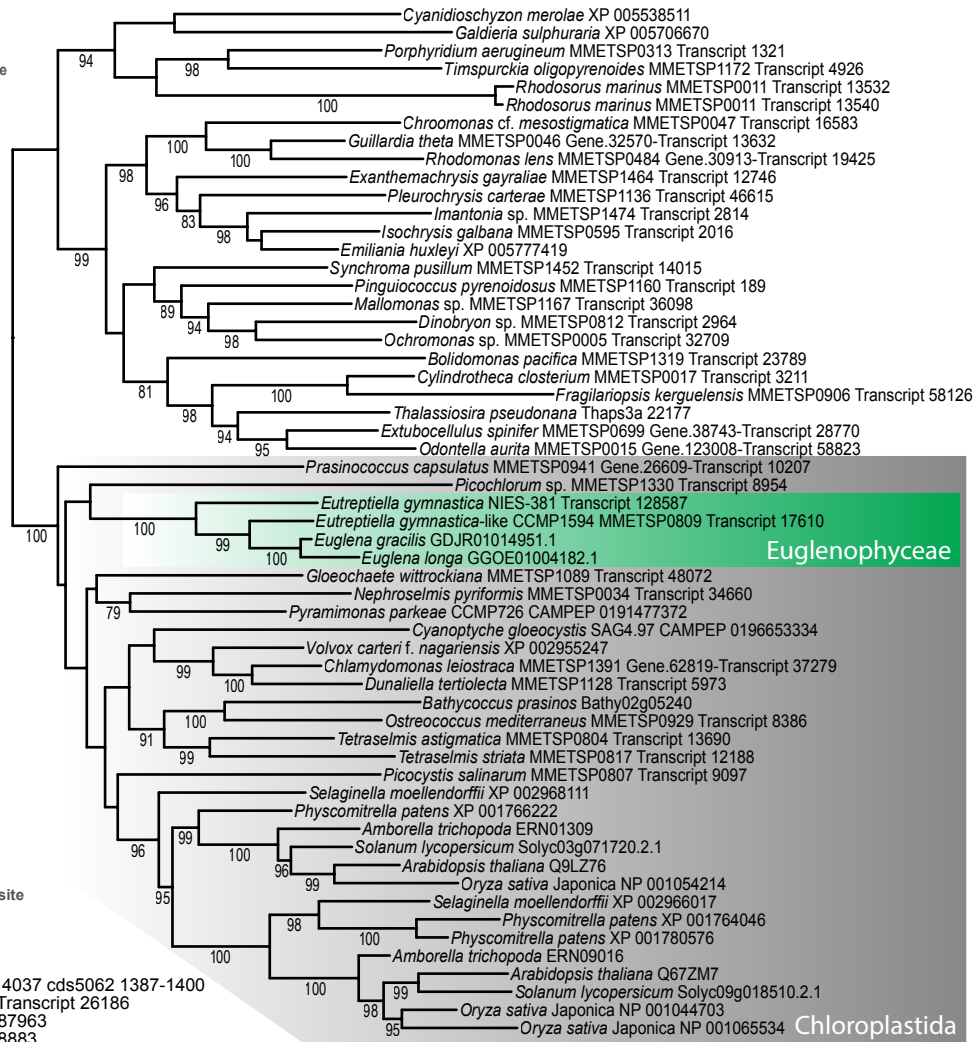


VTE5

0.2 substitution/site



VTE6

0.2 substitution/site



Figure S1: Inferred phylogeny of phytol kinase VTE 5 and phytyl-phosphate kinase VTE6. The maximum-likelihood tree was inferred with IQ-TREE using the LG+I+G4 substitution model and ultra-fast bootstrapping. The UFboot support values are indicated at branches when higher than 75%. Colour of the Euglenophyceae clade reflects possible origin in Chloroplastida (green) or Rhodophyta-derived complex algae (ochre).

