358 which catalyses all reactions from uroporphyrinogen III to sirohaem, but the respective *Euglena* homolog
359 is devoid of any presequence in either species and is thus most likely cytosolic (Table S4). The other is a
360 stand-alone version of the enzyme (uroporphyrinogen-III C-methyltransferase) catalysing the first step of
361 sirohaem synthesis, called CobA or SirA (Table S4). The *Euglena* homologs exhibit a predicted signal
362 peptide, which is however immediately followed by a region belonging to the mature enzyme, so it is unclear
363 whether this protein is routed to the plastid (indeed it is not among the proteomically confirmed plastid
364 protein in *E. gracilis*). Hence, the role of the enzymes of the tetrapyrrole biosynthesis pathway in the *E.*
365 *longa* plastid remains unexplained.
366

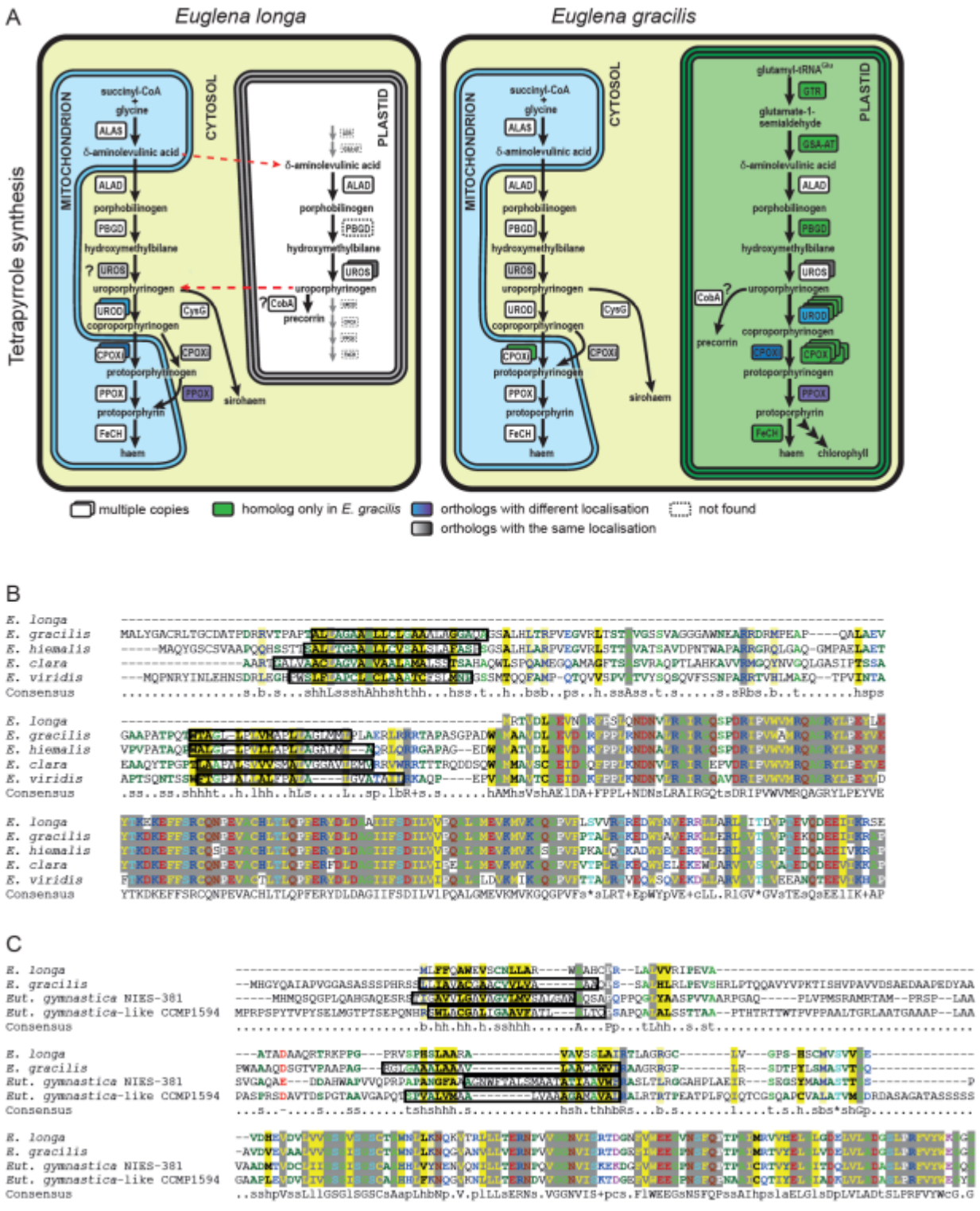


Figure 2: Tetrapyrrole biosynthesis in *E. longa* and *E. gracilis*. A: Schematic comparison of the localization and evolutionary origin of enzymes (see colour-coding graphical legend). Abbreviations: ALAD – delta-aminolevulinate dehydrogenase, ALAS – delta-aminolevulinate synthase, CobA –

371 uroporphyrinogen-III C-methyltransferase, CPOX – coproporphyrinogen III oxidase, CPOXi – oxygen-
372 independent CPOX, CysG – trifunctional enzyme of sirohaem synthesis (see text), FeCH – ferrochelatase,
373 GSA-AT – glutamate semialdehyde-aminomutase, GTR – glutamyl-tRNA reductase, PBGD –
374 porphobilinogen deaminase, PPOX – protoporphyrinogen oxidase, UROD – uroporphyrinogen
375 decarboxylase, UROS – uroporphyrinogen-III synthase. **B-C:** Sequence alignment of UROD (B) and PPOX
376 (C) presequences, showing the loss of the plastid-targeting motifs in *E. longa*. These targeting motifs
377 represented by transmembrane domains are marked with frames in other euglenophyte sequences.

378

379 ***E. longa* plastid does not make fatty acids but maintains phospholipid and glycolipid synthesis**

380 Eukaryotes synthesize even-chain fatty acids by a single large multi-modular fatty acid synthase I (FASI)
381 in the cytosol or by a multi-enzyme type II fatty acid synthesis complex in the plastid. *E. gracilis* possesses
382 both systems (Zimorski et al. 2017). In contrast, *E. longa* encodes only a homolog of the FASI enzyme
383 (Table S5), whereas enzymes of the type II fatty acid synthesis are absent (Fig. 3A) except for the acyl
384 carrier protein (ACP). The loss of plastid-localized fatty acid synthesis in *E. longa* is not without precedent,
385 as it has been also reported for the apicoplast of *Theileria parva* (which is fully dependent on fatty acid
386 supply from host), the plastid of *Perkinsus marinus* (Janouškovec et al. 2015) and the chrysophyte
387 "*Spumella*" sp. NIES-1846 (Dorrell et al. 2019). Nevertheless, the *E. longa* plastid has kept plastid-targeted
388 enzymatic steps downstream of fatty acid synthesis. These include ACP and 4'-phosphopantetheinyl
389 transferases (or holo-ACP synthase) crucial for the synthesis of an active form of ACP, which serves as a
390 carrier of acyl chains in phospholipid and glycolipid biosynthesis (Lambalot and Walsh 1995). Next, *E.*
391 *longa* possesses predicted plastid-targeted homologs of acyl-ACP synthetases (presumably activating fatty
392 acids imported into the plastid from outside) and all enzymes required for the synthesis of phosphatidic acid
393 (PA) and its subsequent conversion to phosphatidylglycerol (PG) (Fig. 3A; Table S5). It is worth noting that
394 *E. longa* also has a parallel, plastid-independent, route of phosphatidylglycerol synthesis (Table S6).

395 No other reactions of phospholipid synthesis or decomposition beyond PG synthesis seem to operate
396 in the *E. longa* plastid, but interestingly, we found homologs of the enzymes of the synthesis of galactolipids
397 monogalactosyldiacylglycerol (MGDG) and digalactosyldiacylglycerol (DGDG) in the *E. longa*
398 transcriptome (Fig. 3A, Table S5). All the proteins have a predicted N-terminal plastid-targeting
399 presequence, consistent with the plastidial localization of galactolipid synthesis in all other plastid-bearing
400 eukaryotes studied so far (Yuzawa et al. 2012). In support of the bioinformatic predictions, both MGDG
401 and DGDG could be detected in lipid extracts from *E. longa* and *E. gracilis*, although galactolipid levels
402 were significantly lower in *E. longa* than in the control sample of *E. gracilis* (Fig. 3B). The presence of
403 DGDG was further confirmed by immunofluorescence using an anti-DGDG antibody, which showed
404 DGDG to be present in small foci in the *E. longa* cells (white arrowheads in Fig. 3C), presumably

405 representing individual small plastids. In comparison, most of the volume of the photosynthetic *E. gracilis*
406 cells was stained, whereas the negative control, the primary osmotrophic (i.e. plastid-lacking) euglenoid
407 *Rhabdomonas costata*, did not stain at all.

408 The presence of galactolipids in plastids is generally being explained by their essentiality for the
409 proper functioning of the photosynthetic apparatus, but this view has been challenged by demonstration that
410 cyanobacterial mutants lacking galactolipids can photosynthesize normally (Awai et al. 2001). A
411 photosynthesis-independent role of galactolipids in plastid biology was proposed, too. Transit peptides of
412 plastid-targeted proteins exhibit affinities for MGDG and DGDG in the plastid envelope (Pinnaduwage and
413 Bruce 1996), suggesting a direct role of these lipids in plastid protein import. The photosynthesis-
414 independent role of galactolipids is indicated not only by their presence in the *E. longa* plastid documented
415 here, but also by previous reports from other non-photosynthetic algae, including the diatom *Nitzschia alba*
416 (Anderson et al. 1978) and the chlorophyte *Prototheca wickerhamii* (Borza et al. 2005), and from non-
417 photosynthetic tissues of plants (Awai et al. 2001; Kobayashi 2016). On the other hand, the apicoplast (Botté
418 et al. 2008; Botté et al. 2013) and most likely also the relic plastid of *Helicosporidium* (based on our analysis
419 of the respective genome sequence data generated by Pombert et al. 2014) lack galactolipid synthesis
420 completely. The reason for the differential retention of galactolipids in different colourless plastids remains
421 unclear.

422 In addition to MGDG and DGDG, we identified in samples from both *Euglena* spp. another
423 common glycolipid characteristic for plastids, sulfoquinovosyldiacylglycerol (SQDG) (Fig. 3B) (Hori et al.
424 2016). The presence of SQDG in *E. longa* is also not unprecedented among non-photosynthetic plastid-
425 bearing eukaryotes; see, e.g., its documented occurrence in the diatom *N. alba* or the dinoflagellate *Oxyrrhis*
426 *marina* (Anderson et al. 1978; Goddard-Borger and Williams 2017; Yoon et al. 2017). Of the three enzymes
427 of SQDG biosynthesis normally localized to the plastid, we found in both *Euglena* species only that
428 catalysing the final step (sulfoquinovosyltransferase; Fig. 3A). Interestingly, the standard eukaryotic version
429 of the enzyme, SQD2, is present only in *E. gracilis*, but both species proved to share another isoform
430 phylogenetically affiliated to bacterial SqdX version of the enzyme (Fig. 4). To the best of our knowledge,
431 this is the first encounter of SqdX in any eukaryote. The presence of SQD2 only in *E. gracilis* probably
432 relates to specific needs of its photosynthetic plastid. Indeed, *E. gracilis* contains a much larger amount of
433 SQDG compared to *E. longa* (Fig. 3B) and the profile of esterified fatty acids differs between the two species
434 (*E. longa* lacks SQDG forms with unsaturated longer chains; Table S7).

