The net-like heterotrophic amoeba *Leukarachnion salinum* sp. nov. (Ochrophyta, Stramenopiles) has a cryptic plastid

Karin Jaške[1†], Dovilė Barcytė[1†], Tomáš Pánek[1], Tereza Ševčíková[1], Anežka Eliášová[1], Marek Eliáš[1*]

[1]Department of Biology and Ecology, Faculty of Science, University of Ostrava, Chittussiho 10, 710 00 Ostrava, Czech Republic
[†]these authors have contributed equally to this work
[*]corresponding author
e-mail: marek.elias@osu.cz

Abstract
Numerous lineages of heterotrophs evolved from photosynthetic eukaryotic ancestors and usually have retained a plastid, although cases of secondary plastid loss are known, too. Based on a previous investigation by transmission electron microscopy (TEM), *Leukarachnion* sp. PRA-24, an amoeboid colourless protist related to the photosynthetic class Synchromophyceae within the algal phylum Ochrophyta, is a candidate for another case of a plastid loss. Here we provide a detailed characterisation of this organism and formally describe it as *Leukarachnion salinum*, sp. nov. While we could not find any unambiguous candidate for a plastid organelle by TEM, genome sequencing recovered a complete plastid genome with a reduced gene set similar to plastid genomes of other non-photosynthetic ochrophytes yet even more extreme in the sequence divergence. The presence in the *L. salinum* transcriptome assembly of homologs of hallmark plastid proteins, including components of the plastid protein import complexes, supported the notion that *L. salinum* has a cryptic plastid organelle. Based on an analysis of plastid-targeting signals in *L. salinum* proteins, the plastid presumably contains a unique combination of biosynthetic pathways producing haem, the folate precursor aminodeoxychorismate, and, surprisingly, tocotrienols. Our work thus uncovers the existence of a novel form of a relict plastid organelle.

Keywords: Ochrophyta, *Leukarachnion*, non-photosynthetic plastid, plastid evolution, plastid genome, Stramenopiles

1. Background
In the turmoil of the Second World War, Lothar Geitler working in Vienna discovered and described *Leukarachnion batrachospermi*, a freshwater colourless amoeboid organism forming large anastomosing networks and walled cysts (Geitler 1942). The phylogenetic position of this organism could not be properly deduced with the knowledge of protist phylogeny and methods available at that time, and no further study that would illuminate the identity of *Leukarachnion* appeared for the next 65 years. A critical step forward was eventually made by the work of Grant et al. (2009), who studied a strain originally isolated from a salt marsh in Virginia and identified as a potentially new species of the genus *Leukarachnion*. Phylogenetic analysis of four nuclear gene sequences obtained from *Leukarachnion* sp. PRA-24 revealed that it is a stramenopile, specifically a secondarily heterotrophic ochrophyte most closely related to the amoeboid algae *Chlamydomyxa labyrinthuloides* and *Synchroma grande*. This raised a possibility that it has retained a colourless plastid, but Grant et al. did not find any candidates for it when studying the *Leukarachion* ultrastructure by transmission electron microscopy (TEM). No further details on this organism have been published since then, leaving the status of the plastid, and generally the degree of cellular and metabolic reduction related to the loss of photosynthesis, unknown.

Leukarachnion is not the sole colourless member of the predominantly photosynthetic algal group Ochrophyta. The loss of photosynthesis is a particularly common phenomenon in the class Chrysophyceae, where at least 13 independently evolved non-photosynthetic lineages have been documented so far (Dorrell et al. 2019). Other ochrophyte groups lose photosynthesis less commonly. Two such lineages were described within diatoms (Frankovich et al. 2018; Onyshchenko et al. 2019), at least two separate apochlorotic lineages are known in the class Dictyochophyceae (Sekiguchi et al.

2002; Kayama et al. 2020), and the class Phaeophyceae includes at least one lineage of non-photosynthetic parasites or endophytes of other brown algae (Heesch et al. 2008; Bringloe et al. 2021). In addition, ochrophytes include *Picophagus flagellates*, a tiny marine heterotrophic flagellate that based on 18S rRNA phylogenies constitutes an independent lineage of its own (Guillou et al. 1999), and there is molecular phylogenetic evidence suggesting that the heterotrophic predatory actinophryids, a subgroup of the traditional protozoan group Heliozoa, belongs to the Ochrophyta, potentially closest to the class Raphidophyceae (Cavalier-Smith and Scoble 2013).

Non-photosynthetic ochrophytes span a broad gradient of reduction of photosynthesis-related structures and functions. The chrysophyte *Cornospumella fuschlensis* still retains many plastid-targeted proteins associated with photosystems and light-harvesting complexes, although it lacks enzymes of reductive CO_2 fixation (Dorrell et al. 2019). Most other colourless ochrophytes studied in a sufficient detail lack any traces of the photosynthetic machinery, yet still contain reduced plastids (leucoplasts) with their own genome encoding some critical components of required for specific functional pathways taking place in the organelle (Kamikawa et al. 2015; Kamikawa et al. 2017; Kamikawa et al. 2018; Dorrell et al. 2019; Onyshchenko et al. 2019; Kayama et al. 2020; Bringloe et al. 2021). The set of functional processes retained by the ochrophyte leucoplasts differ between the different independently evolved non-photosynthetic lineages. On the one hand, it may theoretically still include CO_2 fixation, as suggested before for the colourless dictyochophytes *Pteridomonas danica* and *Ciliophrys infusionum* based on identification of the *rbcL* gene encoding the large subunit of the enzyme RuBisCO (Sekiguchi et al. 2002); strikingly though, more recent investigations suggested that these *rbcL* sequences might have been contaminants (Kayama et al. 2020). The most extreme reduction has been encountered in the chrysophyte genus *Paraphysomonas*, which contains a vestigial plastid without a genome and hence lacking also the apparatus involved in transcription and translation (Dorrell et al. 2019). Whether there are any ochrophytes that have secondarily lost the plastid as such is presently unknown. Potential candidates are not only *Leukarachnion* sp. PRA-24, but also *P. flagellates* and actinophryids, as no structures identifiable as vestigial plastids have been observed in TEM preparations (Guillou et al. 1999; Grant et al. 2009; Cavalier-Smith and Scoble 2013). However, no leucoplast could be identified by TEM in *C. infusionum* despite the putative molecular evidence (the *rbcL* gene amplified from the organism) for the presence of a plastid genome (Sekiguchi et al. 2002; but see Kayama et al. 2020), indicating that the apparent absence of the leucoplast at the cytological level must be interpreted with caution.

In this study we set out to test the presence of a cryptic plastid in *Leukarachnion* sp. PRA-24 by molecular approaches. Specifically, we generated transcriptomic and genomic data from the organism and searched for molecular signatures of a plastid. This effort proved fruitful and here we report a complete sequence of a plastid genome of *Leukarachnion* sp. PRA-24 and an *in silico* reconstruction of the metabolic functions of the elusive plastid organelle beyond any doubts present in this peculiar ochrophyte. In addition, we provide a detailed morphological and ultrastructural characterization of the organisms and formally describe it as a species new to science.