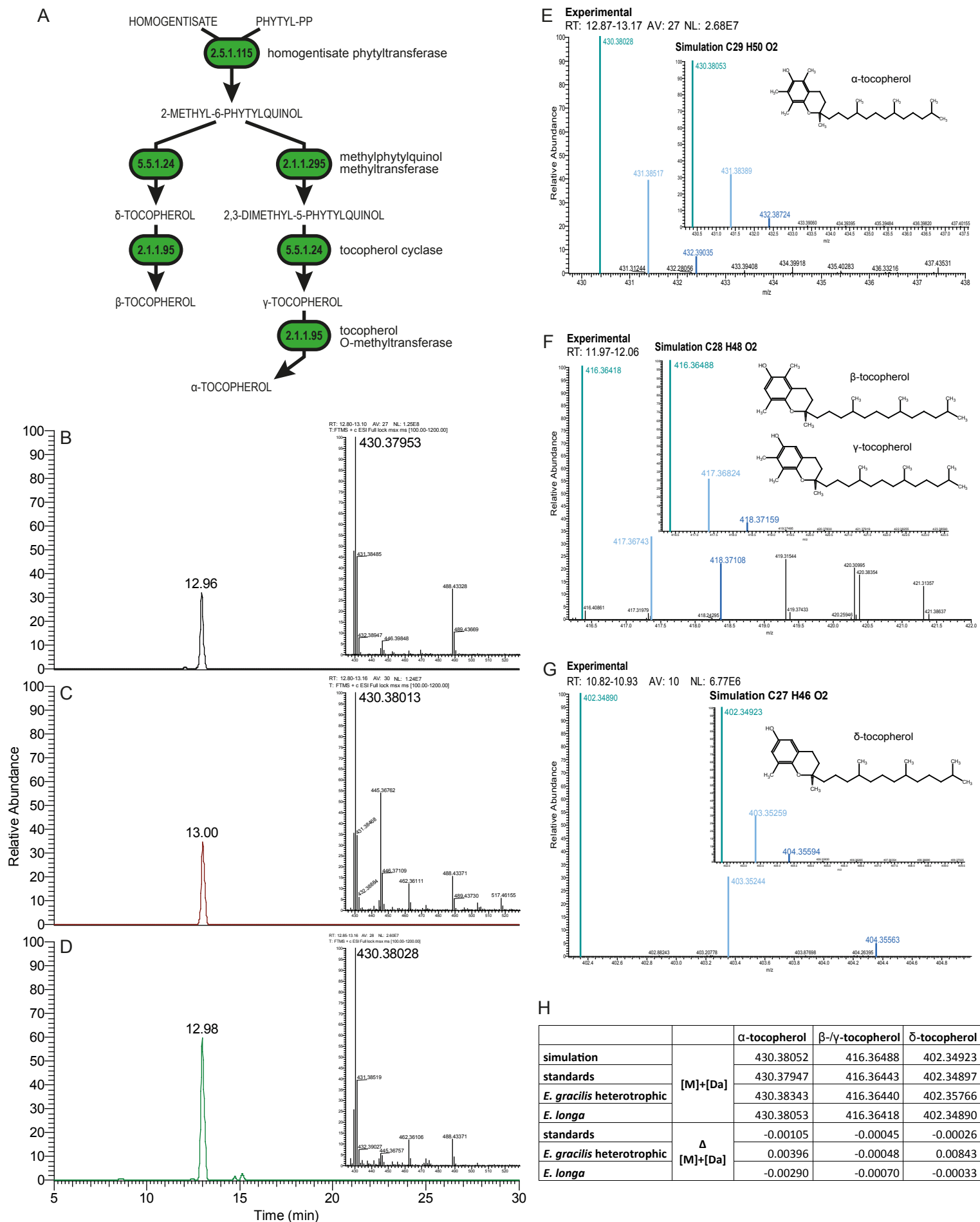
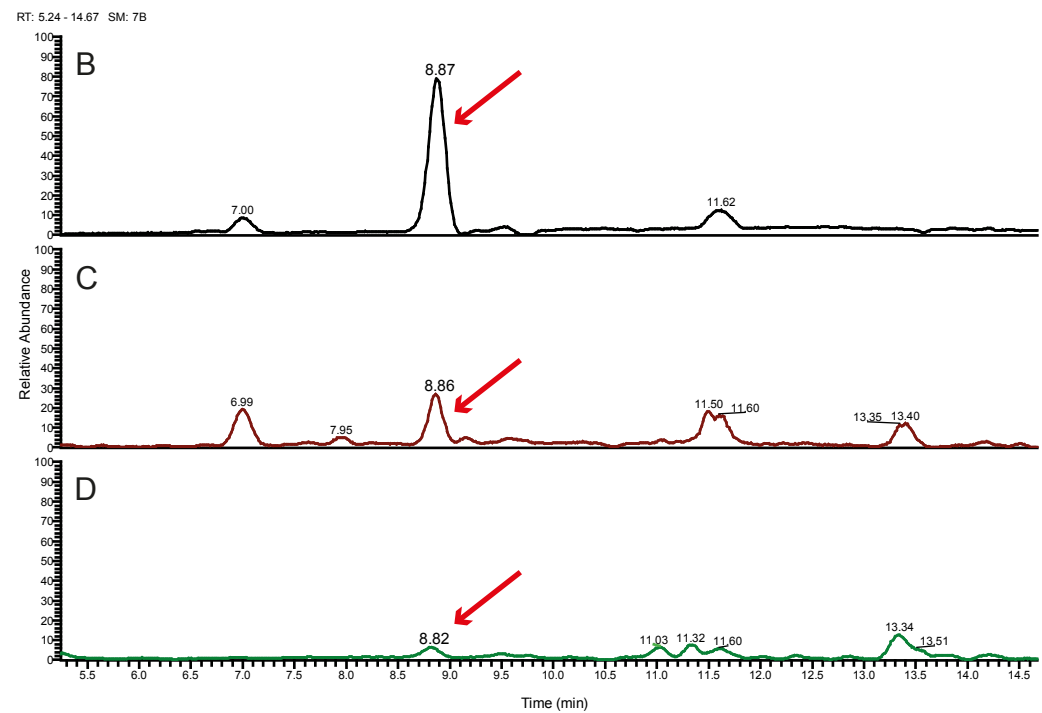