435 Details of the synthesis of the saccharide moieties of glycolipids in *E. longa* are also worth
436 considering (Fig. 3A). *E. longa* exhibits an ortholog of the *E. gracilis* UDP-glucose epimerase previously
437 identified in the plastid proteome (Table S5), explaining the source of UDP-galactose for galactolipid
438 synthesis. This enzyme seems to have been acquired by euglenophytes from a bacterial source (Fig. S6),

439 but, interestingly, *E. gracilis* encodes another enzyme, which corresponds to the plastidial UDP-glucose
440 epimerase, also called PHD1, known from plants and various algae (Li et al. 2011). The *E. gracilis* PHD1
441 possesses a predicted plastid-targeting presequence (Table S5) and is thus also likely plastidial (although
442 this is not confirmed by proteomic data reported by Novák Vanclová et al. 2019). This putative redundancy
443 in UDP-galactose is apparently not shared by *E. longa*, possibly because of a presumably much lower need
444 for galactolipid synthesis (Fig. 3B). The origin of the SQDG precursor UDP-sulfoquinovose in *E. longa*
445 remains obscure, because as noticed before (Novák Vanclová et al. 2019), euglenophytes in general lack the
446 conventional UDP-sulfoquinovose synthase SQD1/SqdB and probably employ an alternative, unrelated
447 enzyme. UDP-glucose, i.e. the common precursor of both UDP-galactose and UDP-sulfoquinovose, is most
448 likely imported from the cytoplasm, owing to the absence in *E. longa* of a candidate plastid-targeted version
449 of any of the relevant enzymes (UDP-glucose/UDP-sugar pyrophosphorylase).
450

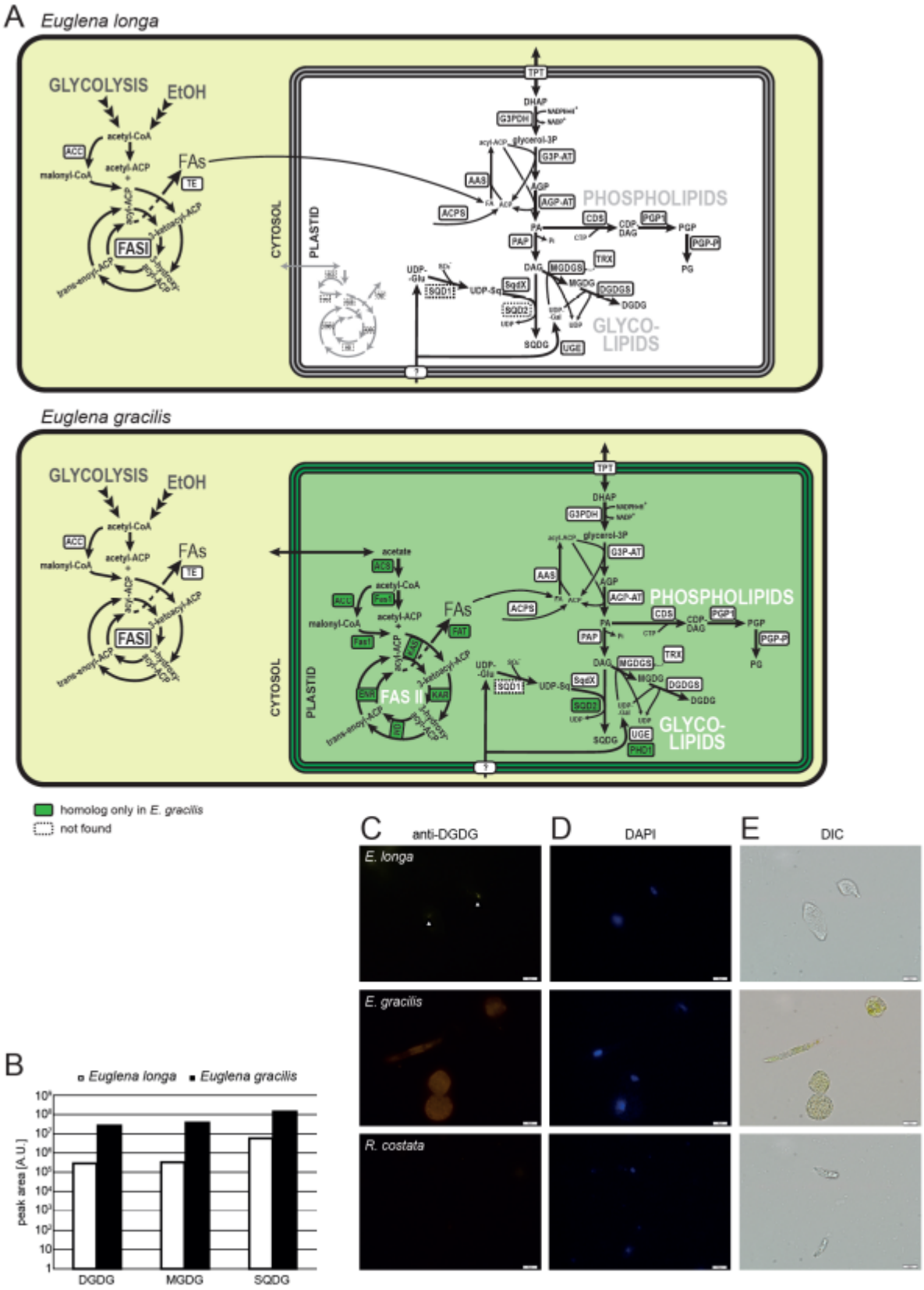


Figure 3: Fatty acid and lipid biosynthesis in *E. longa* and *E. gracilis*. A: Schematic comparison of the

453 localization and evolutionary origin of enzymes. Abbreviations, fatty acid synthesis: ACC – acetyl-CoA
454 carboxylase, ACS – acetyl-CoA synthetase, ENR – enoyl-CoA reductase, Fas1 – malonyl-CoA/acetyl-CoA:
455 ACP transacylase, FASI – type I fatty acid synthase, FAT – fatty acyl-ACP thioesterase, HD – hydroxyacyl-
456 ACP dehydratase, KAR – ketoacyl-ACP reductase, KAS – ketoacyl-ACP synthase, TE – fatty acid
457 thioesterase, TRX – thioredoxin-regulated enzyme; glycolipid synthesis: AAS – acyl-ACP synthase, ACPS
458 – holo-ACP synthase, AGP-AT – acylglycerophosphate acyltransferase, G3P-AT – glycerol-3-phosphate
459 acyltransferase, G3PDH – glycerol-3-phosphate dehydrogenase, MGDG/DGDG – mono-/digalactosyl
460 diacylglycerol, MGDGS/DGDGS – MGDG/DGDG synthase, PAP – phosphatidic acid phosphatase, SQD1
461 – UDP-sulfoquinovose synthase, SQD2/SQDX – sulfoquinovosyl diacylglycerol (SQDG) synthase,
462 UGE/PHD1 – UDP-glucose epimerase, UGP3 – UDP-glucose pyrophosphorylase 3; phospholipid
463 synthesis: CDS – CDP-diacylglycerol synthase, PGP1 – phosphatidylglycerophosphate synthase, PGP-P –
464 phosphateidylglycerophosphate phosphatase. **B:** Semiquantitative comparison of glycolipids present in *E.*
465 *longa* and autotrophic *E. gracilis*. Note the logarithmic scale of the quantification units (peak area). Peak
466 area is an arbitrary unit expressing the intensity of the signal of a particular lipid species, recalculated
467 according to their respective ionization promptitude. As each lipid species have different ionization
468 promptitude, note that direct comparison can be done only within lipid class (for details, see Tomčala et al.
469 2017). **C-E:** Immunofluorescence micrographs using anti-DGDG antibody (C), DAPI (D) and differential
470 interference contrast (E). Autotrophic *E. gracilis* represents a positive control, while the aplastidic
471 euglenozoan *R. costata* was used as negative control.
472

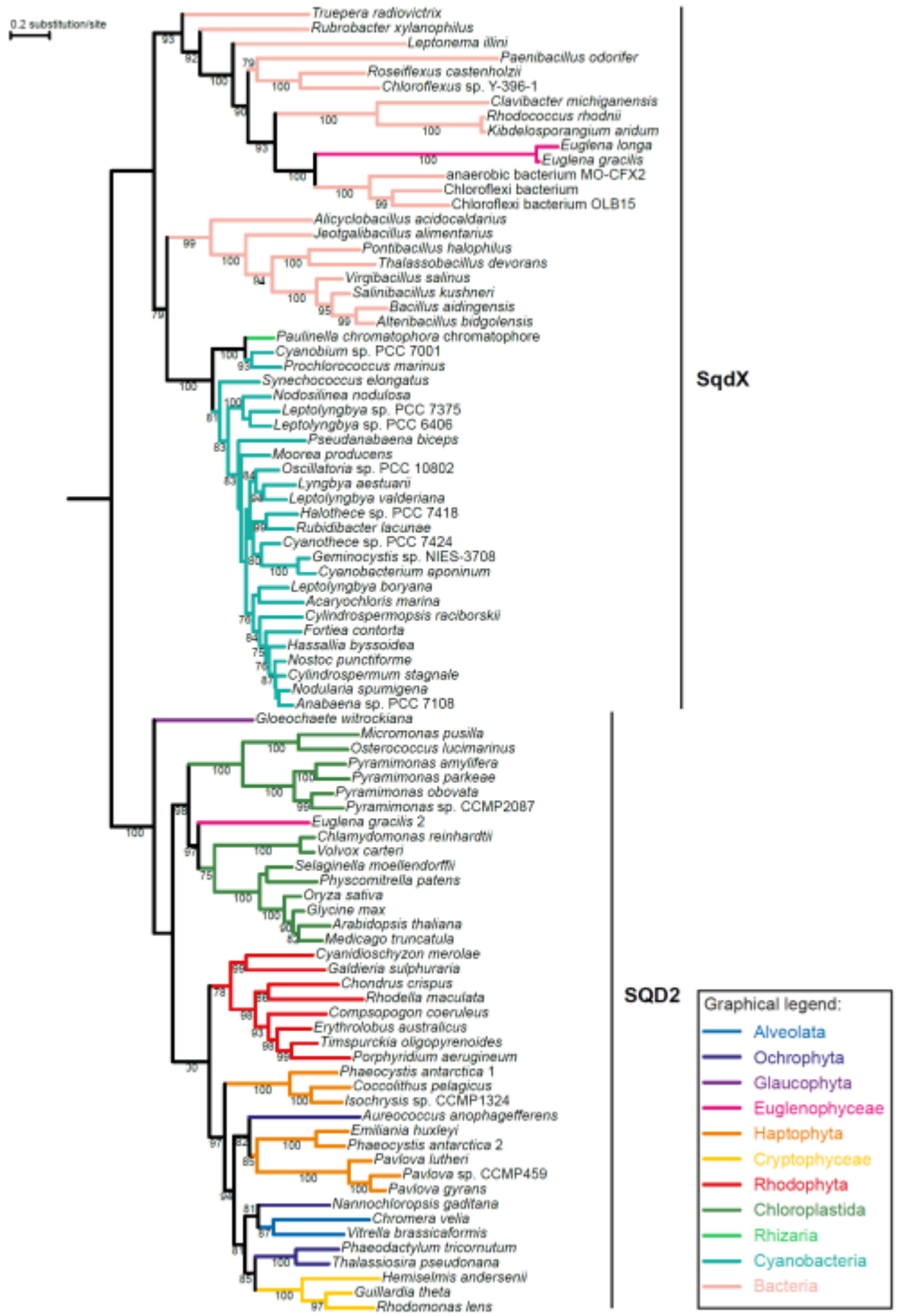


Figure 4: Euglenophytes have replaced the eukaryotic form of sulfoquinovosyltransferase (SQD2) with a bacterial version (SqdX). The maximum-likelihood tree was inferred with IQ-TREE using the LG+F+G4 substitution model and ultra-fast bootstrapping. The UFboot support values are indicated at branches when higher than 75%.

479 **A linearized Calvin-Benson pathway in the *E. longa* plastid**

480 The presence of genes encoding both subunits of the enzyme RuBisCO in *E. longa* (Záhonová et al. 2016)
481 raises the question whether the Calvin-Benson cycle (CBC) as a whole has been preserved in this organism
482 to provide ribulose-1,5-bisphosphate, the RuBisCO substrate. A putative *E. longa* plastid triose-phosphate
483 isomerase was described previously (Sun et al. 2008). We confirmed not only this candidate, but additionally
484 identified homologs with putative plastid-targeting presequences for almost all remaining CBC enzymes
485 (Table S8). Our identification was confirmed by phylogenetic analyses (supplementary Dataset S1), which
486 showed specific relationships of the *E. longa* proteins to the previously characterized CBC enzymes from
487 other euglenophytes (Markunas and Triemer 2016). However, two key CBC enzymes are apparently
488 missing from the *E. longa* transcriptome: phosphoglycerate kinase (ptPGK) and glyceraldehyde-phosphate
489 dehydrogenase (ptGAPDH). The homologs of these enzymes present in *E. longa* are not orthologous to the
490 plastid-targeted isoenzymes from other euglenophytes and all clearly lack a plastid-targeting presequence
491 (Table S8). Hence, these are presumably cytosolic enzymes involved in glycolysis/gluconeogenesis. The
492 lack of ptPGK and ptGAPDH in *E. longa* means that the product of the RuBisCO carboxylase activity, 3-
493 phosphoglycerate (3PG), cannot be converted (via 2,3-bisphosphoglycerate; 2,3-BPG) to glyceraldehyde-
494 3-phosphate (GA3P) in the plastid and the cycle becomes a linear pathway (Fig. 5).