2. Material and Methods

2.1. Cultivation and microscopy

Leukarachnion sp. PRA-24 was obtained from The American Type Culture Collection (ATCC®; www.atcc.org). The strain was cultivated in Thermo Scientific™ Biolite 25 cm^2 cell culture flasks with vented lids (Thermo Fisher Scientific) in seawater cereal grass media: cereal grass (2,5 g/l), NaCl (24,72 g/l), KCl (0,68 g/l), $CaCl_2 \cdot 2H_2O$ (1,36 g/l), $MgCl_2 \cdot 6H_2O$ (4,66 g/l), $MgSO_4 \cdot 7H_2O$ (6,29 g/l), $NaHCO_3$ (0,18 g/l). The cultures were kept in dark at room temperature and they contained bacteria that served as a food source for *Leukarachnion* cells. The cultures of *Leukarachnion* sp. PRA-24 were examined using an Olympus BX53 (Olympus, Tokyo, Japan) fluorescence microscope employing differential interference contrast (DIC). Micrographs were captured using an Olympus DP73 digital camera. Cells were measured using an Olympus micro imagining software cellSens v1.6. For the transmission electron microscopy (TEM) observations, the cells were prepared using high-pressure freezing (HPF). Centrifuged cells were immersed in a cryoprotectant (20% BSA in cell medium) and immediately frozen with a Leica HPM100 high-pressure freezer (Leica Microsystems, Vienna,

Austria). The cells were freeze substituted in 2% OsO_4 and 100% acetone using an automatic freeze substitution system Leica EM AFS2. The sample was then washed three times in 100% acetone followed by a gradual embedding in EMbed-812 resin at room temperature (resin:acetone – 1:2, 1:1, and 2:1 for one hour per each) with a final overnight embedding in 100% resin. Finally, it was embedded into a fresh degassed resin and polymerized at 60°C for 48 hours. Thin sections were cut on a Reichert-Jung Ultracut E ultramicrotome (Reichert-Jung, Vienna, Austria) and stained using uranyl acetate and lead citrate. Sections were examined and photographed using a JEOL JEM-1011 (JEOL, Tokyo, Japan) electron microscope, equipped with a Veleta camera and the iTEM 5.1 software (Olympus Soft Imaging Solution GmbH, Münster, Germany).

2.2. RNA and DNA extraction

For RNA extraction, 100 ml of the culture was used, after spinning at 1,500 g for 15 min, the pellet was resuspended in 1ml of TRI Reagent® (Genbiotech), incubated for 5 min, centrifuged at 12,000 g for 15 min on 4 °C and the supernatant was transferred to a fresh tube. After adding of 200 ul of chloroform, shaking, and 15 min of incubation at room temperature, the sample was spun at 12,000 g for 15 min on 4 °C and the aqueous phase was transferred to a fresh tube. It was precipitated with 500 ul of isopropanol, incubated for 30 min on -80 °C, then spun at 12000 g for 8 min on 4 °C, the supernatant was discarded, the pellet washed with 70 % ethanol, dried and dissolved in 50 µl of RNase free water on 55 °C for 15 min. The concentration was checked on NanoDrop Lite Spectrophotometer (Thermo Scientific), after DNase treatment (Thermo Scientific), the concentration was checked on NanoDrop Lite Spectrophotometer again and the quality of RNA was checked by gel electrophoresis (0.8 % (w/v) agarose gel stained with ethidium bromide) using the Gene Ruler 1 kb DNA Ladder (Fermentas) as a standard. For DNA extraction, approximately 200 ml of media was used. It was centrifuged at 1,500 g for 25 min, the supernatant was discarded and the pellet used for DNA extraction by modified protocol originally used for plants (Dellaporta et al. 1983). Modifications included additional steps of RNAse (RNAse H; the final concentration 0.1 mg/ml) and proteinase treatments (proteinase K; the final concentration 0.2 mg/ml) followed by phenol-chloroform extraction before the final DNA precipitation. The quality and concentration of the isolated DNA quality were checked by gel electrophoresis as for RNA (see above).

2.3. Transcriptome and genome sequencing and assembly

The transcriptome was sequenced using Illumina HiSeq 2000 2x150bp platform from cDNA libraries prepared using TruSeq RNA sample prep kit v2 (Illumina, San Diego, CA) at Macrogen Inc. (Seoul, South Korea). The raw reads were error-corrected by Rcorrector (Song and Florea 2015), trimmed with Trimmomatic v. 0.32, and assembled into contigs by Trinity assembler v. 2.8.4. (Haas et al. 2013). Protein sequences were then predicted with TransDecoder (https://github.com/TransDecoder/). DNA library was prepared using TruSeq DNA PCR-Free Protocol (Illumina, San Diego, CA) at Macrogen Inc. (Seoul, South Korea), where it was also sequenced on Illumina HiSeq 2000 2x150bp platform. The total amount of 57,000,000 raw reads were trimmed by Trimmomatic v. 0.32. (Bolger et al. 2014). An initial (meta)genome assembly was built with Spades 3.13.0 (Bankevich et al. 2012) and filtered from bacterial contamination as follows: all contigs > 100,000 bp were investigated by BLAST (Altschul et al. 1997) to determine bacterial contamination; using Burrows-Wheeler Aligner (BWA; Li and Durbin 2010) and SAMtools (Li et al. 2009), reads aligned with identified bacterial contigs were removed, surviving reads were reassembled and decontamination procedure was repeated. The final set of surviving reads (7,400,000 reads) was assembled by Spades 3.13.0 using user defined k-mer sizes (-k 21,33,55,77) and „--careful“ option to minimize number of mismatches in the final contig. Such a way, we obtained relatively clean genome assembly containing 11,440 contigs > 1000 bp, three of them identified as fragments of a ptDNA based on TBLASTN searches with common plastid proteins. Subsequent iterative searches of the raw Illumina reads enabled us to unambiguously extend the ptDNA fragments to obtain a full circular-mapping genome sequence including two copies of the inverted repeat typical for plastid genomes. The final assembly was verified by visual inspection (using Tablet; Milne et al. 2016) of sequencing reads mapped onto the assembled plastid genome sequence with BWA.

2.4. Plastid genome annotation

The assembled plastid genome of *Leukarachnion* sp. PRA-24 was initially annotated by using MFannot (https://megasun.bch.umontreal.ca/cgi-bin/mfannot/mfannotInterface.pl) and the annotation provided by the program was carefully manually checked to ensure that the genes and coding sequences are correctly delimited and that no gene present in the genome is missed. The identity of tRNA genes was additionally checked using tRNAscan-SE v. 2.0 (Chan and Lowe 2019). No candidate gene for 5S rRNA was recognised by MFannot, so we employed Infernal cmscan (https://www.ebi.ac.uk/Tools/rna/infernal_cmscan/), which did not suggest any candidate either. Further attempts to identify a possible 5S rRNA gene was based on searching for a sequence segment in the unannotated region downstream of the 23S rRNA gene that would fold into an RNA secondary structure characteristic for 5S rRNA. The RNAfold WebServer (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi) was fed with various subsections of the aforementioned region and the predicted structures were inspected to detect possible structural resemblance to 5S rRNA, including its circularly permuted variants (Valach et al. 2014). Possible homology of the several unidentified ORFs to proteins from other eukaryotes were sought using BLAST, HHpred (Zimmermann et al. 2018), and Phyre2 (Kelley et al. 2015). The presence of transmembrane domains in plastid-encoded proteins was analysed using TMHMM server v. 2.0 (http://www.cbs.dtu.dk/services/TMHMM/). The plastid genome map was obtained using OGDRAW v.1.3.1 (https://chlorobox.mpimp-golm.mpg.de/OGDraw.html; Greiner et al. 2019).