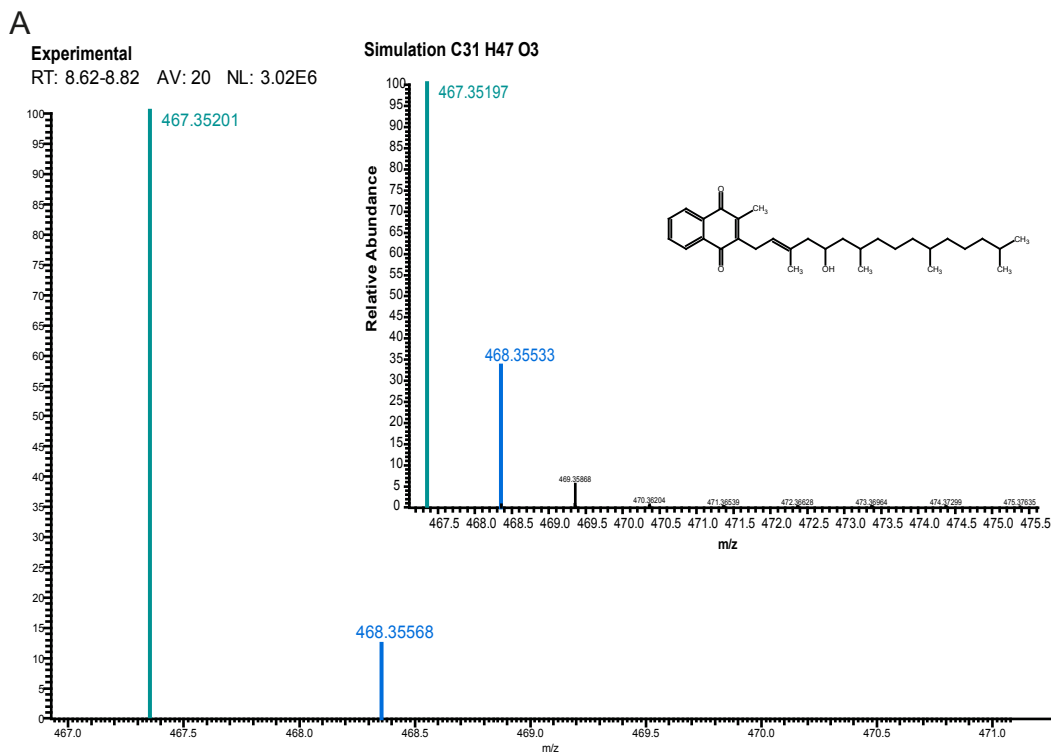