495 Assuming that the reactions catalyzed by fructose bisphosphatase, phosphoribulokinase, and
496 RuBisCO are not reversible (Raines and Lloyd 2001), the flux through this linearized CB pathway most
497 likely goes from GA3P to 3PG, with a net production of six molecules of 3PG from five molecules of GA3P
498 due to fixation of three CO_2 molecules catalysed by RuBisCO. We thus need to define the origin of GA3P
499 entering the pathway. Euglenophytes do not store starch in the plastid (instead they have cytosolic
500 paramylon as a storage polysaccharide; Kiss et al. 1987), and indeed we did not find any glucose
501 metabolism-related enzymes in the predicted *E. longa* plastid proteome. Hence, GA3P cannot be produced
502 by a glycolytic route in the *E. longa* plastid. The presence of the plastid-targeted glycerophosphate
503 dehydrogenase (Table S5) in principle allows for generation of GA3P from glycerol-3-phosphate (via
504 dihydroxyacetone phosphate; DHAP; Fig. 3), which could possibly come from the degradation of
505 glycerolipids in the plastid. However, no phospholipid-degradation enzymes (phospholipases) appear to
506 localize to the plastid in *E. longa*. Hence, the primary function of glycerophosphate dehydrogenase is
507 perhaps to operate in the reverse direction, i.e. to provide glycerol-3-phosphate for the plastid phospholipid
508 and glycolipid synthesis (see above). The *E. longa* plastid thus most likely imports GA3P or DHAP from
509 the cytosol (Fig. 5). This assumption is supported by the presence of several members of the plastid
510 phosphate translocator (pPT) family (Fig. S7; Facchinelli and Weber 2011), including one phylogenetically
511 closest to a previously characterized cryptophyte transporter with a preference for DHAP (Haferkamp et al.
512 2006). Concerning the opposite end of the hypothesized linear CB pathway, we did not identify any

513 candidate *E. longa* plastid-targeted enzyme that would metabolize 3PG further (see the absence of 3PG
514 dehydrogenase discussed above), so this intermediate is most likely exported from the plastid into the
515 cytosol, probably also by one of the members of the pPT family of transporters (Fig. 5).
516 The operation of CBC is inherently linked with the oxygenase side-activity of RuBisCO, which
517 converts ribulose-1,5-bisphosphate into 3PG and phosphoglycolate instead of two molecules of 3PG that
518 are produced by the regular carboxylase activity (Tabita et al. 2007). Phosphoglycolate is metabolized in
519 the photorespiration pathway, initiated by phosphoglycolate phosphatase yielding glycolate. We did find a
520 candidate plastid-targeted phosphoglycolate phosphatase in *E. longa* (Table S8), orthologous to a protein
521 detected in *E. gracilis* plastid proteome, suggesting that the *E. longa* RuBisCO has the oxygenase activity
522 and phosphoglycolate is metabolized in this species. However, we did not find in our *E. longa* transcriptome
523 assembly any discernible homolog of the recently characterized transporter PLGG1 mediating glycolate
524 export from the plastid in the canonical photorespiratory pathway (Pick et al. 2013), although *E. gracilis*
525 does have it (Table S8). Since there is no obvious candidate for a plastid-targeted glycolate-metabolizing
526 enzyme (glycolate oxidase, glyoxylate reductase, glycolaldehyde dehydrogenase) in *E. longa*, it is unclear
527 how glycolate is removed from the plastid of this species. It is possible that the amount of glycolate produced
528 in the *E. longa* plastid is low and that it can be exported by an alternative (PLGG1-independent) route,
529 whose existence has been proposed also for plant plastids (Walker et al. 2016) and which might be sufficient
530 for glycolate recycling in the semi-parasitic plant *Cuscuta campestris* capable of low-efficiency
531 photosynthesis (Vogel et al. 2018).
532

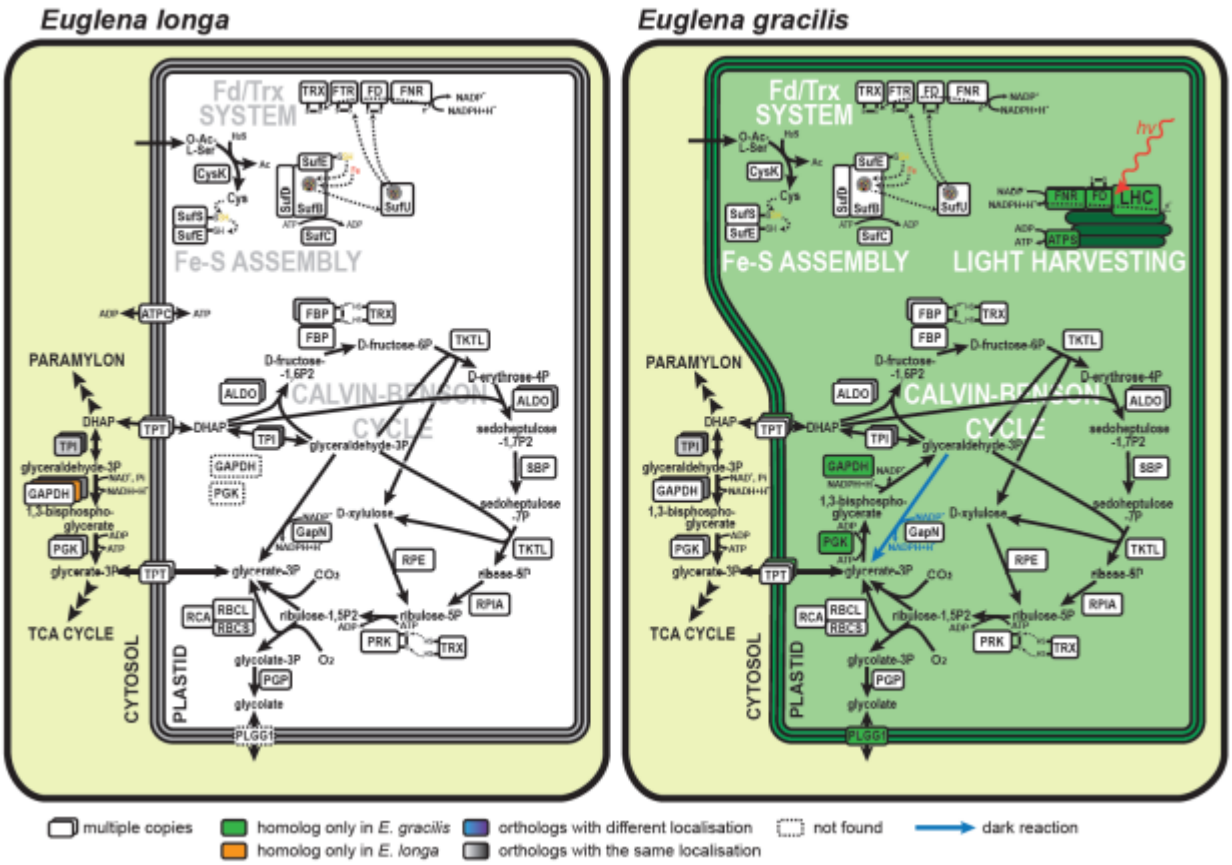


Figure 5: Carbon metabolism in the plastids of *E. longa* and *E. gracilis*. The Calvin-Benson cycle (CBC) resident to this organelle is central to the plastid carbon metabolism, regulated by the ferredoxin/thioredoxin (Fd/Trx) system. Reduction of disulfide bonds by the Fd/Trx system activates FBP and PRK. FTR and FD of the Fd/Trx system require for their function a post-translationally added Fe-S prosthetic group provided by the Fe-S assembly system. GapN apparently mediates shuttling of reducing equivalent (NADPH) through the exchange of DHAP/GA3P and 3PG, reflecting the cytosolic NADPH/NADP$^+$ ratio and thus an overall metabolic state of the cell. In contrast, *E. gracilis* plastid is an energy-converting organelle, harvesting light into chemical energy bound as NADPH and ATP and subsequently using this bound energy to fix CO_2 into organic carbohydrates via the CBC. Enzyme abbreviations are colour-coded according to their inferred evolutionary origin, see the graphical legend. Abbreviations, CBC: ALDO – aldolase, DHAP – dihydroxyacetone-phosphate, FBP – fructose-1,6-bisphosphatase, GAPDH – glyceraldehyde-3-phosphate dehydrogenase, PGK – 3-phosphoglygerate kinase, PGP – phosphoglycolate phosphatase, PLGG1 – plastid glycolate/glycerate transporter, PRK – phosphoribulokinase, RBCL/RBCS – RuBisCO large/small subunit, RCA – RuBisCO activase, RPE – ribulose-5-phosphate epimerase; RPIA – ribulose-phosphate isomerase A, SBP – sedoheptulose-1,7-bisphosphatase, TKTL – transketolase, TPI – triose-phosphate isomerase, TPT – triose-phosphate translocator; Fd/Trx system: FD – ferredoxin; FNR – FD/NADP+ oxidoreductase, FTR

550 – FD/TRX oxidoreductase, TRX – thioredoxin, ATPS – ATP synthase, ATPC – ADP/ATP translocase,
551 LHC – light-harvesting complex.

552

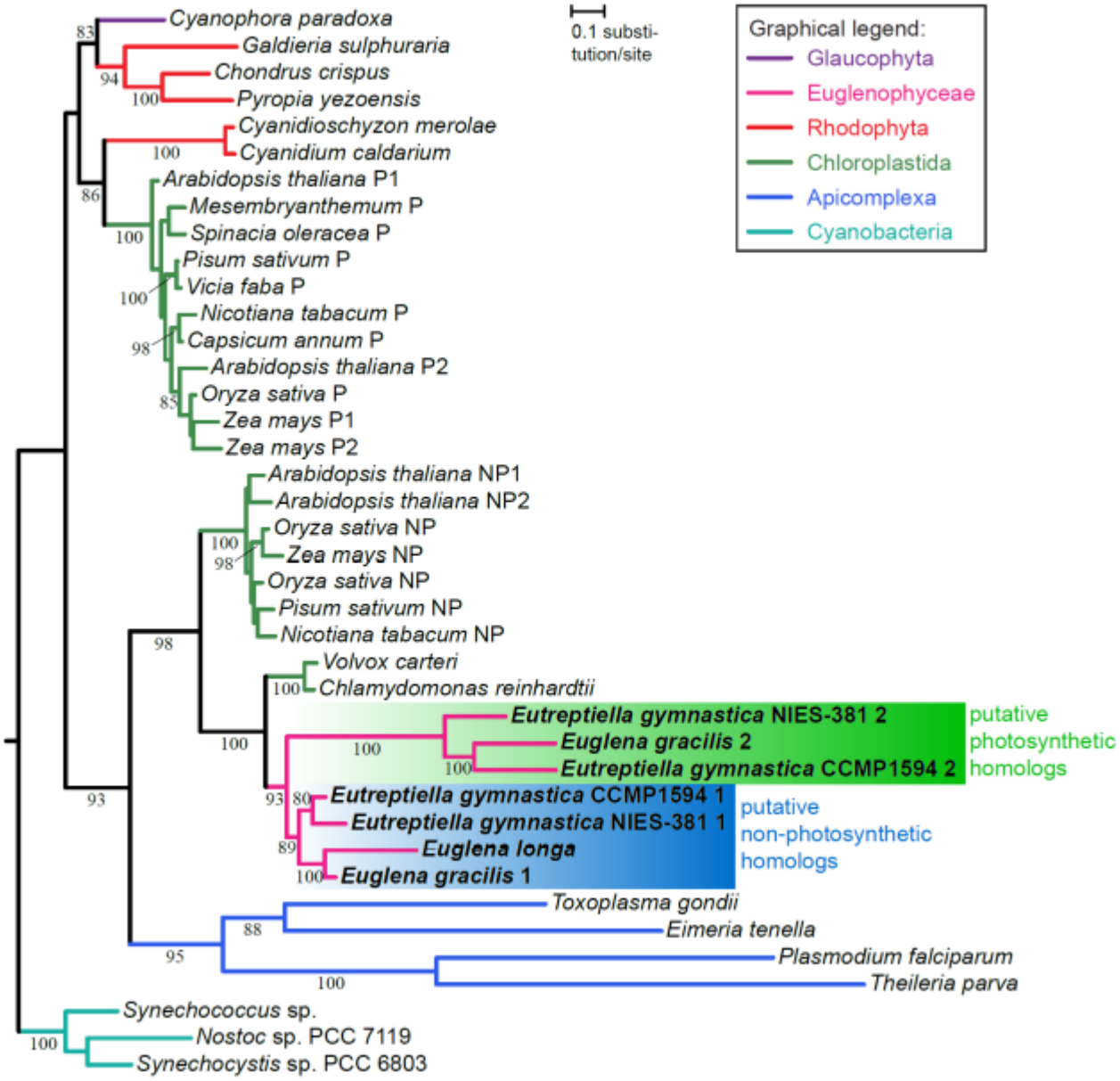
553 **The system of redox regulation of the Calvin-Benson pathway is conserved in *E. longa***