2.5. Analysis of nucleus-encoded plastid-targeted proteins

Candidates for plastid proteins encoded by the *Leukarachnion* sp. PRA-24 nuclear genome were searched by using two different homology tools. TBLASTN was employed to search the transcriptome assembly for homologs of well-characterised reference plastid proteins gathered from databases (especially KEGG; https://www.genome.jp/kegg/pathway.html) or the literature. HMMER 3.0 (Eddy 2011) was employed as a more sensitive tool to search against the database of *Leukarachnion* sp. PRA-24 protein sequences inferred from the transcriptome assembly. HMM models were built from seed alignments of appropriate protein families or domains downloaded from the database Pfam (http://pfam.xfam.org/; Mistry et al. 2021) or from custom protein sequence alignments (obtained by using MAFFT version 7; https://mafft.cbrc.jp/alignment/server/; Katoh and Standley 2013) and used as queries for the HMMER searches. Hits with significantly low e-values (typically <0.001) were evaluated by BLASTX or BLASTP searches against the NCBI non-redundant protein sequence database to remove false hits, bacterial contaminants (most frequently encountered were contaminants from a *Vibrio* sp.), or sequences representing a different member (paralog or pseudoparalog) of the respective protein family than the protein searched for. The presence of an N-terminal signal peptide, a prerequisite for the protein to be potentially plastid targeted (see section 3.4.) in the candidates was evaluated by using TargetP 2.0 ("non-plant" setting; https://services.healthtech.dtu.dk/service.php?TargetP-2.0; Almagro Armenteros et al. 2019), TargetP 1.1 (https://services.healthtech.dtu.dk/service.php?TargetP-1.1; Emanuelsson et al. 2000), PredSL (http://aias.biol.uoa.gr/PredSL/input.html; Petsalaki et al. 2006), and PrediSi (http://www.predisi.de/; Hiller et al. 2004). The former three tools provided also prediction of possible mitochondrial targeting of the proteins, which was additionally tested using MitoFates (set to "fungi"; http://mitf.cbrc.jp/MitoFates/cgi-bin/top.cgi; Fukasawa et al. 2015).

2.6. Phylogenetic analyses

Thirty plastid-genome encoded proteins present in at least one species of the pair of non-photosynthetic ochrophytes "*Spumella*" sp. NIES-1846 (Dorrell et al. 2019) and *Leukarachnion* sp. PRA-24 and suitable for phylogenetic analysis (marked in supplementary electronic material, table S1) were selected to build a supermatrix for phylogenomic inference. For each protein, orthologs were collected from a set of reference ochrophyte plastid genomes (including representatives of all classes for which plastid genome sequence is available, and including also various non-photosynthetic ochrophytes for comparison) as well as plastid genomes of selected non-ochrophytes (to provide an outgroup); the list of the genomes analysed and the sequence identifiers are provided in supplementary electronic material, table S2. The orthologous protein sequences (in some cases modified compared to the available annotation to fix obvious mistakes in the definition of the coding sequence start or to account for apparent plastid genome sequence assembly errors) were aligned using MAFFT version 7

with the E-INS-i option. The alignments were checked visually, trimmed with trimAl (Capella-Gutierrez et al. 2009), and concatenated with FASconCAT (Kück and Meusemann 2010) to yield a supermatrix of 6,627 aligned amino acid positions. The maximum likelihood (ML) tree was inferred using IQ-TREE multicore version 2.0.3 (Nguyen et al. 2015) with the substitution model LG+C60+F+G and 100 nonparametric bootstrap replications. A dataset of the plastid 16S rRNA gene sequences was aligned and trimmed in the same manner as described above, followed by a separate ML analysis employing the GTR+F+R4 substitution model and 100 nonparametric bootstrap replications. All phylogenetic trees were visualised using FigTree 1.4.4 (Rambaut 2009) and post-processed in a graphical editor.

3. Results and Discussion

3.1. Reinvestigation of the life cycle and morphology of *Leukarachnion* sp. PRA-24 leads to its description as *Leukarachnion salinum*, sp. nov.

The life cycle of the strain studied alternated among vegetative amoeboid cells with different pseudopodial forms, flagellates, and cysts (figure 1). The uninucleate amoebae formed lobopodia (figure 1A,B). Meanwhile, the multinuclear plasmodium formed either filopodia (figure 1C–E) or reticulopodia, resulting in a meroplasmodium. The latter stage was formed only for a very short period of time at the beginning of every sub-culturing. The networks were small and highly fractured, encompassing only several cell bodies of different sizes. The previous study, on the other hand, reported a large network for the same strain (Grant et al. 2009). Our culturing conditions are likely responsible for the lack of significant anastomosing. Notably, several days after re-inoculation, when bacterial densities increased, cells always became less active. The filopodia-forming plasmodium was very dynamic in terms of its shape and size, with the main contracted body measuring up to 15 µm in length. The smallest measured cell body was 1.5 µm in length. The filopodia were of different thickness with the finest threads usually stretching out posteriorly and laterally from the bulging ones directly connected to the main body. In our cultures, one to six, mostly two or three, filopodia radiated from the main body (figure 1C–E). The active movement of the main body and pseudopodia was observed, including building, stretching and pulling in the filopodia and reticulopodia, resulting in an actively migrating plasmodium. Little bulbous protrusions could be seen to move inside the threads towards the main body or opposite.

Although the amoeboid form is most characteristic for *Leukarachnion*, our cultures were mostly dominated by resting stages, i.e. irregularly roundish cells with a slight movement or none at all. Apart from nuclei with prominent nucleoli, they also contained numerous randomly distributed vacuoles (figure 1F–H). We have also observed cells containing a single huge vacuole (figure 1I). The same type of cell was also found, for example, in another amoeboid alga, *Synchroma grande* (Synchomophyceae, Ochrophyta), and considered an intermediate stage between migrating and sessile amoebae (Koch et al. 2010). During the several weeks of cultivation, almost all *Leukarachnion* cells transformed into cysts. They could be distinguished from the vegetative cells by having a regular spherical shape and being covered by a thick double-layered cell wall (figure 1J,K). Their size varied from 2.5–7.0 µm in diameter (n=90). This size range generally matched the one of vegetative non-branching cells, suggesting that every cell is capable of encystation when growth conditions are no longer favourable. The cell wall of the cysts was smooth and did not contain any obvious protrusions or germination pores (figure 1J,K). Sometimes, the cell content inside the cyst was considerably shrunken but appeared to be active (figure 1K). Such coated cells were possibly undergoing mitotic division.

In addition to active amoeboid cells, pre-cysts and cysts, we have also observed two types of flagellated cells (figure 1L,M). Geitler (1942) studied *Leukarachnion* from field samples and did not mention any flagellates. Grant et al. (2009), on the other hand, noted the presence of flagellates in their cultures of *Leukarachnion* sp. PRA-24 but they did not provide any details. The first type of observed swimming cells had two flagella of unequal length. They were inserted subapically (figure 1L). The shape of biflagellate cells ranged from oval to bean-shaped and they were approximately 3.0 µm long and 2.5 µm in wide. They moved fast and rotated while swimming. Such cells likely represent zoospores. The second type of motile cells included uniflagellate cells with a flagellum being inserted

apically (figure 1M). Their size reached up to 3.0 in length and 2.0 in width. They could possibly represent gametes. This hypothesis is also supported by the fact that we did observe a peculiar interaction of the two cells with one of them exhibiting an active single-flagellum movement and dragging the second one while it slowly crawls towards the swimming one (figure 1N). However, the elucidation of a full life cycle of *Leukarachnion* sp. PRA-24, including the alternation of different morphotypes, awaits further detailed investigations.