Figure S2: Experimental detection of tocopherols in *E. longa* and *E. gracilis*. **A:** Overview of tocopherol biosynthesis (enzymes indicated by their EC numbers). **B-D:** Extracted chromatograms of exact mass of α-tocopherol (m/z 430.3805; inset) and recorded spectra from a particular peak in an α-tocopherol standard (B), raw lipid extracts of heterotrophic *E. gracilis* (C) and *E. longa* (D). **E-G:** Comparison of simulation of α- (E), β- and γ- (F), and δ-tocopherol (G) chemical formula spectrum and experimentally gained data from the *E. longa* sample. The chemical structure of the particular tocopherol is shown as inset. **H:** Comparison of tocopherols' monoisotopic mass simulated and obtained high-resolution data in examined euglenophyte samples.



E

5 - hydroxyphyllolquinone	[M+H] ⁺ [Da]	Δ [M+H] ⁺ [Da]
simulation	467.35197	
<i>E. gracilis</i> autotrophic	467.35191	0.00006
<i>E. gracilis</i> heterotrophic	467.35196	0.00001
<i>E. longa</i>	467.35202	-0.00005

Figure S3. Experimental confirmation of 5-hydroxyphyllolquinone (OH-PhQ) in *E. longa* and *E. gracilis*. **A:** Comparison of simulation of OH-PhQ chemical formula spectrum and experimentally gained data from autotrophic *E. gracilis* sample. The chemical structure of OH-PhQ is shown as inset. **B-D:** Extracted chromatograms of exact mass of protonated OH-PhQ (m/z 467.35) of raw lipid extracts. The red arrow points to a peak of OH-PhQ determined by high-resolution and fragmentation pattern in autotrophic *E. gracilis* (B), heterotrophic *E. gracilis* (C), and *E. longa* (D). **E:** Comparison of OH-PhQ monoisotopic mass simulated and obtained high-resolution data in examined euglenophyte samples.