554 Although the photosynthetic machinery, including all photosystem I subunits, is missing from *E. longa*
555 (Záhonová et al. 2018), we found homologs (with clear plastidial localization) of the typical
556 "photosynthetic" (PetF-related) ferredoxin (Fd) and ferredoxin-$NADP^+$ reductase (FNR) (Table S9). These
557 two proteins are primarily involved in passing electrons from an activated photosystem I to $NADP^+$, which
558 is thus reduced to NADPH, the main donor of reducing equivalents for anabolic reactions in the plastid
559 (above all, CO_2 fixation). Phylogenetic analysis of FNR homologs from euglenophytes revealed the
560 existence of two different, yet related, clades affiliated to FNR from green algae (Fig. 6). One clade
561 comprises the *E. longa* FNR plus its orthologs from photosynthetic euglenophytes (*E. gracilis* and two
562 *Eutreptiella* strains), whereas the second clade is restricted to the photosynthetic species. Two different FNR
563 forms also exist in plants, one functioning in photosynthesis (production of NADPH dependent on the
564 function of the photosystem I; marked with "P" in Fig. 6) and the other being a "non-photosynthetic"
565 homolog ("NP" in Fig. 6) that allows the electron flow in the reverse direction, from NADPH to Fd (Vollmer
566 et al. 2001). In analogy with plants, we suggest that the two euglenophyte FNR clades functionally differ,
567 with one (that lacking a representative in *E. longa*) serving in photosynthesis and the other (present in *E.*
568 *longa*) mediating light-independent production of reduced Fd.

569 Multiple plastid enzymes depend on reduced Fd as an electron donor, namely glutamate synthase,
570 lipid desaturases, nitrite reductase, and sulfite reductase (Neuhaus and Emes 2000). As discussed in previous
571 sections, glutamate synthase and nitrite reductase are missing from *E. longa*, whereas all identified lipid
572 desaturases are predicted as targeted to the mitochondrion or the ER (Table S5). A two-subunit sulfite
573 reductase seems to be present in the plastid of *E. gracilis* (Novák Vanclová et al. 2019) and both subunits
574 have highly similar homologs in *E. longa* (Table S4), but this form of the enzyme utilizes NADPH rather
575 than ferredoxin as electron donor (Patron et al. 2008).

576 Another crucial role of Fd in plastids is to provide electrons to ferredoxin:thioredoxin reductase
577 (FTR) mediating reduction of the protein thioredoxin (Trx). The Fd/Trx system regulates several plastid
578 CBC enzymes in response to the redox status in the stroma and, in extension, to the photosynthetic activity
579 of the plastid (Fig. 5). An excess of NADPH leads to electrons being relayed from reduced ferredoxin to the
580 Fd/Trx system, which eventually reduces certain disulfide bonds in the target enzymes, thus changing their
581 activity (Schürmann and Buchanan 2008). This ensures activation of the CBC only when the photosynthetic
582 machinery works properly. Notably, FTR and Trx homologs with evident plastid-targeting presequences are
583 both present in *E. longa* (Table S9). Specific motifs necessary for the function of the Fd/Trx system

584 (Schürmann and Buchanan 2008) are conserved in the respective *E. longa* proteins (Fig. S8), consistent with
585 the Fd/Trx system being functional in this species. In addition, two *E. longa* CB pathway enzymes, fructose
586 bisphosphatase (two of the three isoforms present) and phosphoribulokinase, exhibit the conserved Trx
587 regulatory cysteine motifs similar to their orthologs in *E. gracilis* (Fig. S8, Table S10). Thus, we suggest
588 that the CB pathway in *E. longa* is sensitive to the redox status in the plastid, specifically to the concentration
589 of NADPH (Fig. 5).
590



591

592 **Figure 6: The inferred phylogeny of FNR.** The maximum-likelihood tree was inferred with IQ-TREE
593 using the LG+F+G4 substitution model and ultra-fast bootstrapping. The UFboot support values are

594 indicated at branches when higher than 75%. Euglenophyte species are in bold, and their putative
595 photosynthetic and non-photosynthetic homologs are depicted. P, photosynthetic; NP, non-photosynthetic.
596

597 **May the *E. longa* plastid be involved in keeping the redox balance of the cell?**

598 How to interpret in functional terms the retention of the linearized CB pathway and its putative redox
599 regulation in the non-photosynthetic plastid of *E. longa*? The key role of the pathway is supported by the
600 fact that production of the large RuBisCO subunit seems to be the *raison d'être* for the preservation of the
601 plastid genome in *E. longa* (Záhonová et al. 2016). The presence of CB enzymes in a non-photosynthetic
602 plastid is not without precedent, as it has been reported from a set of unrelated colourless algae and plants.
603 Some of them, e.g. the dinoflagellate *Crypthecodinium cohnii*, the dictyochophytes *Pteridomonas danica*
604 and *Ciliophrys infusionum*, the cryptophyte *Cryptomonas paramecium*, and some parasitic or
605 mycoheterotrophic land plants, are known to encode RuBisCO (Sekiguchi et al. 2002; Sanchez-Puerta et al.
606 2007; Donaher et al. 2009; Wicke et al. 2013; Hadariová et al. 2018), but how complete is the complement
607 of other CBC enzymes in these species is unknown. In contrast, transcriptomic or genomic analyses of other
608 colourless plastid-bearing taxa, such as the dinoflagellate *Pfiesteria piscicida*, the green alga
609 *Helicosporidium* sp. ATCC50920, the diatom *Nitzschia* sp. NIES-3581, and the non-photosynthetic
610 chrysophytes, revealed the presence of a subset of CB enzymes, including ptPGK and ptGAPDH, but not
611 of RuBisCO (Kim et al. 2013; Pombert et al. 2014; Kamikawa et al. 2017; Graupner et al. 2018). Hence,
612 the constellation of the CB enzymes retained in the *E. longa* plastid seems to be unique.

613 The CBC enzymes retained in various non-photosynthetic eukaryotes obviously do not serve to
614 sustain autotrophic growth, as ATP and the reducing power generated by photosynthesis are unavailable.
615 The CB pathway in *Nitzschia* sp. NIES-3581 was proposed to serve as a source of erythrose-4-P for the
616 synthesis of aromatic amino acid via the shikimate pathway in the plastid (Kamikawa et al. 2017). Although
617 not discussed in the respective report (Pombert et al. 2014), the CB pathway in the *Helicosporidium* plastid
618 may likewise serve to feed the co-localized shikimate pathway with erythrose-4-P. However, such a
619 rationalization of the CB pathway in the *E. longa* plastid would not work, since enzymes for aromatic amino
620 acid biosynthesis in this species are apparently localized to the cytosol (Table S2) and thus have access to
621 erythrose-4-P produced by the pentose phosphate pathway. In addition, the need to produce erythrose-4-P
622 in the *E. longa* plastid would not explain the retention of RuBisCO (absent in both *Nitzschia* and
623 *Helicosporidium*). A photosynthesis- and CBC-independent role of RuBisCO was described in oil formation
624 in developing seeds of *Brassica napus*, where refixation of CO_2 released during carbohydrate-to-fatty acid
625 conversion increases carbon use efficiency (Schwender et al. 2004). A similar explanation is unlikely to
626 hold for RuBisCO retention in *E. longa*, given the lack of fatty acid synthesis in its plastid and the apparently
627 much smaller significance of oil as a reserve substance in *E. longa* compared to *B. napus*.

628 We believe that the identification of the putatively functional Fd/Trx system, despite the absence of
629 the photosynthetic electron transport chain in this species, provides one of the key hints to understanding
630 the physiological role of the linear CB pathway in the *E. longa* plastid. The second potentially important
631 piece of the puzzle is provided by the proteomic data from *E. gracilis* (Novák Vanclová et al. 2019), which
632 indicated the presence of a unique form of GAPDH, the so-called non-phosphorylating GAPDH also
633 referred to as GapN, in the plastid (Table S8). This enzyme uses $NADP^+$ to directly oxidize GA3P to 3PG,
634 skipping the intermediate 2,3-BPG (and hence not leading to ATP generation) and producing NADPH rather
635 than NADH (Iddar et al. 2003). In plants, this enzyme is localized to the cytosol and is involved in shuttling
636 of reducing equivalents from the plastid by the exchange of GA3P and 3PG between the two compartments
637 (Rius et al. 2006). *E. longa* possesses a GapN homolog highly similar to the *E. gracilis* protein, including
638 an N-terminal presequence (Table S8), consistent with its presumed plastidial localization. It thus appears
639 that in *Euglena* spp. GapN mediates shuttling of reducing equivalents in the opposite direction than in plants,
640 i.e. from the cytosol to the plastid (Fig. 5). In case of *E. longa* this may be the main (if not the only)
641 mechanism of providing NADPH for the use in the plastid, whereas *E. gracilis* would utilize it when
642 photosynthetic production of NADPH is shut down. At the same time, the shuttle provides a mechanism of
643 linking the level of NADPH in the plastid with the cytosolic concentration of GA3P.

644 Taken together, we propose that in *E. longa* (at specific circumstances possibly also in *E. gracilis*)
645 the plastidial $NADPH/NADP^+$ ratio is directly influenced by the redox status of the cell, i.e. that it rises in
646 an excess of reducing power that slows down the glycolytic oxidation of GA3P in the cytosol. This
647 stimulates the linear CB pathway via the Fd/Trx system, effectively decreasing the level of GA3 by
648 converting it to 3PG without further increasing the reducing power in the cell. This conclusion is apparent
649 from considering the overall stoichiometries of the two alternative pathways from GA3 to 3PG:

650 glycolysis: $5\ GA3P + 5\ NAD^+ + 5\ ADP + 5\ PO_3^{2-} \rightarrow 5\ 3PG + 5\ NADH + 5\ H^+ + 5\ ATP$

651 CB pathway: $5\ GA3P + 3\ CO_2 + 3\ H_2O + 3\ ATP \rightarrow 6\ 3PG + 3\ ADP + 2\ PO_3^{2-}$

652 The key difference is that the CB pathway does not produce NADH that needs to be reoxidized to keep the
653 glycolytic pathway running, since the fixed CO_2 effectively serves as an electron acceptor. Hence, turning
654 the CB bypass on may help the cell keep the redox balance when reoxidation of NADH is not efficient, e.g.
655 at hypoxic (or anoxic) conditions (although this happens at the expense of ATP). Indeed, euglenophytes in
656 their natural settings are probably often exposed to the shortage of oxygen, and anaerobiosis in *E. gracilis*
657 has been studied to some extent (Tucci et al. 2010; Zimorski et al. 2017). The anaerobic heterotrophic
658 metabolism of *E. gracilis* relies on fermentative degradation of the reserve polysaccharide paramylon
659 leading to production of wax esters (Yoshida et al. 2016). It is likely that *E. longa* exhibits a similar
660 metabolic adaptation to low levels of ambient oxygen as *E. gracilis*. However, details of the euglenophyte
661 anaerobic metabolism need to be worked out yet, and we propose that the plastid may be involved in it as a