Leukarachnion batrachospermi was described as having a very specialized lifestyle, including occurrence, behaviour and diet. It was found to parasitize two species of the freshwater *Batrachospermum* (Rhodophyta) by seeking out their spermatangia and carposporangia, spinning around them, and removing reproductive cells which are digested in large digestive vacuoles. Geitler (1942) noted that apart from the reproductive *Batrachospermum* cells, the plasmodium generally did not contain any other food bodies, and only in very few cases small bacterial rods were found in digestive vacuoles. This led the author to conclude that *L. batrachospermi* is a parasite specializing in *Batrachospermum* and its occurrence is also limited to the respective macroalga. *Leukarachnion* sp. PRA-24 was isolated from a salt marsh in Virginia, an ecologically different habitat. Behaviourally, the strain PRA-24 matches *Leukarachnion* description only in terms of anastomosing and often and abundant formation of cysts. Morphologically, the depicted and described cysts of *L. batrachospermi* differ from the cysts of PRA-24 by containing stubby or spiky protrusions outside the cyst wall, while cysts of PRA-24 do not form such structures. Moreover, the cyst size measured for PRA-24 is twice as small of *L. batrachospermi* (10–14 µm). On the contrary, Grant et al. (2009) found the cysts of the PRA-24 to be up to 15 µm (likely being pre-cysts), but getting smaller and reaching a uniform size of 5 µm. Finally, no prominent large digestive vacuoles were present in the plasmodia of the PRA-24 strain. Considering the fundamental differences of *Leukarachnion* sp. PRA-24 from *L. batrachospermi* in terms of being found in different habitats and living different life styles, having smaller size and cyst appearance, we propose to recognise the strain PRA-24 as a new species of the genus *Leukarachnion*, *L. salinum* (for the formal diagnosis see the Taxonomic appendix below).

3.2. *Leukarachnion salinum* has a plastid genome

As already mentioned, previous investigations by TEM did not reveal any candidate structures that might be identified as a plastid in *L. salinum* (Grant et al. 2009). To address the question of the plastid status in *L. salinum* by an alternative approach, we used the Illumina HiSeq sequencing technology to perform a survey sequencing of a DNA sample isolated from the *L. salinum* culture. Strikingly, searching the assembled sequence contigs with reference plastid genes retrieved three contigs that resembled parts of a plastid genome lacking photosynthesis-related genes. The three contigs could be manually joint to obtain a single circular-mapping sequence of 42,214 bp that exhibits the conventional plastid genome architecture with two inverted repeats separated by the long and short single-copy regions (figure 2A). We are certain that the assembled sequence does not correspond to the mitochondrial genome, as another contig was retrieved from the same assembly that represents a typical mitochondrial genome and whose assignment to *L. salinum* has been confirmed by detailed analyses (to be published elsewhere as port of a comparative study of ochrophyte mitogenomes). Hence, we refer to the assembled ~42 kbp sequence as to the *L. salinum* plastome, which is further corroborated by phylogenetic analyses presented below. Disregarding the duplications due to the presence of the inverted repeats, we identified in the plastome 38 protein-coding genes (including three unidentified ORFs, see below), 26 tRNA genes, and two rRNA genes (16S and 23S). Only the 23S rRNA gene is present in the inverted repeat in the *L. salinum* plastid genome, whereas the 16S rRNA gene is located in a single copy in the short single-copy region (figure 2A).

A comparison of the repertoire of biological functions mediated by protein-coding genes in the plastomes of *L. salinum* and selected other non-photosynthetic eukaryotes is shown in figure 2B, a comparison of the occurrence of protein-coding plastid genes in *L. salinum* and a smaller subset of non-photosynthetic taxa is provided in figure 2C, and a detailed look at the gene content of plastid genomes of *L. salinum* and diverse ochrophytes is provided in supplementary electronic material, table S1. The *L. salinum* plastid genome lacks all genes encoding components of photosystems and the electron-transport chain, the respective biogenesis factors (*ccs1*, *ccsA*), and chlorophyll-synthesis

enzymes, indicating complete loss of photosynthetic functions. In contrast to some other non-photosynthetic plastids (Hadariová et al. 2018; Bringloe et al. 2021; Pánek et al. 2022), it also lacks genes for the F_oF_1 ATP synthase, RuBisCO subunits, and proteins involved in *rbcL* and *rbcS* expression and regulation (*ccbX*, *rbcR*). Lost are also all other plastid genes that in other ochrophytes encode enzymes of plastid-localized biosynthetic pathways, including synthesis of fatty acids (*acpP*), branch amino acids (*ilvB*, *ilvH*), and thiamin (*thiG*, *thiS*). Further missing are genes for subunits of the membrane protein-targeting system SEC (including the non-coding RNA that makes part of the plastidial signal recognition particle), the twin-arginine protein translocase (TAT), and the chaperones DnaK and GroEL. An interesting feature of the *L. salinum* plastid genome is the absence of the genes *sufB* and *sufC* encoding the critical components of the SUF system of the iron-sulfur (Fe-S) cluster assembly. Both *sufB* and *sufC* are retained by plastid genomes of virtually all ochrophytes, including the non-photosynthetic ones (figure 2C; Kamikawa et al. 2018; Dorrell et al. 2019; Bringloe et al. 2021). Both are also kept in the plastid genome of parasitic colourless red algae (Salomaki et al. 2015; Preuss et al. 2020) and non-photosynthetic cryptophytes (Tanifuji et al. 2020), and *sufB* is present in plastid genomes of apicomplexans (except for *Plasmodium*) and some other non-photosynthetic myzozoans (Mathur et al. 2019). The only other exception among ochrophytes, reported recently, is the dictyochophyte genus *Pteridomonas* (Kayama et al. 2020) lacking both *sufB* and *sufC* like *L. salinum* (figures 2B and 2C).

Thus, the *L. salinum* plastid genes retained seem to belong to only three functional categories: transcription (RNA polymerase subunits), translation (ribosomal proteins, the translation elongation factor Tu, rRNAs and tRNAs), and protein turnover (ClpC). However, this list may be incomplete, as the function of three putative protein-coding genes could not be deduced because of the lack of discernible homology to other proteins even when using the highly sensitive tools HHpred and Phyre2, so they are annotated as ORFs. One of them, *orf293*, is localized downstream of *rpoA* that in other plastid genomes is occupied by the *rpl13* gene. No *rpl13* is recognizable elsewhere in the *L. salinum* plastid genome, so we visually expected a multiple alignment of Orf293 and ochrophyte Rpl13 proteins, but the former really does not seem to fit the pattern of Rpl13; if *orf293* is an *rpl13* ortholog, it must have diverged extremely. The other two, *orf374* and *orf241*, are present in the inverted repeat region and flank the gene *rps16*. The putative proteins encoded by these ORFs include two predicted transmembrane domains each (being thus the only transmembrane proteins encoded by the whole plastid genome). We tried to match them to plastid proteins with a similar general architecture or those encoded by genes with an analogous position in the genome, but without finding any plausible candidates for homologs; the origin of these proteins and their function thus remain completely unknown.