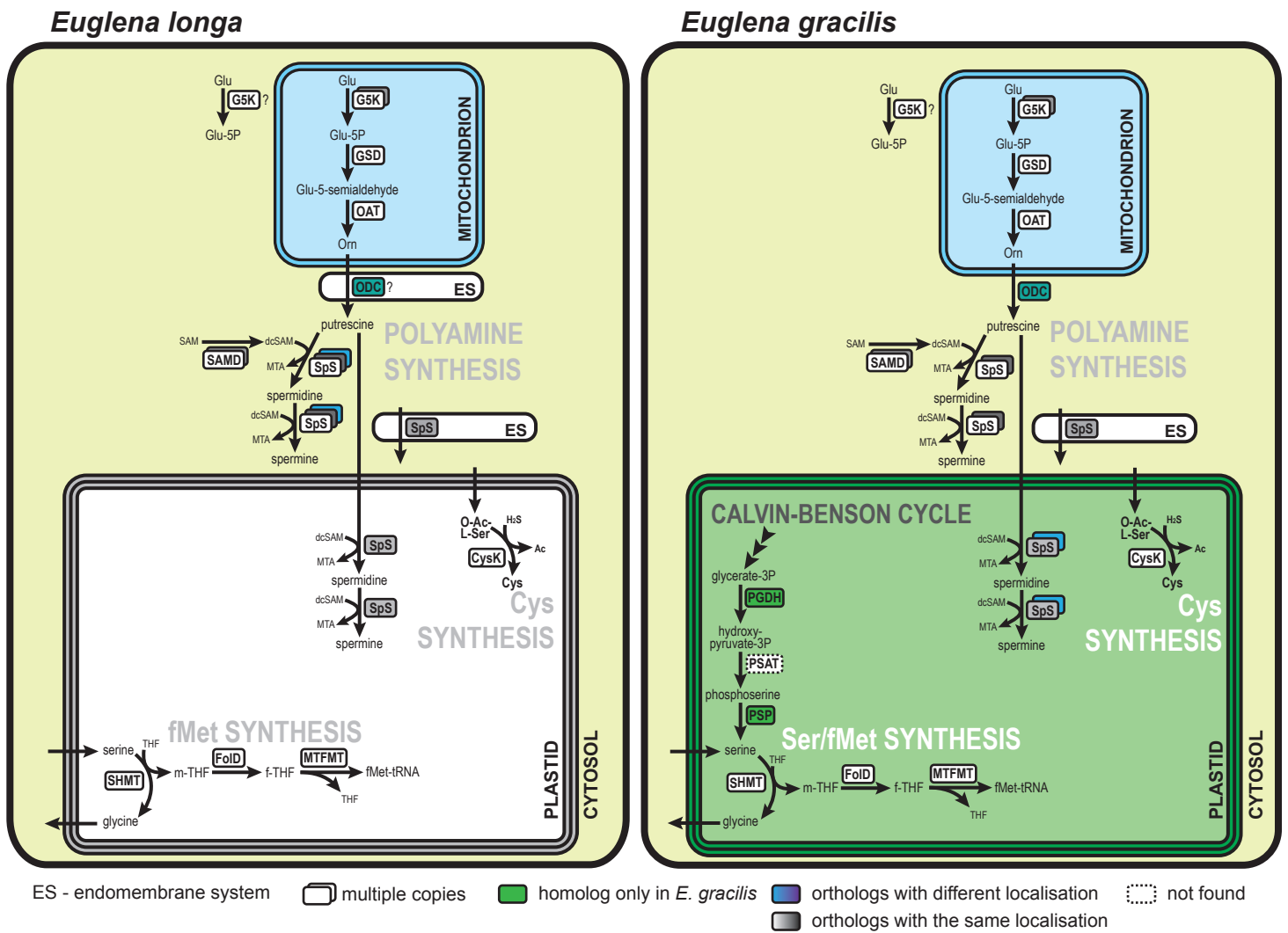


Figure S4: Plastid-linked amino acid metabolism in *E. longa* and *E. gracilis*. Schematic comparison of the localization of enzymes of plastid serine, formylmethionine, cysteine and polyamine synthesis. Abbreviations, CysK – cysteine synthase A; FoID – bifunctional methylenetetrahydrofolate dehydrogenase (NADP+) / cyclohydrolase; G5K – glutamate 5-kinase; GSD – glutamate semialdehyde dehydrogenase; MTA – 5'-methylthioadenosine; MTFMT – Met-tRNA formyltransferase; OAT – ornithine--oxo-acid transaminase; ODC – ornithine decarboxylase; PSAT – phosphoserine aminotransferase; PSP – phosphoserine phosphatase; (dc)SAM – (decarboxy-)S-adenosylmethionine; SAMD – SAM decarboxylase; SHMT – serine hydroxymethyltransferase; SpS – spermidine/spermin synthase; (f-/m-)THF – (10-formyl-/5,10-methylene-)tetrahydrofolate.