662 kind of a redox valve. Work is ongoing to test this hypothesis and to illuminate further details of
663 physiological role of the linear CB pathway in the *E. longa* plastid.
664
665
666 **Conclusions**
667 Endosymbiotic organelles have proven to be extremely evolutionarily versatile. One of the manifestations
668 is the recurrent loss of key metabolic functions of the canonical forms of both mitochondria and plastids,
669 resulting in anaerobic mitochondria-related organelles (such as hydrogenosomes and mitosomes) and non-
670 photosynthetic plastids distributed across diverse eukaryotic branches. A lot of attention has been paid to
671 various mitochondrial derivatives and it is now well documented that they vary substantially in the
672 complement of functional pathways they have retained (Roger et al. 2017). The variation in functional
673 profiles of non-photosynthetic plastids is less well known, as the only example studied in detail is the
674 apicoplast of apicomplexan parasites. Nevertheless, a picture is emerging that independently evolved
675 colourless plastids may also exhibit a surprising degree of diversity in terms of their metabolic capacity.
676 Our analyses of the *E. longa* plastid stretch the breadth of variation among non-photosynthetic
677 plastids even further. The combination of pathways present (tocopherol and phylloquinone synthesis,
678 glycolipid synthesis and a linearized CB pathway including RuBisCO), absent (fatty acid, amino acid, and
679 isoprenoid precursor synthesis), and potentially residual (tetrapyrrole synthesis) makes the *E. longa* plastid
680 unlike any of the previously investigated non-photosynthetic plastids, including the apicoplast. However,
681 further work, combining additional *in silico* analyses (aimed, e.g., at potential plastid membrane transporters
682 mediating metabolite exchange with the cytosol) with biochemical and cytological investigations is needed
683 to achieve a more precise idea about the protein composition of the *E. longa* plastid and a better
684 understanding of its physiological roles.
685
686
687 **Materials and Methods**
688 **Identification and annotation of plastid-targeted proteins**
689 The analyses reported in this study were done using the *E. longa* transcriptome assembly reported previously
690 (Záhonová et al. 2018). Protein models for annotation were generated by a custom Geneious 8.1.6 (Kearse
691 et al. 2012) script that extracted all open reading frames longer than 297 bp, translated the sequences and
692 then filtered the protein models by a local BLAST+ ver.2.2.30 (Altschul et al. 1997) search against the
693 Swiss-Prot database (version 10/5/15, max E-value=10). Transcript models containing at least a partial
694 spliced leader sequence (TTTTTCG) at their 5'-end or 3'-end (within the first or last 35 nt) were translated
695 in the forward or reverse direction only, respectively.

696 Candidates for plastid-targeted proteins were identified using criteria for prediction of plastid-
697 targeting presequences described in detail by Záhonová et al. (2018). In the first step protein sequences were
698 gathered that fulfilled at least one of the following requirements: (i) the signal peptide was predicted by the
699 PrediSi v.2004 (Hiller et al. 2004) or PredSL v.2005 (Petsalaki et al. 2006) standalone programs; (ii) one or
700 two transmembrane domains at the N-terminus of the protein were predicted by standalone TMHMM 2.0c
701 (online version where graphical output was considered) (Krogh et al. 2001). The resulting set was then
702 filtered by checking for the presence of a plastid transit peptide, which was predicted by standalone
703 MultiLoc2.5 (Blum et al. 2009) after an *in silico* removal of the signal peptide or the first transmembrane
704 domain. Finally, protein models with a putative plastid-targeting presequence were automatically annotated
705 using InterProScan 5.21 (Jones et al. 2014) and the annotations were manually scrutinized to identify
706 proteins with an assignable specific metabolism-related function (enzymes of the metabolism of nucleic
707 acids and proteins were ignored in this study).
708 In parallel we searched with BLAST v.2.2.30 (including tBLASTn against the transcriptome
709 assembly and, in special cases, even against the whole set of raw RNAseq reads) for putative plastid proteins
710 by direct identification and evaluation of homologs of enzymes of specific biochemical pathways potentially
711 localized to the plastid; as queries we used respective protein sequences from *E. gracilis* (as identified by
712 Novák Vanclová et al. 2019), reference sequences from the KEGG PATHWAY Database
713 (https://www.genome.jp/kegg/pathway.html), or sequences identified by literature searches. In some cases,
714 a more sensitive homology detection algorithm HMMER 3.0 (Mistry et al. 2013) to identify homologs of
715 poorly conserved enzymes (e.g., UROS). For comparative purposes we used the same approach to identify
716 plastid-targeted proteins encoded by the transcriptome assemblies from *E. gracilis* reported by (Yoshida et
717 al. 2016) (accession GDJR00000000.1) and (Ebenezer et al. 2017) (accession GEFR00000000.1). Where
718 accessions are given for *E. gracilis* sequences, those starting with GDJR and GEFR belong to the former
719 and latter dataset, respectively.
720 For MenA cDNA resequencing, RNA was isolated using TRI Reagent (Thermo Fisher Scientific,
721 San Jose, USA) and mRNA was then extracted using the Dynabeads mRNA Purification kit (Thermo Fisher
722 Scientific). Reverse-transcription was performed with random hexamers and StrataScript III Reverse
723 Transcriptase (Thermo Fisher Scientific). For cDNA amplification, forward 5'-
724 GGTGCTGTTCTGCTCTCACT-3' and reverse 5'-CAGTGGGGATCAGAGATGCG-3' primers, and the
725 Q5 High-Fidelity DNA polymerase in a standard buffer (New England Biolabs) were used. Amplicons were
726 purified on MinElute PCR Purification columns (Qiagen, Hilden, Germany) and sequenced at the GATC
727 sequencing facility (Konstanz, Germany). The MenA cDNA sequence is deposited in GenBank with the
728 accession number MK484704.
729

730 **Phylogenetic analyses**

731 Phylogenetic analyses were employed to establish orthologous relationships among *E. longa* and *E. gracilis*
732 genes or to illuminate the origin of the euglenophyte proteins of special interest. Homologs of target proteins
733 were identified by BLAST v.2.2.30 searches in the non-redundant protein sequence database at NCBI
734 (www.ncbi.nlm.nih.gov) and among protein models of selected organisms from JGI (Joint Genome
735 Institute, jgi.doe.gov) and MMETSP (Marine Microbial Eukaryote Transcriptome Sequencing Project,
736 marinemicroeukaryotes.org; Keeling et al. 2014). Sequences were aligned using the MAFFT v7.407
737 (Multiple Alignment using Fast Fourier Transform) tool with L-INS-I setting (Katoh and Standley 2013)
738 and poorly aligned positions were eliminated by the trimAL v1.4.rev22 tool with "-automated1" trimming
739 (Capella-Gutierrez et al. 2009). For presentation purposes, alignments were processed using the CHROMA
740 software (Goodstadt and Ponting 2001). Maximum likelihood (ML) trees were inferred from the alignments
741 using the LG+F+G4 model of IQ-TREE v1.6.9 (Nguyen et al. 2015), and employing the strategy of rapid
742 bootstrapping followed by a "thorough" ML search with 1,000 bootstrap replicates (-m TEST -bb 1000).
743 The list of species, and the number of sequences and amino acid positions are present in Tables S11-S22 for
744 each phylogenetic tree.

745

746 **Culture conditions**

747 *Euglena gracilis* strain *Z* ("autotrophic" conditions) were cultivated statically under constant illumination
748 at 26 °C in Cramer-Myers medium supplemented with ethanol (0.8% v/v) as a carbon source (Cramer and
749 Myers 1952). *E. longa* strain CCAP 1204-17a (a gift from W. Hachtel, Bonn, Germany) and heterotrophic
750 *E. gracilis* strain *Z* were cultivated in identical medium without illumination. The cultures of *E. longa* were
751 not completely axenic, but the contaminating bacteria were kept at as low level as possible. *Rhabdomonas*
752 *costata* strain PANT2 was provided by Vladimír Hampl (Department of Parasitology, Faculty of Science,
753 Charles University in Prague, Czech Republic). It was isolated from a freshwater body in Pantanal (Brasil)
754 and grown with an uncharacterised mixture of bacteria in Sonneborn's *Paramecium* medium (pH 7.4;
755 Sonneborn 1950) at room temperature (inoculated every three to four weeks).

756

757 **Mass spectrometry of structural lipids and terpenoids**

758 For analysis of structural lipids, extracts from *E. longa* and autotrophic *E. gracilis* cellular pellets (four
759 biological samples of different culture age) were obtained with chloroform and methanol solution (ratio –
760 2:1) following the method of (Folch et al. 1957) as modified by (Košťál and Šimek 1998). Samples were
761 homogenized in extraction solution with glass beads using TissueLyser LT mill (Qiagen). Homogenates
762 were dried, weighted, and resolved in 500 µl of chloroform and methanol (1:2) with internal standard PC
763 17:0/17:0 (Sigma Aldrich). Aliquots from each sample extract were used for lipid determination by HPLC

764 using a liquid chromatograph and autosampler Accela (Thermo Fisher Scientific). The samples (5 µL) were
765 injected and separated on the Gemini column 250 × 2 mm; i.d. 3 µm (Phenomenex, Torrance, USA). A
766 linear ion trap LTQ-XL mass spectrometer (Thermo Fisher Scientific) was used in both positive and
767 negative ion ESI mode. The settings of the system followed the methodology published earlier (Tomčala et
768 al. 2017). Data were acquired and processed using Xcalibur software version 2.1 (Thermo Fisher Scientific).
769 Particular compounds were determined based on m/z value, retention time, behaviour in positive and
770 negative ionization mode, and characteristic fragmentation pattern of target molecules (for details see
771 Tomčala et al. 2017).
772 Terpenoids were extracted from an autotrophic culture of *E. gracilis*, a heterotrophic culture of *E.*
773 *gracilis*, and a culture of *E. longa* of the same age in three repetitions. The same extraction procedure as for
774 lipid analysis was used. Sample aliquots were injected into the high-resolution mass spectrometry system
775 powered by Orbitrap Q-Exactive Plus with Dionex Ultimate 3000 XRS pump and Dionex Ultimate 3000
776 XRS Open autosampler (all by Thermo Fisher Scientific) and followed the settings described in (Tomčala
777 et al. 2017). Data were acquired and processed using Xcalibur software version 2.1. Identification of OH-
778 PhQ was achieved by considering the m/z value, fragmentation pattern, and high-resolution data.
779 Tocopherols (α, β/γ, and δ) were determined by the same characteristic as OH-PhQ and results were then
780 compared with commercially purchased standards (α-tocopherol from Sigma Aldrich, the other three
781 variants from SUPELCO).
782
783 **Immunofluorescence assay**
784 Immunofluorescence was performed as described previously (Botté et al. 2011). Briefly, cells were fixed in
785 4% paraformaldehyde for 30 minutes, cellular membranes were permeabilized for 10 minutes on ice with
786 0.1% non-ionic detergent Igepal CA-630 (Sigma-Aldrich) in PHEM buffer pH 6.9 (60 mM PIPES, 25 mM
787 HEPES, 10 mM EGTA, 2 mM $MgCl_2$), and background was masked with 3% BSA in PHEM buffer. DGDG
788 was detected using a polyclonal rabbit anti-DGDG antibody (1:25) that was a kind gift from Cyrille Y. Botté
789 (ApicoLipid Team, Laboratoire Adaptation et Pathogenie des Microorganismes, University of Grenoble I,
790 France), followed by incubation with a secondary Cy3-labeled polyclonal goat anti-rabbit antibody
791 (AP132C, 1:800, Merck Millipore, Billerica, USA). To improve the fluorescence lifetime, Fluoroshield™
792 with DAPI mounting medium (Sigma-Aldrich) was used. Cells were mounted on slides and observed with
793 a fluorescent microscope Olympus BX53 (Olympus, Tokyo, Japan). Photosynthetic *E. gracilis* served as a
794 positive control, and the primary osmotroph (i.e. aplastidic) *R. costata* as a negative control.
795
796

Acknowledgements

We thank Vladimír Hampl (Charles University) for providing the culture of *Rhabdomonas costata* and Cyrille Y. Botté (University of Grenoble I) for providing the anti-DGDG antibody. We acknowledge computation resources provided by CERIT-SC and MetaCentrum, Brno, Czech Republic, and the infrastructure grant "Přístroje IET" (CZ.1.05/2.1.00/19.0388). We thank Laboratory of Analytical Biochemistry and Metabolomics (Biology Centre, Czech Academy of Sciences) for free access to LC–MS instruments. This study was supported by the Czech Science Foundation grants 17-21409S (to ME) and 16-24027S (to MO), the Czech Academy of Sciences fellowship (to ZF), and the Scientific Grant Agency of the Slovak Ministry of Education (grant VEGA 1/0535/17 to JK). This work was also part of the project "Centre for research of pathogenicity and virulence of parasites", supported by the European Regional Development Fund, within the Operational programme for Research, Development and Education (CZ.02.1.01/0.0/0.0/16_019/0000759).