3.3. Transmission electron microscopy does not reveal an unambiguous plastid candidate in *Leukarachnion salinum*

The plastome sequence assembled from DNA sequence data obtained from the *L. salinum* culture strongly suggest the presence of a plastid organelle despite previous negative evidence from TEM. Therefore, we were compelled to reassesses its presence at the cytological level by our own TEM investigations. For the better preservation of membranes, we employed HPF, which indeed yielded successful results in visualizing the membranes of all the detected cell's structures (figure 3). Generally, the vegetative cells of *L. salinum* were bounded by plasma membrane of irregular outline (figure 3A). Meanwhile cysts (which dominated our prepared TEM specimens) had additional thick double-layered cell walls (figure 3B). The thickness of both layers was variable depending likely on the cyst age. The outer layer was fibrous, sometimes prominently extending the fibers forward and forming a fuzzy surface (figure 3C). The inner layer was more electron dense and sometimes of uneven thickness depending of cell membrane invaginations (figure 3B). The cells typically possessed one to three nuclei of variable position. Mitochondria were highly abundant and contained tubular cristae (figure 3D). The cytoplasm was dominated by ribosomes and cell volume was also filled by numerous smaller and bigger cytoplasmic vacuoles. Usually there were one or two big vacuoles and tens of smaller ones prominently encircled by the ribosomes. Multivesicular bodies were also detected (figure 3A) as well as lysosomes (figure 3E). All these vesicles are likely interconnected and together

form endosomal machinery. In addition, cytoplasmic microtubules were commonly observed as well (figure 3A,G).

Our TEM investigation confirmed the presence of flagellated cells with two flagella and their basal bodies detected (figure 3F–H). Though not complete, the observed structures of *L. salinum* flagellar apparatus still make it possible to consider their homology with other ochrophyte representatives. The basal bodies made an angle of about 90 degree and overlapped their proximal ends in a clockwise orientation (figure 3G). Such configuration, for example, resembles that of chrysophytes (Andersen 1990). A transition (proximal) helix with four gyres above the transition (basal) plate was detected in the anterior flagellum. In addition, a dense band linking the axonemal doublets to an unidentified structure was also present (figure 3H). In various representatives from closely related classes (including, for example, *Spumella* spp.), four gyres are quite common and the mentioned dense band connects axonemal doublets to the plasmalemma (Hibberd 1979). Even though out TEM observations on flagellar components are rather superficial, they are consistent with the phylogenetic placement of *Leukarachnion* among ochrophytes.

Of the highest interest was to identify the putative plastid in *L. salinum*, which we expected to manifest itself as a structure laying along the nucleus and sharing its outermost membrane with the nuclear envelope. Such structures, interpreted as leucoplasts, were detected in some other non-photosynthetic ochrophytes, including the dictyochophyte *Pteridomonas danica* (Sekiguchi et al. 2002) or the chrysophyte *Spumella* (Jeong et al. 2021). A notable example is also the four-membrane containing apicoplast of the apicomplexan parasite *Plasmodium* (Aconoidasida) detected in the close proximity to the nucleus despite the lack of continuity between the outermost membranes of the two organelles (Lemgruber et al. 2013). A somewhat similar structure bounded by several membranes and appearing to bridge to the nucleus by its outermost covering was also observed by us (figure 3I). However, since this structure was observed just once in our TEM investigation, we refrain to identify it as a plastid. Residual plastid could either be of a random distribution or typically extremely appressed to the nucleus and hardly noticeable, or plastid membranes could be highly fused and deceive by the appearance. For example, in the closest known relative of *Leukarachnion*, the amoeboid photosynthetic *Chlamydomyxa montana*, no ER membranes were evident under the TEM investigation (Pearlmutter and Timpano 1984), though the presence of the two membranes of the periplastidial endoplasmic reticulum was confirmed in another *Chlamydomyxa* species, *C. labyrinthuloides* (Wenderoth et al. 1999). Other close relatives of *Leukarachnion*, i.e. *Synchroma* have peculiar plastid arrangements termed “plastid complexes” (Horn et al. 2007) not found in its sister species *Guanchochroma* and *Chrysopodocystis* (Schmidt et al. 2015), suggesting that *Leukarachnion* could also potentially possess some peculiar type of residual plastids which are hard to identify due to the lack of thylakoids.

As an alternative approach towards visualising the *Leukarachnion* plastid we employed fluorescence *in situ* hybridization (FISH), testing several different probes designed to match the plastidial 16S rRNA, but we did not obtain any results that would correspond to specific staining of subcellular structures interpretable as plastids (data not shown). Hence, detection of the *L. salinum* plastid at the cytological level remains a challenge for future investigations.

3.4. Transcriptome data corroborate the presence of a cryptic, reduced plastid organelle in *Leukarachnion salinum*

The plastid genome of the *L. salinum* plastid by itself did not provide any clue as to why the plastid has been retained by the organism: all the functionally annotated genes in the genome encode products have house-keeping roles required for the functioning of the plastid as such. Therefore, we set out to investigate the repertoire of plastid proteins encoded by the *L. salinum* nuclear genome. Since the genome assembly from which we retrieved the plastid genome sequence is highly fragmented with regard to the nuclear genome, we sequenced and assembled the transcriptome of the organism, and looked for encoded proteins that bear characteristics of targeting to a plastid. As an ochrophyte, *L. salinum* is expected to harbour a plastid with four bounding membranes, the outermost of which is continuous with the ER (Dorrell and Bowler 2017), and plastid-targeted proteins are expected to bear

presequences with the characteristic bipartite structure comprised of a signal peptide followed by a transit peptide-like region (Patron and Waller 2007). The route of nucleus-encoded proteins to the plastid includes their co-translational transport into the ER (mediated by the signal peptide), subsequent translocation via the second bounding membrane by the specialized ERAD-derived machinery called SELMA, and finally transport through the inner two membranes (equivalent to bounding membranes of the primary plastid) by an apparatus homologous to the conventional plastid protein import machinery consisting of the TOC and TIC complexes (Maier et al. 2015).

To ascertain that these expectations are valid, we searched the transcriptome data for homologs of the key components of the SELMA, TOC, and TIC complexes. Indeed, we found orthologs of both derlin-related proteins constituting the core of the SELMA complex (sDer1-1 and sDer1-2), an ortholog of the pore-forming protein central to the TOC complex (in ochrophytes called Omp85), and homologs of several components of the TIC complex (Tic22 and Tic110 identified with confidence). Furthermore, *L. salinum* encodes and ortholog of the stromal processing peptidase mediating removal of the N-terminal transit peptide like region to yield mature stromal proteins (see supplementary electronic material, table S3 for details on the plastid import machinery components). These findings not only provide an independent evidence that *L. salinum* has a plastid organelle (its theoretical partial reconstruction is rendered in figure 4), but also justify the notion that plastid-targeted proteins in *L. salinum* will exhibit conventional bipartite targeting presequences (BTP) with a signal peptide at the very N-terminus, in principle allowing for their identification by *in silico* approaches. With this in mind, we searched the *L. salinum* transcriptome assembly for homologs of components of common plastid-localized functional modules and pathways and evaluated their likelihood of being targeted to the plastid by checking the nature of their N-termini. This allowed us to build a tentative functional map of the *L. salinum* plastid, schematically depicted in figure 4 and commented on below.