Uroporphyrinogen III synthase

0,3 substitution/site

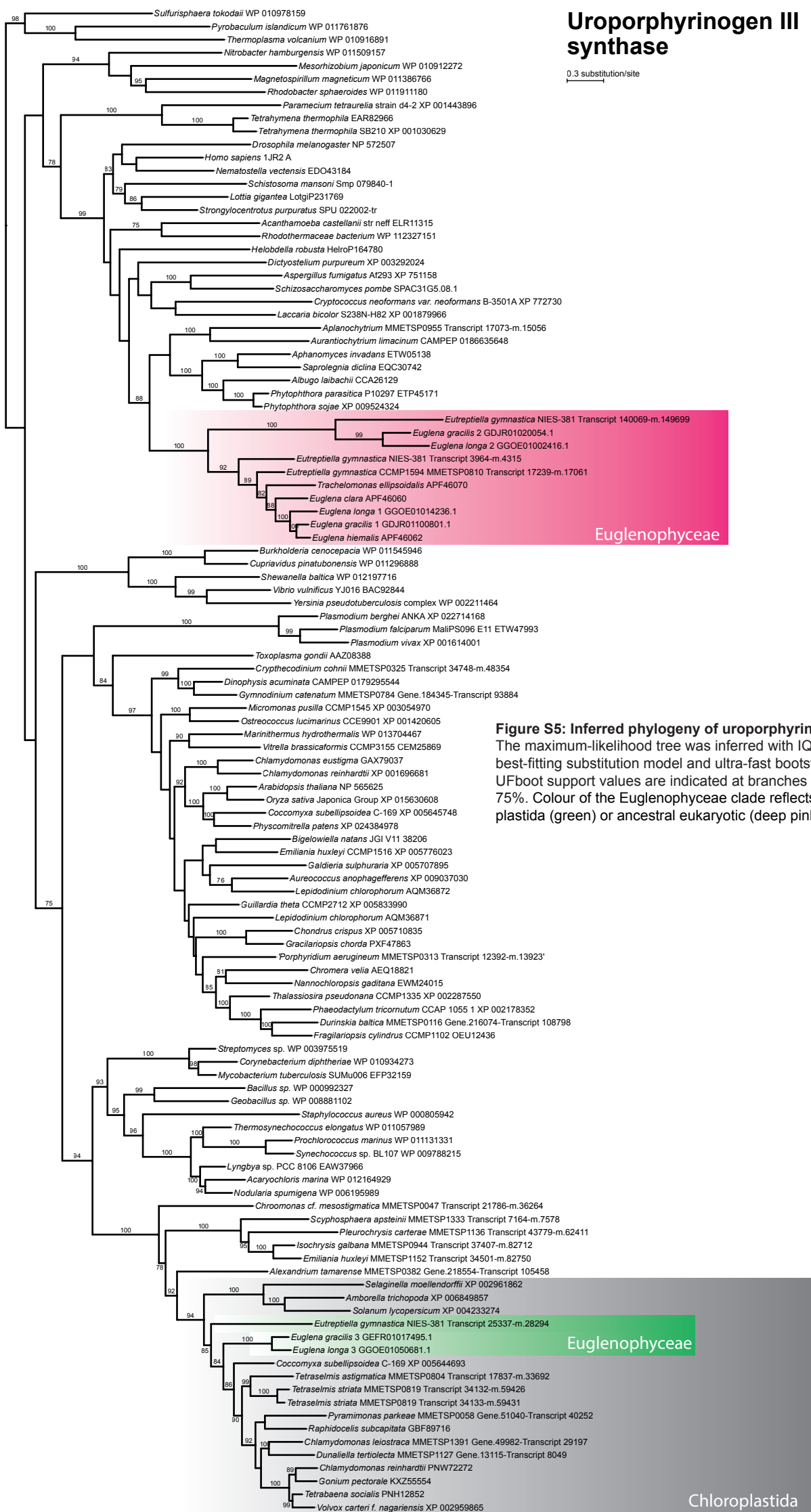


Figure S5: Inferred phylogeny of uroporphyrinogen III synthase. The maximum-likelihood tree was inferred with IQ-TREE using the best-fitting substitution model and ultra-fast bootstrapping. The UFboot support values are indicated at branches when higher than 75%. Colour of the Euglenophyceae clade reflects possible Chloroplastida (green) or ancestral eukaryotic (deep pink) origin.

Euglenophyceae

Euglenophyceae

Chloroplastida

UDP-glucose epimerase

0.1 substitutions/site

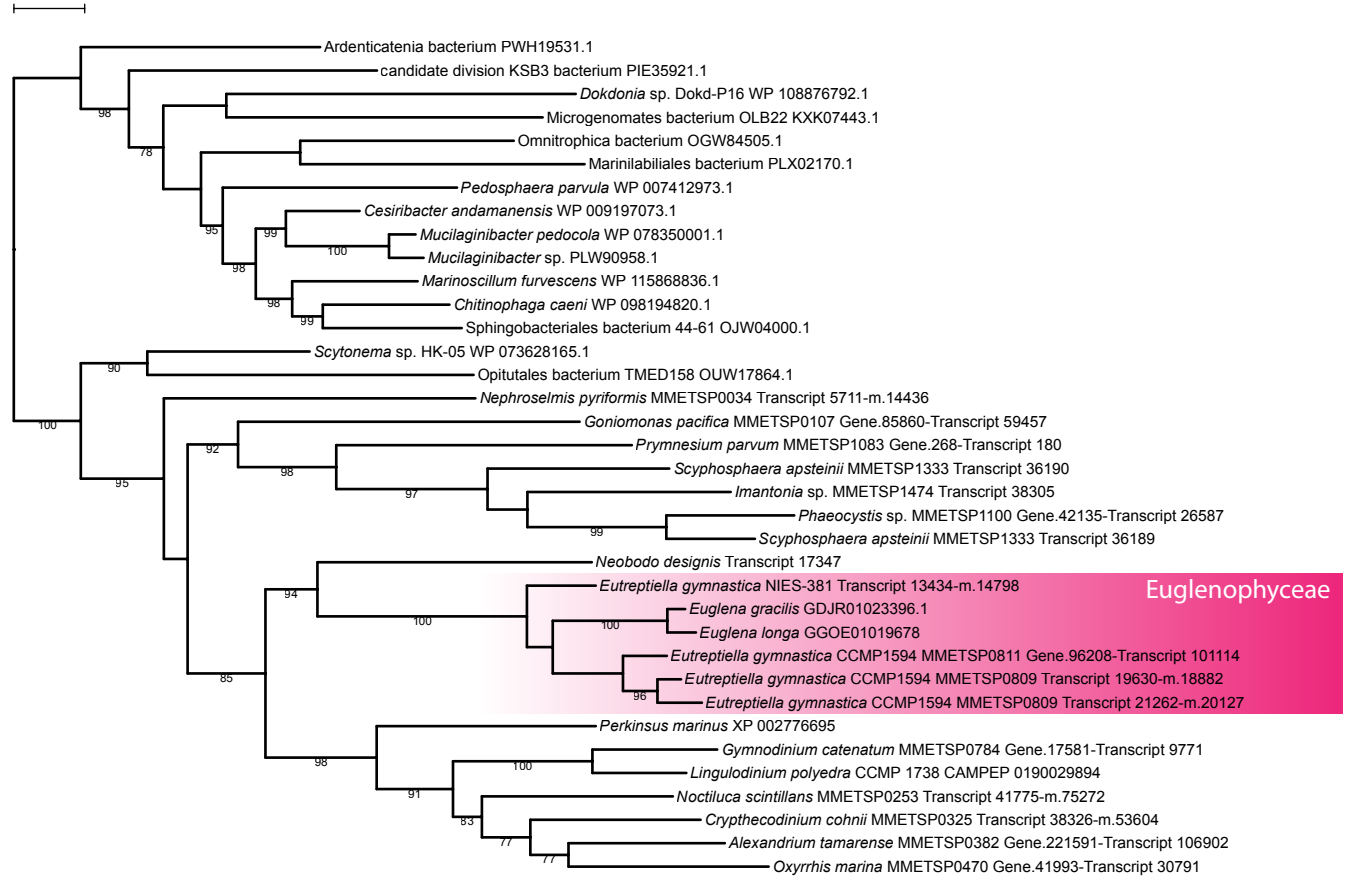


Figure S6: Inferred phylogeny of the putative plastid UDP-glucose epimerase. The maximum-likelihood tree was inferred with IQ-TREE using the best-fitting substitution model and ultra-fast bootstrapping. The UFboot support values are indicated at branches when higher than 75%. Colour of the Euglenophyceae clade reflects possible ancestral eukaryotic origin (deep pink).