Competing interests

The authors declare that they have no financial or non-financial competing interests.

References

Altschul, SF, Madden, TL, Schaffer, AA, Zhang, J, Zhang, Z, Miller, W, Lipman, DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389-3402.

Anderson, R, Livermore, BP, Kates, M, Volcani, BE. 1978. The lipid composition of the non-photosynthetic diatom *Nitzschia alba*. Biochim Biophys Acta 528:77-88.

Awai, K, Maréchal, E, Block, MA, Brun, D, Masuda, T, Shimada, H, Takamiya, K, Ohta, H, Joyard, J. 2001. Two types of MGDG synthase genes, found widely in both 16:3 and 18:3 plants, differentially mediate galactolipid syntheses in photosynthetic and nonphotosynthetic tissues in *Arabidopsis thaliana*. Proc Natl Acad Sci U S A 98:10960-10965.

Blum, T, Briesemeister, S, Kohlbacher, O. 2009. MultiLoc2: integrating phylogeny and Gene Ontology terms improves subcellular protein localization prediction. BMC Bioinformatics 10:274.

Borza, T, Popescu, CE, Lee, RW. 2005. Multiple metabolic roles for the nonphotosynthetic plastid of the green alga *Prototheca wickerhamii*. Eukaryot Cell 4:253-261.

Botté, C, Saidani, N, Mondragon, R, Mondragon, M, Isaac, G, Mui, E, McLeod, R, Dubremetz, JF, Vial, H, Welti, R, et al. 2008. Subcellular localization and dynamics of a digalactolipid-like epitope in *Toxoplasma gondii*. J Lipid Res 49:746-762.

831 Botté, CY, Yamaryo-Botté, Y, Janouškovec, J, Rupasinghe, T, Keeling, PJ, Crellin, P, Coppel, RL,
832 Maréchal, E, McConville, MJ, McFadden, GI. 2011. Identification of plant-like galactolipids in
833 *Chromera velia*, a photosynthetic relative of malaria parasites. J Biol Chem 286:29893-29903.
834 Botté, CY, Yamaryo-Botté, Y, Rupasinghe, TW, Mullin, KA, MacRae, JI, Spurck, TP, Kalanon, M, Shears,
835 MJ, Coppel, RL, Crellin, PK, et al. 2013. Atypical lipid composition in the purified relict plastid
836 (apicoplast) of malaria parasites. Proc Natl Acad Sci U S A 110:7506-7511.
837 Brettel, K. 1997. Electron transfer and arrangement of the redox cofactors in photosystem I. Biochimica et
838 Biophysica Acta (BBA) - Bioenergetics 1318:322-373.
839 Capella-Gutierrez, S, Silla-Martinez, JM, Gabaldon, T. 2009. trimAl: a tool for automated alignment
840 trimming in large-scale phylogenetic analyses. Bioinformatics 25:1972-1973.
841 Cenci, U, Qiu, H, Pillonel, T, Cardol, P, Remacle, C, Colleoni, C, Kadouche, D, Chabi, M, Greub, G,
842 Bhattacharya, D, et al. 2018. Host-pathogen biotic interactions shaped vitamin K metabolism in
843 Archaeplastida. Sci Rep 8:15243.
844 Cihlář, J, Füssy, Z, Horák, A, Oborník, M. 2016. Evolution of the tetrapyrrole biosynthetic pathway in
845 secondary algae: conservation, redundancy and replacement. PLOS ONE 11:e0166338.
846 Cramer, M, Myers, J. 1952. Growth and photosynthetic characteristics of *Euglena gracilis*. Archiv
847 Mikrobiol 17:384-402.
848 Dagenais-Bellefeuille, S, Morse, D. 2013. Putting the N in dinoflagellates. Front Microbiol 4:369.
849 Dailey, HA, Dailey, TA, Gerdes, S, Jahn, D, Jahn, M, O'Brian, MR, Warren, MJ. 2017. Prokaryotic heme
850 biosynthesis: Multiple pathways to a common essential product. Microbiol Mol Biol Rev 81.
851 Disch, A, Schwender, J, Muller, C, Lichtenthaler, HK, Rohmer, M. 1998. Distribution of the mevalonate
852 and glyceraldehyde phosphate/pyruvate pathways for isoprenoid biosynthesis in unicellular algae
853 and the cyanobacterium *Synechocystis* PCC 6714. Biochem J 333 (Pt 2):381-388.
854 Donaher, N, Tanifuji, G, Onodera, NT, Malfatti, SA, Chain, PS, Hara, Y, Archibald, JM. 2009. The
855 complete plastid genome sequence of the secondarily nonphotosynthetic alga *Cryptomonas*
856 *paramecium*: reduction, compaction, and accelerated evolutionary rate. Genome Biol Evol 1:439-
857 448.
858 Dorrell, RG, Azuma, T, Nomura, M, Audren de Kerdrel, G, Paoli, L, Yang, S, Bowler, C, Ishii, KI,
859 Miyashita, H, Gile, GH, et al. 2019. Principles of plastid reductive evolution illuminated by
860 nonphotosynthetic chrysophytes. Proc Natl Acad Sci U S A 116:6914-6923.
861 Ebenezer, TE, Carrington, M, Lebert, M, Kelly, S, Field, MC. 2017. *Euglena gracilis* genome and
862 transcriptome: Organelles, nuclear genome assembly strategies and initial features. Adv Exp Med
863 Biol 979:125-140.
864 Fatihi, A, Latimer, S, Schmollinger, S, Block, A, Dussault, PH, Vermaas, WF, Merchant, SS, Basset, GJ.

865 2015. A dedicated type II NADPH dehydrogenase performs the penultimate step in the biosynthesis
866 of vitamin K_1 in *Synechocystis* and *Arabidopsis*. Plant Cell 27:1730-1741.
867 Fernandez, E, Galvan, A. 2008. Nitrate assimilation in *Chlamydomonas*. Eukaryot Cell 7:555-559.
868 Fernández Robledo, JA, Caler, E, Matsuzaki, M, Keeling, PJ, Shanmugam, D, Roos, DS, Vasta, GR. 2011.
869 The search for the missing link: a relic plastid in *Perkinsus*? Int J Parasitol 41:1217-1229.
870 Figueroa-Martinez, F, Nedelcu, AM, Smith, DR, Reyes-Prieto, A. 2015. When the lights go out: the
871 evolutionary fate of free-living colorless green algae. New Phytologist 206:972-982.
872 Folch, J, Lees, M, Sloane Stanley, GH. 1957. A simple method for the isolation and purification of total
873 lipides from animal tissues. J Biol Chem 226:497-509.
874 Furt, F, Oostende, Cv, Widhalm, JR, Dale, MA, Wertz, J, Basset, GJC. 2010. A bimodular oxidoreductase
875 mediates the specific reduction of phylloquinone (vitamin K1) in chloroplasts. Plant J 64:38-46.
876 Füssy, Z, Oborník, M. 2017. Chapter Six - Chromerids and Their Plastids. In: Y Hirakawa, editor. Advances
877 in Botanical Research: Academic Press. p. 187-218.
878 Füssy Z, Faitová T, Oborník M. 2019. Subcellular compartments interplay for carbon and nitrogen
879 allocation in *Chromera velia* and *Vitrella brassicaformis*. Genom Biol Evol, in press.
880 doi:10.1093/gbe/evz123.
881 Gerdes, S, Lerma-Ortiz, C, Frelin, O, Seaver, SM, Henry, CS, de Crecy-Lagard, V, Hanson, AD. 2012.
882 Plant B vitamin pathways and their compartmentation: a guide for the perplexed. J Exp Bot
883 63:5379-5395.
884 Giordano, M, Raven, JA. 2014. Nitrogen and sulfur assimilation in plants and algae. Aquatic Botany 118:45-
885 61.
886 Gockel, G, Hachtel, W. 2000. Complete gene map of the plastid genome of the nonphotosynthetic euglenoid
887 flagellate *Astasia longa*. Protist 151:347-351.
888 Gockel, G, Hachtel, W, Baier, S, Fliss, C, Henke, M. 1994. Genes for components of the chloroplast
889 translational apparatus are conserved in the reduced 73-kb plastid DNA of the nonphotosynthetic
890 euglenoid flagellate *Astasia longa*. Curr Genet 26:256-262.
891 Goddard-Borger, ED, Williams, SJ. 2017. Sulfoquinovose in the biosphere: occurrence, metabolism and
892 functions. Biochem J 474:827-849.
893 Gomez-Silva, B, Timko, MP, Schiff, JA. 1985. Chlorophyll biosynthesis from glutamate or 5-
894 aminolevulinate in intact *Euglena* chloroplasts. Planta 165:12-22.
895 Goodstadt, L, Ponting, CP. 2001. CHROMA: consensus-based colouring of multiple alignments for
896 publication. Bioinformatics 17:845-846.
897 Graupner, N, Jensen, M, Bock, C, Marks, S, Rahmann, S, Beisser, D, Boenigk, J. 2018. Evolution of
898 heterotrophy in chrysophytes as reflected by comparative transcriptomics. FEMS Microbiol Ecol

899 94.
900 Gu, X, Harding, S, Nyamdari, B, Aulakh, K, Clermont, K, Westwood, J, Tsai, C-J. 2018. A role for
901 phylloquinone biosynthesis in the plasma membrane as revealed in a non-photosynthetic parasitic
902 plant. bioRxiv:257519.
903 Gutbrod, K, Romer, J, Dormann, P. 2019. Phytol metabolism in plants. Prog Lipid Res 74:1-17.
904 Hadariová, L, Vesteg, M, Birčák, E, Schwartzbach, SD, Krajčovič, J. 2017. An intact plastid genome is
905 essential for the survival of colorless *Euglena longa* but not *Euglena gracilis*. Curr Genet 63:331-
906 341.
907 Hadariová, L, Vesteg, M, Hampl, V, Krajčovič, J. 2018. Reductive evolution of chloroplasts in non-
908 photosynthetic plants, algae and protists. Curr Genet 64:365-387.
909 Haferkamp, I, Deschamps, P, Ast, M, Jeblick, W, Maier, U, Ball, S, Neuhaus, HE. 2006. Molecular and
910 biochemical analysis of periplastidial starch metabolism in the cryptophyte *Guillardia theta*.
911 Eukaryot Cell 5:964-971.
912 Hachtel, W. 1996. DNA and gene expression in nonphotosynthetic plastids. In: M Pessarakli, editor.
913 Handbook of Photosynthesis. New York: Marcel Dekker. p. 349-355.
914 Hiller, K, Grote, A, Scheer, M, Munch, R, Jahn, D. 2004. PrediSi: prediction of signal peptides and their
915 cleavage positions. Nucleic Acids Res 32:W375-379.
916 Hori, K, Nobusawa, T, Watanabe, T, Madoka, Y, Suzuki, H, Shibata, D, Shimojima, M, Ohta, H. 2016.
917 Tangled evolutionary processes with commonality and diversity in plastidial glycolipid synthesis
918 in photosynthetic organisms. Biochim Biophys Acta 1861:1294-1308.
919 Iddar, A, Valverde, F, Serrano, A, Soukri, A. 2003. Purification of recombinant non-phosphorylating
920 NADP-dependent glyceraldehyde-3-phosphate dehydrogenase from *Streptococcus pyogenes*
921 expressed in *E. coli*. Mol Cell Biochem 247:195-203.
922 Jackson, C, Knoll, AH, Chan, CX, Verbruggen, H. 2018. Plastid phylogenomics with broad taxon sampling
923 further elucidates the distinct evolutionary origins and timing of secondary green plastids. Sci Rep
924 8:1523.
925 Janouškovec, J, Horák, A, Oborník, M, Lukeš, J, Keeling, PJ. 2010. A common red algal origin of the
926 apicomplexan, dinoflagellate, and heterokont plastids. Proc Natl Acad Sci U S A 107:10949-10954.
927 Janouškovec, J, Tikhonenkov, DV, Burki, F, Howe, AT, Kolísko, M, Mylnikov, AP, Keeling, PJ. 2015.
928 Factors mediating plastid dependency and the origins of parasitism in apicomplexans and their close
929 relatives. Proc Natl Acad Sci U S A 112:10200-10207.
930 Jones, P, Binns, D, Chang, HY, Fraser, M, Li, W, McAnulla, C, McWilliam, H, Maslen, J, Mitchell, A,
931 Nuka, G, et al. 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics
932 30:1236-1240.