In keeping with the non-photosynthetic and presumably highly reduced nature of the *L. salinum* plastid, no candidates for plastid-targeted enzymes were found for most of the pathways investigated, including the Calvin cycle, the plastid-specific isopentenyl pyrophosphate (IPP) biosynthetic pathway (DOXP or MEP), and the pathways for biosynthesis of carotenoids, fatty acids, chlorophyll, plastoquinone, and phylloquinone. The *L. salinum* also apparently lacks ferredoxin-$NADP^+$ reductase and ferredoxin, the latter being a functionally versatile protein with a redox-active iron-sulfur (Fe-S) cluster and involved in multiple plastid-localized processes beyond photosynthesis itself (Hanke and Mulo 2013). This is even more notable considering the fact that a ferredoxin-encoding gene (*petF*) has been retained by plastid genomes of several independently evolved non-photosynthetic chrysophytes (Dorrell et al. 2019; Kim et al. 2020). Related to this is the absence in *L. salinum* of all components of the SUF system, a plastid-specific machinery for the assembly of Fe-S clusters. This is in a striking contrast with the presence of the SUF system in virtually all other plastids investigated, including those lacking any *suf* gene in their plastid genome (see also figure 2) and having the SUF system encoded fully by the nuclear genome. The latter concerns not only photosynthetic taxa (Przybyla-Toscano et al. 2021), but also eukaryotes bearing non-photosynthetic plastids, such as the parasitic trebouxiophyte *Helicosporidium* sp. (Pombert et al. 2014), various non-photosynthetic myzozoans (colpodellids, *Perkinsus marinus*; Janouškovec et al. 2015), or *Euglena longa* (Novák Vanclová et al. 2020). Rare counterexamples exhibiting plastids devoid of the SUF pathway include piroplasmids (Apicomplexa; Sato 2011) and *Pteridomonas* spp. (Kayama et al. 2020).

Plastids membranes generally contain four types of characteristic structural lipids, the phospholipid phosphatidylglycerol (PG), the galactolipids monogalactosyldiacylglycerol (MGDG) and digalactosyldiacylglycerol (DGDG), and the sulfolipid sulfoquinovosyldiacylglycerol (SQDG; Yoshihara and Kobayashi 2022), so it was natural to check *L. salinum* for the presence of enzymes mediating their synthesis. In plastid-bearing eukaryotes PG is synthesized by a plastid-localised pathway that is, however, typically parallel to a separate pathway producing this phospholipid for the use in other cellular compartments (Michaud et al. 2017). While *L. salinum* does have homologs of the enzymes dedicated to the PG biosynthesis, none of them has a putative plastid-targeting presequence (supplementary electronic material, table S3). Completely absent from *L. salinum* are enzymes for the synthesis of MGDG and DGDG. The presence of galactolipids in non-photosynthetic plastids varies;

some have retained them (e.g. that of *E. longa*; Füssy et al. 2020), others function without them, such as the apicoplast of apicomplexan parasites (Botté and Maréchal 2014). In contrast, both enzymes required for the synthesis of SQDG, i.e. UDP-sulfoquinovose synthase (SQD1) and SQDG synthase (SQD2), are present in *L. salinum*, both bearing an N-terminal extension matching the structure of a BTP consistent with their presumed plastid localisation. The *L. salinum* plastid membranes are thus predicted to contain SQDG, similar to the leucoplasts of *E. longa*, non-photosynthetic diatoms, and some non-photosynthetic chrysophytes (Goddard-Borger and Williams 2017; Dorrell et al. 2019; Füssy et al. 2020).

Plastid propagation depends on division of existing organelles, mediated by a machinery consisting of both cyanobacterium-derived and newly-evolved components (Miyagishima 2011; Chen et al. 2018). We searched the *L. salinum* transcriptome assembly for the key components of plastid division machinery previously reported from ochrophytes, namely the protein FtsZ (derived from the cyanobacterial plastid ancestor) that localizes in the plastid stroma and participates in the plastid envelope constriction from the inside, and the eukaryote-specific dynamin-related protein ARC5 (also called DRP5B) that associates with and constricts the plastid from the outside. Three FtsZ homologs were identified, with one showed by phylogenetic analysis to be an ortholog of the mitochondrial form of the protein and the other two representing the two plastidial FtsZ paralogs previously recognised to occur in ochrophytes and labelled FtsZA and FtsZC (supplementary electronic material, table S3; Miyagishima et al. 2004). Together with the presence of an *L. salinum* ARC5 ortholog in the transcriptome data (supplementary electronic material, table S3), the bioinformatic evidence suggests that the "standard" mechanism of plastid division is generally conserved in *L. salinum*.

3.5. Biosynthesis of haem, a folate precursor, and tocotrienols are the key physiological functions of the *Leukarachnion salinum* plastid

The putative *L. salinum* plastid-localised functions described in the previous sections are all house-keeping processes serving in the biogenesis and maintenance of the plastid itself. But what are the key physiological functions of the plastid that underpin its retention by the "host" cell? Our analyses of the transcriptome data indicate that the *L. salinum* plastid hosts of at least three metabolic pathway that clearly or potentially serve to the benefit of the whole cell (figure 4).

The first one is a pathway responsible for the synthesis of haem. All ochrophytes investigated so far possess a single haem biosynthesis pathway that is fully contained in the plastid and serves not only in the production of haem and other tetrapyrroles for the plastid itself, but supplies with haem the cell as a whole (Cihlář et al. 2019). It was, therefore, not surprising to find in the *L. salinum* transcriptome assembly homologs of all enzymes of the pathway from glutamyl-tRNA reductase (HemA) and glutamate-1-semialdehyde 2,1-aminomutase (HemL) to the final protoporphyrin ferrochelatase (HemH). All the enzymes carry a putative plastid-targeting presequence (supplementary electronic material, table S3), consistent with their expected plastid localisation. The presence of HemA and HemL implies that the key intermediate of the haem biosynthesis, 5-aminolevulinate, is created in *L. salinum* via the so called C5 pathway characteristic for plastids and using glutamyl-tRNA as the haem precursor (Cihlář et al. 2019). This means that tRNA(Gln) specified by the *L. salinum* plastid genome serves not only in translation in the plastid, but also is critical for haem production. The alternative route to 5-aminolevulinate, the so-called C4 (or Shemin) pathway based on the reaction of glycine and succinyl-CoA catalysed by the enzyme 5-aminolevulinate synthase (ALAS) localised to the mitochondrion in eukaryotes is absent from *L. salinum*, as we could not detect any ALAS homolog in its transcriptome assembly. This is consistent with the general absence of this enzyme from ochrophytes (Cihlář et al. 2019). Evidence for possible ramifications of the haem biosynthesis pathway that would localise to the *L. salinum* plastid, such as the branches leading to sirohaem or chlorophylls, or further modifications of heam towards bilins, was not identified in our data. While possible function of haem in the *L. salinum* plastid cannot be excluded at the present stage of knowledge, we infer that the primary clients of the plastid-localised haem synthesis are cytochromes of the mitochondrional respiratory chain.