triose-phosphate transporters

1 substitution/site

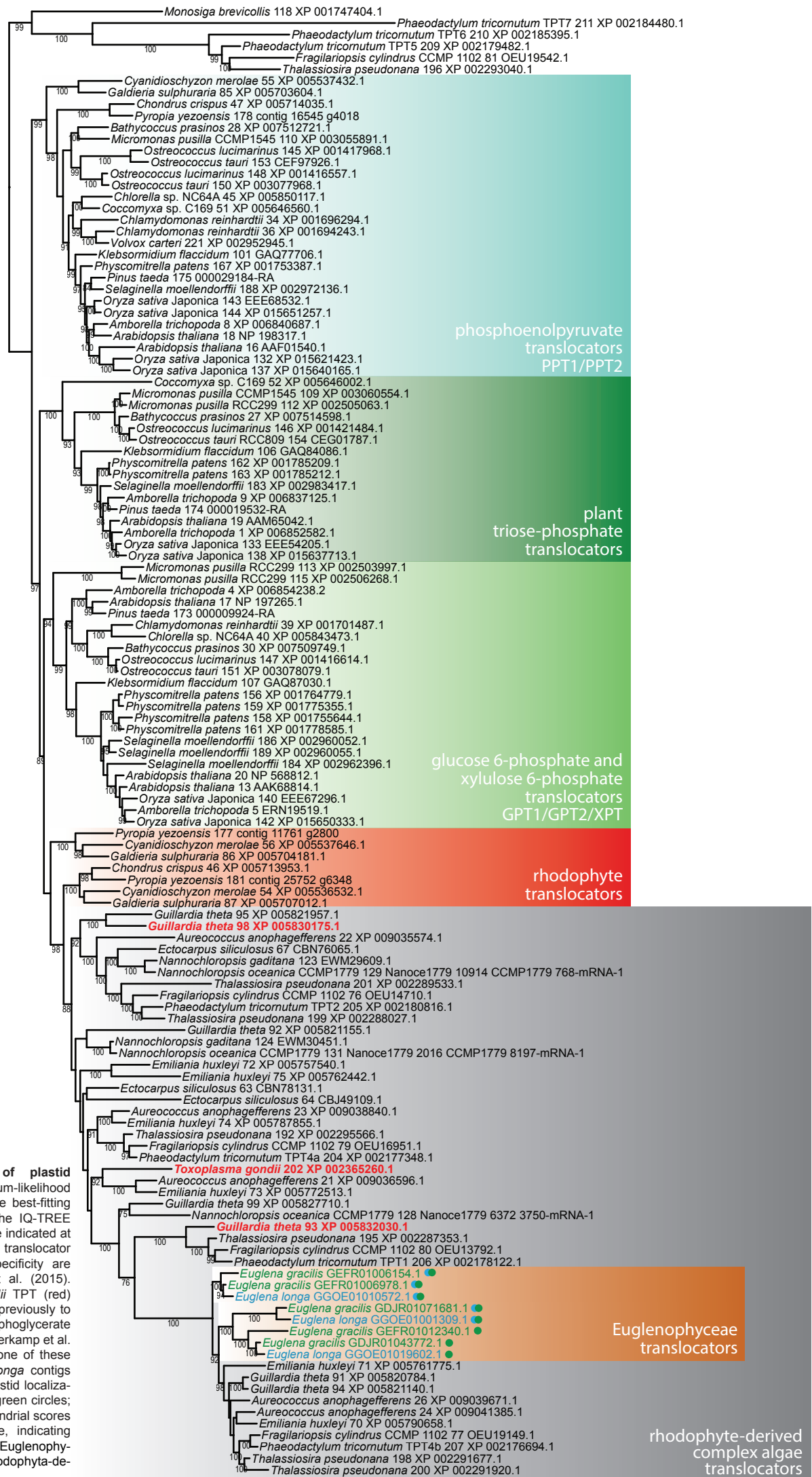
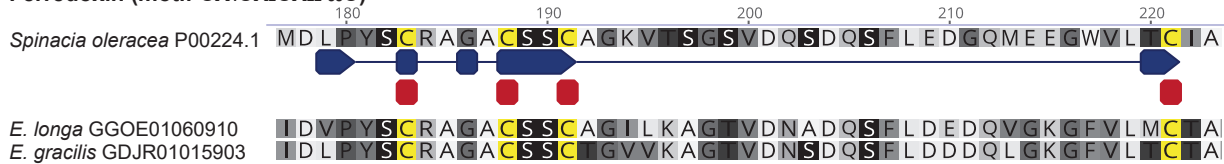