933 Kamikawa, R, Moog, D, Zauner, S, Tanifuji, G, Ishida, KI, Miyashita, H, Mayama, S, Hashimoto, T, Maier,
934 UG, Archibald, JM, et al. 2017. A non-photosynthetic diatom reveals early steps of reductive
935 evolution in plastids. Mol Biol Evol 34:2355-2366.
936 Karamoko, M, Cline, S, Redding, K, Ruiz, N, Hamel, PP. 2011. Lumen thiol oxidoreductase1, a disulfide
937 bond-forming catalyst, is required for the assembly of photosystem II in *Arabidopsis*. The Plant
938 Cell 23:4462.
939 Katoh, K, Standley, DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in
940 performance and usability. Mol Biol Evol 30:772-780.
941 Ke, H, Sigala, PA, Miura, K, Morrisey, JM, Mather, MW, Crowley, JR, Henderson, JP, Goldberg, DE,
942 Long, CA, Vaidya, AB. 2014. The heme biosynthesis pathway is essential for *Plasmodium*
943 *falciparum* development in mosquito stage but not in blood stages. J Biol Chem 289:34827-34837.
944 Kearse, M, Moir, R, Wilson, A, Stones-Havas, S, Cheung, M, Sturrock, S, Buxton, S, Cooper, A,
945 Markowitz, S, Duran, C, et al. 2012. Geneious Basic: an integrated and extendable desktop software
946 platform for the organization and analysis of sequence data. Bioinformatics 28:1647-1649.
947 Keeling, PJ. 2013. The number, speed, and impact of plastid endosymbioses in eukaryotic evolution. Annu
948 Rev Plant Biol 64:583-607.
949 Keeling, PJ, Burki, F, Wilcox, HM, Allam, B, Allen, EE, Amaral-Zettler, LA, Armbrust, EV, Archibald,
950 JM, Bharti, AK, Bell, CJ, et al. 2014. The Marine Microbial Eukaryote Transcriptome Sequencing
951 Project (MMETSP): illuminating the functional diversity of eukaryotic life in the oceans through
952 transcriptome sequencing. PLOS Biol 12:e1001889.
953 Kim, D, Filtz, MR, Proteau, PJ. 2004. The methylerythritol phosphate pathway contributes to carotenoid
954 but not phytol biosynthesis in *Euglena gracilis*. J Nat Prod 67:1067-1069.
955 Kim, GH, Jeong, HJ, Yoo, YD, Kim, S, Han, JH, Han, JW, Zuccarello, GC. 2013. Still acting green:
956 continued expression of photosynthetic genes in the heterotrophic Dinoflagellate *Pfiesteria*
957 *piscicida* (Peridiniales, Alveolata). PLOS ONE 8:e68232.
958 Kiss, JZ, Vasconcelos, AC, Triemer, RE. 1987. Structure of the euglenoid storage carbohydrate, paramylon.
959 Am J Bot 74:877-882.
960 Kitaoka, S, Nakano, Y, Miyatake, K, Yokota, A. 1989. Enzymes and their functional location. In: DE
961 Buetow, editor. Subcellular Biochemistry and Molecular Biology: Academic Press. p. 1-135.
962 Kivic, PA, Vesk, M. 1974. An electron microscope search for plastids in bleached *Euglena gracilis* and in
963 *Astasia longa*. Can J Bot 52:695-699.
964 Kobayashi, K. 2016. Role of membrane glycerolipids in photosynthesis, thylakoid biogenesis and
965 chloroplast development. J Plant Res 129:565-580.
966 Kořený, L, Oborník, M. 2011. Sequence evidence for the presence of two tetrapyrrole pathways in *Euglena*

967 *gracilis*. Genome Biol Evol 3:359-364.
968 Košťál, V, Šimek, P. 1998. Changes in fatty acid composition of phospholipids and triacylglycerols after
969 cold-acclimation of an aestivating insect prepupa. Journal of Comparative Physiology B 168:453-
970 460.
971 Krajčovič, J, Ebringer, L, Schwartzbach, SD. 2002. Reversion of endosymbiosis? In: J Seckbach, editor.
972 Symbiosis: Mechanisms and model systems. Dordrecht: Springer Netherlands. p. 185-206.
973 Krogh, A, Larsson, B, von Heijne, G, Sonnhammer, EL. 2001. Predicting transmembrane protein topology
974 with a hidden Markov model: application to complete genomes. J Mol Biol 305:567-580.
975 Kwong, WK, Del Campo, J, Mathur, V, Vermeij, MJA, Keeling, PJ. 2019. A widespread coral-infecting
976 apicomplexan with chlorophyll biosynthesis genes. Nature 568:103-107.
977 Lakey, B, Triemer, R. 2016. The tetrapyrrole synthesis pathway as a model of horizontal gene transfer in
978 euglenoids. J Phycol 53:198-217.
979 Lambalot, RH, Walsh, CT. 1995. Cloning, overproduction, and characterization of the *Escherichia coli*
980 holo-acyl carrier protein synthase. J Biol Chem 270:24658-24661.
981 Leander, BS, Lax, G, Karnkowska, A, Simpson, AGB. 2017. Euglenida. In: JM Archibald, AGB Simpson,
982 CH Slamovits, L Margulis, M Melkonian, DJ Chapman, JO Corliss, editors. Handbook of the
983 Protists. Cham: Springer International Publishing. p. 1-42.
984 Li, C, Wang, Y, Liu, L, Hu, Y, Zhang, F, Mergen, S, Wang, G, Schlappi, MR, Chu, C. 2011. A rice plastidial
985 nucleotide sugar epimerase is involved in galactolipid biosynthesis and improves photosynthetic
986 efficiency. PLoS Genet 7:e1002196.
987 Lim, L, McFadden, GI. 2010. The evolution, metabolism and functions of the apicoplast. Philos Trans R
988 Soc Lond B Biol Sci 365:749-763.
989 Lohr, M, Schwender, J, Polle, JE. 2012. Isoprenoid biosynthesis in eukaryotic phototrophs: a spotlight on
990 algae. Plant Sci 185-186:9-22.
991 Lochner, K, Doring, O, Bottger, M. 2003. Phylloquinone, what can we learn from plants? Biofactors 18:73-
992 78.
993 Maeda, H, DellaPenna, D. 2007. Tocopherol functions in photosynthetic organisms. Curr Opin Plant Biol
994 10:260-265.
995 Marin, B, Palm, A, Klingberg, M, Melkonian, M. 2003. Phylogeny and taxonomic revision of plastid-
996 containing euglenophytes based on SSU rDNA sequence comparisons and synapomorphic
997 signatures in the SSU rRNA secondary structure. Protist 154:99-145.
998 Markunas, CM, Triemer, RE. 2016. Evolutionary history of the enzymes involved in the Calvin-Benson
999 cycle in euglenids. J Eukaryot Microbiol 63:326-339.
1000 Martins-Duarte, ES, Carias, M, Vommaro, R, Surolia, N, de Souza, W. 2016. Apicoplast fatty acid synthesis

1001 is essential for pellicle formation at the end of cytokinesis in *Toxoplasma gondii*. J Cell Sci
1002 129:3320-3331.
1003 Matsuzaki, M, Kuroiwa, H, Kuroiwa, T, Kita, K, Nozaki, H. 2008. A cryptic algal group unveiled: a plastid
1004 biosynthesis pathway in the oyster parasite *Perkinsus marinus*. Mol Biol Evol 25:1167-1179.
1005 McFadden, GI. 2014. Origin and evolution of plastids and photosynthesis in eukaryotes. Cold Spring Harb
1006 Perspect Biol 6:a016105.
1007 McFadden, GI, Yeh, E. 2017. The apicoplast: now you see it, now you don't. Int J Parasitol 47:137-144.
1008 Miller, LH, Ackerman, HC, Su, XZ, Wellems, TE. 2013. Malaria biology and disease pathogenesis: insights
1009 for new treatments. Nat Med 19:156-167.
1010 Mistry, J, Finn, RD, Eddy, SR, Bateman, A, Punta, M. 2013. Challenges in homology search: HMMER3
1011 and convergent evolution of coiled-coil regions. Nucleic Acids Res 41:e121.
1012 Moffatt, BA, Ashihara, H. 2002. Purine and pyrimidine nucleotide synthesis and metabolism. Arabidopsis
1013 Book 1:e0018.
1014 Moore, RB, Oborník, M, Janouškovec, J, Chrudimský, T, Vancová, M, Green, DH, Wright, SW, Davies,
1015 NW, Bolch, CJ, Heimann, K, et al. 2008. A photosynthetic alveolate closely related to apicomplexan
1016 parasites. Nature 451:959-963.
1017 Nagaraj, VA, Sundaram, B, Varadarajan, NM, Subramani, PA, Kalappa, DM, Ghosh, SK, Padmanaban, G.
1018 2013. Malaria parasite-synthesized heme is essential in the mosquito and liver stages and
1019 complements host heme in the blood stages of infection. PLOS Pathog 9:e1003522.
1020 Neuhaus, HE, Emes, MJ. 2000. Nonphotosynthetic metabolism in plastids. Annu Rev Plant Physiol Plant
1021 Mol Biol 51:111-140.
1022 Nguyen, LT, Schmidt, HA, von Haeseler, A, Minh, BQ. 2015. IQ-TREE: a fast and effective stochastic
1023 algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32:268-274.
1024 Novák Vanclová, AMG, Zoltner, M, Kelly, S, Soukal, P, Záhonová, K, Füssy, Z, Ebenezer, TE, Lacová
1025 Dobáková, E, Eliáš, M, Lukeš, J, et al. 2019. Proteome of the secondary plastid of *Euglena gracilis*
1026 reveals several metabolic quirks and traces of a colourful history. bioRxiv:573709.
1027 Nudelman, MA, Rossi, MS, Conforti, V, Triemer, RE. 2003. Phylogeny of Euglenophyceae based on small
1028 subunit rDNA sequences: Taxonomic implications. J Phycol 39:226-235.
1029 Oborník, M, Green, BR. 2005. Mosaic origin of the heme biosynthesis pathway in photosynthetic
1030 eukaryotes. Mol Biol Evol 22:2343-2353.
1031 Oborník, M, Janouškovec, J, Chrudimský, T, Lukeš, J. 2009. Evolution of the apicoplast and its hosts: from
1032 heterotrophy to autotrophy and back again. Int J Parasitol 39:1-12.
1033 Oda, Y, Miyatake, K, Kitaoka, S. 1979. Inability of *Euglena gracilis* Z to utilize nitrate, nitrite and urea as
1034 the nitrogen sources. Bulletin of the University of Osaka Prefecture. Series B. Agriculture and