Plastids commonly house the shikimate pathway supplying precursors of aromatic compounds, including aromatic amino acids and folate (Richards et al. 2006; Tzin et al. 2012). Our analysis of the *L. salinum* transcriptome assembly revealed that this organism has homologs of all seven enzymes of the pathway (supplementary electronic material, table S3). The protein corresponding to the first enzyme of the pathway, 3-deoxy-D-arabino-heptulosonate 7-phosphate synthetase catalysing condensation of phosphoenolpyruvate and D-erythrose 4-phosphate, does have an N-terminal extension compared to the targeting signal-free prokaryotic homologs, but a possible signal peptide was not recognized by any of the prediction programs employed and possible mitochondrial targeting presequence was predicted inconsistently (only by some of the programs and with weak scores only; supplementary electronic material, table S3), so we interpret the enzyme to most likely be cytosolic (figure 4). The protein representing the second enzyme of the pathway, 3-dehydroquinate synthase, has no presequence and the prediction programs do not suggest any specific targeting (supplementary electronic material, table S3), consistent with the notion that this enzyme is cytosolic in *L. salinum*. The following two enzymes of the shikimate pathway, 3-dehydroquinate dehydratase and shikimate dehydrogenase, are parts of a single fusion protein possessing an N-terminal extension compared to bacterial homologs, but it is not recognised as any characteristic targeting presequence by any of the programs used by us, again suggesting cytosolic localisation of the protein. In contrast, the homologs of the remaining three enzymes, namely shikimate kinase, 5-enolpyruvylshikimate-3-phosphate synthetase (also called 3-phosphoshikimate 1-carboxyvinyltransferase), and chorismate synthase, are all predicted to be targeted to the plastid. Thus, the organisation of the shikimate pathway in *L. salinum* (figure 4) seems to differ from that in other ochrophytes, including the non-photosynthetic members of the diatom genus *Nitzschia*, which contain the whole pathway in the plastid (Kamikawa et al. 2017). The relocation of the initial steps of the pathway from the plastid in *L. salinum* may relate to the loss of the Calvin cycle and hence a source of D-erythrose 4-phosphate in the organelle.

The final product of the shikimate pathway, chorismate, is a branching point leading to multiple aromatic compounds. One branch, starting with the enzyme chorismate mutase, leads to phenylalanine and tyrosine. However, the *L. salinum* transcriptome assembly lacks any discernible homologs of this or other enzymes of phenylalanine and tyrosine biosynthesis. Another branch leads to tryptophan, with the first two reactions being catalysed by anthranilate synthase and anthranilate phosphoribosyltransferase. We did not detect any candidates for these enzymes in the *L. salinum* transcriptome data, so we were surprised to find a protein comprising regions of homology to the two following enzymes of the tryptophan biosynthesis pathway, i.e. phosphoribosylanthranilate isomerase (TrpF) and indole-3-glycerol phosphate synthase (TrpC), fused in the reversed order from the N- to the C-terminus (supplementary electronic material, table S3). However, a closer inspection of the amino acid sequence of both the TrpF and TrpC region revealed mutations at invariant positions that were identified as critical for the catalytic activity of both enzymes: a cysteine residue, crucial for the TrpF function (Henn-Sax et al. 2002) is mutated to a serine reside in the *L. salinum* protein (supplementary electronic material, figure S1A), whereas a lysine residue essential for the catalytic activity of TrpC (Darimont et al. 1998) is mutated to a alanine residue in the *L. salinum* protein (supplementary electronic material, figure S1B). These changes suggest that the *L. salinum* TrpC-TrpF fusion protein is catalytically inactive, i.e. that it is a pseudoenzyme (Jeffery 2019) with presently undefined function in the *L. salinum* plastid. Since *L. salinum* apparently lacks candidates for the remaining enzyme of tryptophan biosynthesis, we conclude that this organism is auxotrophic for trypthophan as well as phenylalanine and tyrosine.

The only metabolic sink for chorismate produced in the *L. salinum* plastid thus seems to the biosynthetic branch leading to folate. Specifically, we identified in the *L. salinum* transcriptome assembly homologs of all enzymes of the pathway, with the one catalysing the first step, 4-amino-4-deoxychorismate synthase, having a putative plastid-targeting presequence (supplementary electronic material, table S3). Enzymes for the following steps, from 4-amino-4-deoxychorismate lyase to the bifunctional dihydrofolate synthase / folylpolyglutamate synthase, are most likely all localised outside the plastid based on the features of their N-termini. Hence, the *L. salinum* plastid seems to a site of production of 4-amino-4-deoxychorismate, which is then exported to the cytosol for further processing up to the one-carbon carrier tetrahydrofolate (figure 4).

The thirds pathway, whose presence in the *L. salinum* plastid can be reconstructed based on the transcriptomic data, may look surprising for a non-photosynthetic plastid, as it is commonly associated with photosynthesis. Namely, *L. salinum* possesses homologs of three enzymes of tocopherol and tocotrienol biosynthesis pathway, all with predicted plastidial localisation: homogentisate phytyltransferase / homogentisate geranylgeranyltransferase, MPBQ/MSBQ methyltransferase, and tocopherol cyclase (supplementary electronic material, table S3). The first enzyme catalyses the first committed step of the pathway, prenylation of homogentisate with a phytyl or geranylgeranyl residue (using phytyl-PP or geranylgeranyl-PP as the source), initiating thus synthesis of tocopherols or tocotrienols, respectively. *L. salinum* possesses both enzymes required for the production of homogentisate from L-tyrosine, predicted to be localized in the mitochondrion and the cytosol, implying homogentisate import into the plastid (figure 4; supplementary electronic material, table S3). Geranylgeranyl-PP is a ubiquitous metabolite and *L. salinum* contains a single form of the enzyme catalysing its production (geranylgeranyl diphosphate synthase, type III), which is apparently cytosolic (supplementary electronic material, table S3). Phytyl-PP is a specific derivative of geranylgeranyl-PP synthesised by the action of geranylgeranyl diphosphate reductase (GGDR), but no candidate for this enzyme is discernible in the *L. salinum* sequence data. This organism thus probably lacks the capability of synthesizing phytyl-PP, which is consistent with the lack of other pathways utilising this compound (chlorophyll and phylloquinone biosynthesis).

Based on these findings we infer that *L. salinum* produces tocotrienols but not tocopherols, specifically γ- and δ-tocotrienol, which assumes inport of geranylgeranyl-PP from the cytosol (figure 4). It should, however, be noted that *Euglena* spp., while containing phytyl chains in chlorophyll and phylloquinone, also lack a discernible GGDR homolog (Novák Vanclová et al. 2020; Füssy et al. 2020), suggesting the existence of an alternative, currently unknown form of the enzyme that may theoretically be present also in *L. salinum*. The enzyme for the additional methylation step converting γ- and δ-tocotrienol (or γ-/δ-tocopherol) to α- and β-tocotrienol (or α-/β-tocopherol), i.e. tocopherol O-methyltransferase, seems to be absent from *L. salinum* (we could not retrieve it even when using as the query the tocopherol O-methyltransferase sequence from the closely related photosynthetic species *Synchroma pusillum*). The physiological role of γ- and δ-tocotrienols in a non-photosynthetic plastid is unclear, but they may be used in other parts of the cell – especially mitochondria and peroxisomes as general lipophilic antioxidants protecting membrane lipids against reactive oxygen species. Interestingly, a recent study of the non-photosynthetic euglenophyte *Euglena longa* provided bioinformatic as well as direct biochemical evidence for the production of tocopherols in this organism (Füssy et al. 2020), so the situation we have no encountered in *L. salinum* is not unprecedented.