Figure S7: Inferred phylogeny of plastid phosphate translocators. The maximum-likelihood tree was inferred with RAxML using the best-fitting substitution model as determined by the IQ-TREE software. The UFboot support values are indicated at branches when higher than 75%. Plant translocator families with determined substrate specificity are marked. Dataset taken from Moog et al. (2015). *Guillardia theta* and *Toxoplasma gondii* TPT (red) substrate specificities were determined previously to be DHAP and triose-phosphate/3-phosphoglycerate /phosphoenolpyruvate, respectively (Haferkamp et al. 2006; Brooks et al. 2010). However, none of these translocators clusters closely with *E. longa* contigs (blue) or *E. gracilis* (green). Putative plastid localizations of these sequences is marked by green circles; sequences with high plastid and mitochondrial scores are marked by green and blue circle, indicating possible dual targeting. Colour of the Euglenophyceae clade reflects possible origin in Rhodophyta-derived complex algae (ochre).

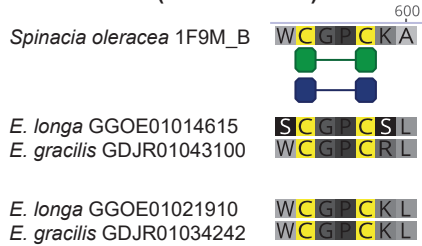
Ferredoxin (motif CX₄CX₂CX₂₂₋₃₃C)



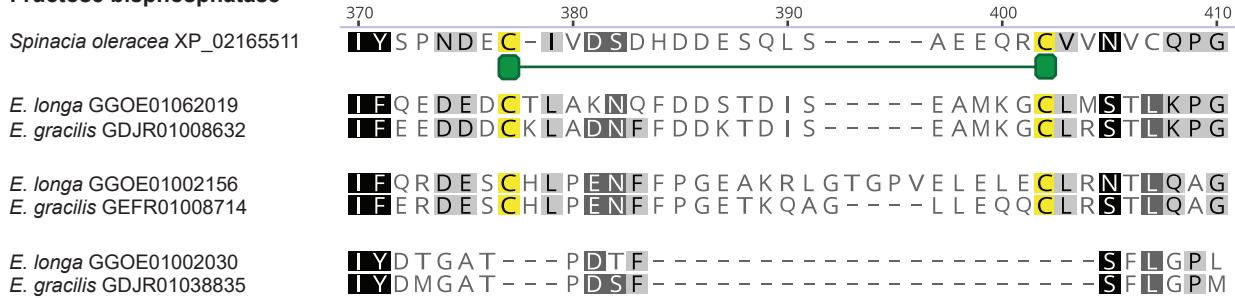
Ferredoxin-thioredoxin reductase (motif CPCX₁₆CPCX₆CHC)



Thioredoxin f (motif WCGPC)



Fructose biphosphatase



Phosphoribulokinase

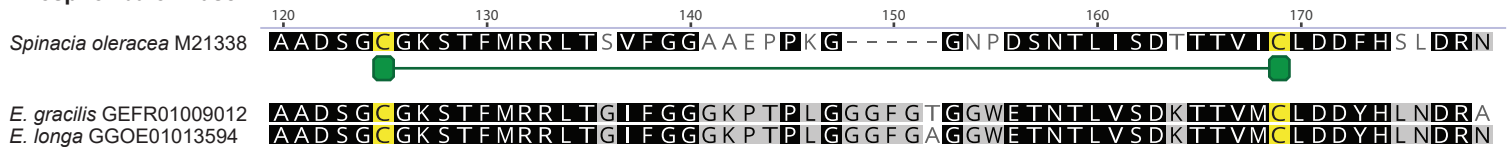


Figure S8: Conserved cysteine motifs in ferredoxin/thioredoxin system and their target proteins of Calvin-Benson cycle. Shown are alignments of *E. longa* and *E. gracilis* plastid-localized homologs with reference sequences from *Spinacia oleracea*. Connected blue boxes - catalytic residues; red boxes - iron/sulfur cluster binding site; green boxes - redox active cysteine bonds.