1035 Biology 31:43-48.
1036 Patron, NJ, Durnford, DG, Kopriva, S. 2008. Sulfate assimilation in eukaryotes: fusions, relocations and
1037 lateral transfers. BMC Evol Biol 8:39.
1038 Petsalaki, EI, Bagos, PG, Litou, ZI, Hamodrakas, SJ. 2006. PredSL: a tool for the N-terminal sequence-
1039 based prediction of protein subcellular localization. Genomics Proteomics Bioinformatics 4:48-55.
1040 Pick, TR, Brautigam, A, Schulz, MA, Obata, T, Fernie, AR, Weber, AP. 2013. PLGG1, a plastidic glycolate
1041 glycerate transporter, is required for photorespiration and defines a unique class of metabolite
1042 transporters. Proc Natl Acad Sci U S A 110:3185-3190.
1043 Pinnaduwage, P, Bruce, BD. 1996. In vitro interaction between a chloroplast transit peptide and chloroplast
1044 outer envelope lipids is sequence-specific and lipid class-dependent. J Biol Chem 271:32907-
1045 32915.
1046 Pombert, JF, Blouin, NA, Lane, C, Boucias, D, Keeling, PJ. 2014. A lack of parasitic reduction in the
1047 obligate parasitic green alga *Helicosporidium*. PLoS Genet 10:e1004355.
1048 Ponce-Toledo, RI, Deschamps, P, Lopez-Garcia, P, Zivanovic, Y, Benzerara, K, Moreira, D. 2017. An
1049 Early-Branching Freshwater Cyanobacterium at the Origin of Plastids. Curr Biol 27:386-391.
1050 Raines, CA, Lloyd, JC. 2001. C3 carbon reduction cycle. eLS: John Wiley & Sons, Ltd.
1051 Rius, SP, Casati, P, Iglesias, AA, Gomez-Casati, DF. 2006. Characterization of an *Arabidopsis thaliana*
1052 mutant lacking a cytosolic non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase. Plant
1053 Mol Biol 61:945-957.
1054 Roger, AJ, Muñoz-Gómez, SA, Kamikawa, R. 2017. The origin and diversification of mitochondria. Curr
1055 Biol 27:R1177-R1192.
1056 Sanchez-Puerta, MV, Lippmeier, JC, Apt, KE, Delwiche, CF. 2007. Plastid genes in a non-photosynthetic
1057 dinoflagellate. Protist 158:105-117.
1058 Sanz-Luque, E, Chamizo-Ampudia, A, Llamas, A, Galvan, A, Fernandez, E. 2015. Understanding nitrate
1059 assimilation and its regulation in microalgae. Front Plant Sci 6:899.
1060 Seeger, JW, Bentley, R. 1991. Phylloquinone (vitamin K_1) biosynthesis in *Euglena gracilis* strain *Z.*
1061 Phytochemistry 30:3585-3589.
1062 Sekiguchi, H, Moriya, M, Nakayama, T, Inouye, I. 2002. Vestigial chloroplasts in heterotrophic
1063 stramenopiles *Pteridomonas danica* and *Ciliophrys infusionum* (Dictyochophyceae). Protist
1064 153:157-167.
1065 Shashidhara, LS, Smith, AG. 1991. Expression and subcellular location of the tetrapyrrole synthesis enzyme
1066 porphobilinogen deaminase in light-grown *Euglena gracilis* and three nonchlorophyllous cell lines.
1067 Proc Natl Acad Sci U S A 88:63-67.
1068 Schopfer, P, Heyno, E, Drepper, F, Krieger-Liszkay, A. 2008. Naphthoquinone-dependent generation of

1069 superoxide radicals by quinone reductase isolated from the plasma membrane of soybean. Plant
1070 Physiol 147:864.
1071 Schürmann, P, Buchanan, BB. 2008. The ferredoxin/thioredoxin system of oxygenic photosynthesis.
1072 Antioxid Redox Signal 10:1235-1274.
1073 Schwender, J, Goffman, F, Ohlrogge, JB, Shachar-Hill, Y. 2004. Rubisco without the Calvin cycle improves
1074 the carbon efficiency of developing green seeds. Nature 432:779-782.
1075 Siemeister, G, Buchholz, C, Hachtel, W. 1990a. Genes for ribosomal proteins are retained on the 73 kb
1076 DNA from *Astasia longa* that resembles *Euglena* chloroplast DNA. Curr Genet 18:457-464.
1077 Siemeister, G, Buchholz, C, Hachtel, W. 1990b. Genes for the plastid elongation factor Tu and ribosomal
1078 protein S7 and six tRNA genes on the 73 kb DNA from *Astasia longa* that resembles the chloroplast
1079 DNA of *Euglena*. Mol Gen Genet 220:425-432.
1080 Slamovits, CH, Keeling, PJ. 2008. Plastid-derived genes in the nonphotosynthetic alveolate *Oxyrrhis*
1081 *marina*. Mol Biol Evol 25:1297-1306.
1082 Smith, DR, Lee, RW. 2014. A plastid without a genome: evidence from the nonphotosynthetic green algal
1083 genus *Polytomella*. Plant Physiol 164:1812-1819.
1084 Sonneborn, TM. 1950. Methods in the general biology and genetics of *Paramecium aurelia*. Journal of
1085 Experimental Zoology 113:87-147.
1086 Sun, GL, Shen, W, Wen, JF. 2008. Triosephosphate isomerase genes in two trophic modes of euglenoids
1087 (euglenophyceae) and their phylogenetic analysis. J Eukaryot Microbiol 55:170-177.
1088 Ševčíková, T, Horák, A, Klimeš, V, Zbránková, V, Demir-Hilton, E, Sudek, S, Jenkins, J, Schmutz, J, Přibyl,
1089 P, Fousek, J, et al. 2015. Updating algal evolutionary relationships through plastid genome
1090 sequencing: did alveolate plastids emerge through endosymbiosis of an ochrophyte? Sci Rep
1091 5:10134.
1092 Tabita, FR, Hanson, TE, Li, H, Satagopan, S, Singh, J, Chan, S. 2007. Function, structure, and evolution of
1093 the RubisCO-like proteins and their RubisCO homologs. Microbiol Mol Biol Rev 71:576-599.
1094 Tomčala, A, Kyselová, V, Schneedorferová, I, Opekarová, I, Moos, M, Urajová, P, Kručinská, J, Oborník,
1095 M. 2017. Separation and identification of lipids in the photosynthetic cousins of Apicomplexa
1096 *Chromera velia* and *Vitrella brassicaformis*. J Sep Sci 40:3402-3413.
1097 Tripathy, BC, Sherameti, I, Oelmuller, R. 2010. Siroheme: an essential component for life on earth. Plant
1098 Signal Behav 5:14-20.
1099 Tucci, S, Vacula, R, Krajčovič, J, Proksch, P, Martin, W. 2010. Variability of wax ester fermentation in
1100 natural and bleached *Euglena gracilis* strains in response to oxygen and the elongase inhibitor
1101 flufenacet. J Eukaryot Microbiol 57:63-69.
1102 Turmel, M, Gagnon, MC, O'Kelly, CJ, Otis, C, Lemieux, C. 2009. The chloroplast genomes of the green

1103 algae *Pyramimonas*, *Monomastix*, and *Pycnococcus* shed new light on the evolutionary history of
1104 prasinophytes and the origin of the secondary chloroplasts of euglenids. Mol Biol Evol 26:631-648.
1105 Van Dingenen, J, Blomme, J, Gonzalez, N, Inzé, D. 2016. Plants grow with a little help from their organelle
1106 friends. J Exp Bot 67:6267-6281.
1107 Vogel, A, Schwacke, R, Denton, AK, Usadel, B, Hollmann, J, Fischer, K, Bolger, A, Schmidt, MH, Bolger,
1108 ME, Gundlach, H, et al. 2018. Footprints of parasitism in the genome of the parasitic flowering
1109 plant *Cuscuta campestris*. Nat Commun 9:2515.
1110 Vollmer, M, Thomsen, N, Wiek, S, Seeber, F. 2001. Apicomplexan parasites possess distinct nuclear-
1111 encoded, but apicoplast-localized, plant-type ferredoxin-NADP+ reductase and ferredoxin. J Biol
1112 Chem 276:5483-5490.
1113 Walker, BJ, South, PF, Ort, DR. 2016. Physiological evidence for plasticity in glycolate/glycerate transport
1114 during photorespiration. Photosynth Res 129:93-103.
1115 Waller, RF, Gornik, SG, Kořený, L, Pain, A. 2016. Metabolic pathway redundancy within the
1116 apicomplexan-dinoflagellate radiation argues against an ancient chromalveolate plastid. Commun
1117 Integr Biol 9:e1116653.
1118 Watanabe, F, Yoshimura, K, Shigeoka, S. 2017. Biochemistry and physiology of vitamins in *Euglena*. In:
1119 SD Schwartzbach, S Shigeoka, editors. *Euglena*: Biochemistry, Cell and Molecular Biology. Cham:
1120 Springer International Publishing. p. 65-90.
1121 Webster, DA, Hackett, DP, Park, RB. 1967. The respiratory chain of colorless algae: III. Electron
1122 microscopy. J Ultrastruct Res 21:514-523.
1123 Wedel, N, Soll, J, Paap, BK. 1997. CP12 provides a new mode of light regulation of Calvin cycle activity
1124 in higher plants. Proc Natl Acad Sci U S A 94:10479-10484.
1125 Weinstein, JD, Beale, SI. 1983. Separate physiological roles and subcellular compartments for two
1126 tetrapyrrole biosynthetic pathways in *Euglena gracilis*. J Biol Chem 258:6799-6807.
1127 Wicke, S, Müller, KF, de Pamphilis, CW, Quandt, D, Wickett, NJ, Zhang, Y, Renner, SS, Schneeweiss,
1128 GM. 2013. Mechanisms of functional and physical genome reduction in photosynthetic and
1129 nonphotosynthetic parasitic plants of the broomrape family. The Plant Cell 25:3711.
1130 Williamson, DH, Gardner, MJ, Preiser, P, Moore, DJ, Rangachari, K, Wilson, RJM. 1994. The evolutionary
1131 origin of the 35 kb circular DNA of *Plasmodium falciparum*: new evidence supports a possible
1132 rhodophyte ancestry. Molecular and General Genetics MGG 243:249-252.
1133 Wilson, RJ, Denny, PW, Preiser, PR, Rangachari, K, Roberts, K, Roy, A, Whyte, A, Strath, M, Moore, DJ,
1134 Moore, PW, et al. 1996. Complete gene map of the plastid-like DNA of the malaria parasite
1135 *Plasmodium falciparum*. J Mol Biol 261:155-172.
1136 Yeh, E, DeRisi, JL. 2011. Chemical rescue of malaria parasites lacking an apicoplast defines organelle

1137 function in blood-stage *Plasmodium falciparum*. PLOS Biol 9:e1001138.
1138 Yoon, EY, Yang, AR, Park, J, Moon, SJ, Jeong, EJ, Rho, JR. 2017. Characterization of a new trioxilin and
1139 a sulfoquinovosyl diacylglycerol with anti-inflammatory properties from the dinoflagellate
1140 *Oxyrrhis marina*. Mar Drugs 15.
1141 Yoshida, Y, Tomiyama, T, Maruta, T, Tomita, M, Ishikawa, T, Arakawa, K. 2016. De novo assembly and
1142 comparative transcriptome analysis of *Euglena gracilis* in response to anaerobic conditions. BMC
1143 Genomics 17:182.
1144 Yuzawa, Y, Nishihara, H, Haraguchi, T, Masuda, S, Shimojima, M, Shimoyama, A, Yuasa, H, Okada, N,
1145 Ohta, H. 2012. Phylogeny of galactolipid synthase homologs together with their enzymatic analyses
1146 revealed a possible origin and divergence time for photosynthetic membrane biogenesis. DNA Res
1147 19:91-102.
1148 Záhonová, K, Füssy, Z, Birčák, E, Novák Vanclová, AMG, Klimeš, V, Vesteg, M, Krajčovič, J, Oborník,
1149 M, Eliáš, M. 2018. Peculiar features of the plastids of the colourless alga *Euglena longa* and
1150 photosynthetic euglenophytes unveiled by transcriptome analyses. Sci Rep 8:17012.
1151 Záhonová, K, Füssy, Z, Oborník, M, Eliáš, M, Yurchenko, V. 2016. RuBisCO in non-photosynthetic alga
1152 *Euglena longa*: divergent features, transcriptomic analysis and regulation of complex formation.
1153 PLOS ONE 11:e0158790.
1154 Ziegler, K, Maldener, I, Lockau, W. 1989. 5'-monohydroxyphylloquinone as a component of photosystem
1155 I. Zeitschrift für Naturforschung C 44:468-472.
1156 Zimorski, V, Rauch, C, van Hellemond, JJ, Tielens, AGM, Martin, WF. 2017. The mitochondrion of
1157 *Euglena gracilis*. In: SD Schwartzbach, S Shigeoka, editors. *Euglena*: Biochemistry, cell and
1158 molecular biology. Cham: Springer International Publishing. p. 19-37.
1159