3.6. Plastid genes in *Leukarachnion salinum* are extremely divergent

Plastid genome sequences have proven instrumental for reconstructing the phylogenetic relationships within ochrophytes (Barcytė et al. 2021; Di Franco et al. 2022). Hence, the availability of the *L. salinum* plastid genome sequence prompted us to employ it for casting more light on the phylogenetic position of this organisms in ochrophytes. To this end we constructed a supermatrix of 30 plastid genome-encoded proteins, sampling selected ochrophytes (to have represented all classes with plastid genome data available) and several non-ochrophyte algae as an outgroups. Analyses of the data (6,627 aligned amino acid positions) using the maximum likelihood method and the complex model LG+C60+F+G yielded a tree topology (figure 5) that is generally congruent with recent phylogenetic analyses based on multiple plastid genome-encoded proteins (Ševčíková et al. 2019; Han et al. 2019; Kim et al. 2020; Barcytė et al. 2021; Di Franco et al. 2022). However, thanks to new plastid genome data that became available only recently, our tree is more comprehensive in terms of the main lineages of ochrophytes represented, and exhibits several notable features. First, it for the first time includes *Schizocladia ischiensis*, the sole known representatives of the ochrophyte class Schizocladiophyceae, and confirms it as a sister lineage of brown algae (Phaeophyceae). Second, the two representatives of the class Pinguiophyceae constitute a clade that is with high bootstrap support (87%) sister to *Olisthodiscus luteus*, a representative of the recently erected class Olisthodiscophyceae that was hypothesised to be a potential sister lineage of pinguiophytes based on a much less well resolved analysis of a smaller sequence dataset (Barcytė et al. 2021). Finally, and in direct relevance of the

main theme of this study, the tree has the *L. salinum* branch placed exactly where it is expected to be based on the results of phylogenetic analyses of nuclear genes (Grant et al. 2009), i.e. in a sister position to the Chrysophyceae (figure 5).

In addition to informing about relationships of the diverse ochrophytes including *L. salinum*, the tree indicates the differences in the relative substitution rates among different plastid genomes. The long branches formed by all non-photosynthetic ochrophytes included in the tree make it immediately clear that their plastid genomes exhibit much higher substitution rates that the genomes of photosynthetic taxa. This is a common phenomenon observed in diverse non-photosynthetic taxa outside ochrophytes at the level of whole plastid genomes as well as individual genes (Záhonová et al. 2016). The branch lengths of the different colourless taxa differ substantially, and it is to be understood that this reflect not only how much the evolutionary rate increased upon the loss of photosynthesis in the particular lineage, but also how recent the switch is. Given the present knowledge, it is likely that the loss of photosynthesis in the colourless diatoms of the genus *Nitzschia* or in the unidentified non-photosynthetic phaeophyceaen encountered as an endophyte of a photosynthetic brown alga (Bringloe et al. 2021) occurred more recently than the analogous transition in the *Leukarachnion* lineage, so it not surprising that the *L. salinum* branch is much longer that the one of the non-photosynthetic diatoms or the brown alga.

Although we cannot precisely define the relative contribution of different evolutionary rates and different ages of the separate non-photosynthetic lineages to the resulting branch length, it is noticeable that the plastid genome of *L. salinum* is by far most divergent among all ochrophyte plastid genomes sequenced so far. The divergent nature of the plastid protein sequences of the non-photosynthetic taxa poses challenges to the phylogenetic inference and may be responsible for some problematic aspects of the tree obtained. Specifically, it may possibly account for the incomplete support for the monophyly of the class Chrysophyceae (85%) and of the sisterhood of chrysophytes and *L. salinum* (94%), and also for the apparently incorrect placement of the *Pteridomonas* lineage, which is classified in the order Pedinellales and is expected to branch with *Pseudopedinella elastica* based on 18S rRNA phylogenies (Kayama et al. 2020), yet in our tree makes a sister lineage to all other dictyochophytes (albeit only with moderate support). Future analyses with an improved taxon sampling and employment of ever more complex substitution models that can better account with the heterogeneity of the evolutionary process across lineages (see, e.g., Muñoz-Gómez et al. 2022), will provide more accurate results. With specific regard of *L. salinum* it is particularly pertinent to include plastid genome data from its photosynthetic relatives, such as *Chlamydomyxa* and *Synchroma* (plastid genome sequences are not yet available from them and only photosynthesis-related protein-coding genes could be extracted from a previously published transcriptome assembly from *Synchroma pusillum*, precluding representation of this species in the analysis presented here).

Theoretically, the extreme substitution rate estimated for plastome-encoded *L. salinum* proteins may be artefactually increased by inaccurate conceptual translation of plastid genes, i.e. by assuming a genetic code variant that is not in fact employed by the respective translation system. Indeed, alternative genetic codes have been described from plastids of certain algal and plant plastids, including the case of a sense-to-sense reassignment affecting the meaning of the AUA codon normally decoded as isoleucine but reassigned to methionine in plastids of the dinoflagellate genus *Lepidodinium* and the green algal genus *Chloroparvula* (Matsumoto et al. 2011; Turmel et al. 2019). To rule out this or other possible reassignments in *L. salinum*, we analysed its plastome with the program FACIL dedicated to predicting the codon meaning in a given genome based on comparisons of homologous sequence (Dutilh et al. 2011). The result (supplementary electronic material, figure S2) shows no signal for codon reassignment in the *L. salinum* plastid. This is in agreement with the fact that the plastome contains a standard set of tRNA genes directly comparable to tRNA gene sets in plastomes of other ochrophytes (supplementary electronic material, table S1).

To test the notion that the elevated evolutionary rate observed for the plastid-coding genes in the *L. salinum* plastome holds also for non-coding genes, we carried out a phylogenetic analysis of the 16S rRNA gene, employing a set of sequences covering the whole diversity of ochrophyte lineages and

appropriate non-ochrophyte taxa as an outgroup. In this case we could also include a sequence *S. pusillum*, as this could be extracted from the available transcriptome data. In addition, we included several 16S rRNA sequences from environmental DNA surveys identified among the best BLASTN hits in searches against the NCBI non-redundant nucleotide sequence database with the *L. salinum* and *S. pusillum* sequences as queries. The phylogenetic tree inferred from the plastidial 16S rRNA sequence comparison (supplementary electronic material, figure S3) revealed two salient things. First, consistent with the result of the analysis of plastid genome-encoded proteins, the branch representing *L. salinum* is extremely divergent, with its root-to-tip cumulative length as high or higher compared to 16S rRNA sequences of other non-photosynthetic ochrophytes included in the analysis. Second, several sequences representing environmental DNA (eDNA) clones were found to be more closely related to the *L. salinum* sequence than to any other taxon. One (GenBank accession number EF043559.1) was a partial sequence from an agricultural soil sample annotated as "uncultured soil bacterium clone 0127" and branching as the immediate sister lineage of the *L. salinum* 16S rRNA. More distant but still specifically affiliated to the *L. salinum* 16S rRNA was a group of three nearly identical sequences (GenBank accession numbers LN565456.1-LN565458.1) representing eDNA clones that appear to have been obtained from microbial communities associated with the leaf-cutter ant *Atta colombica* (although we could not find any formal publication reporting further details on the origin of the sequences) and each is assigned to an "uncultured bacterium" in the respective database record. Our analyses suggest that all these eDNA clones, instead of being derived from uncultured bacteria, come from plastid genomes of *Leukarachnion*-related organisms, presumably also non-photosynthetic (considering the divergence of the sequences similar to that of the *L. salinum* plastidial 16S rRNA gene). It is notable that the clade combining the *L. salinum* sequence and the environmental clones does not cluster together with the sequence from *S. pusillum*, which is likely an artefact resulting from the extreme divergence of the former sequences.

4. Conclusions

Before we stare the work reported in this paper, *Leukarachnion* sp. PRA-24 was a candidate for an ochrophyte that has secondarily lost the plastid organelle. While confirming this would open up an exciting and rare opportunity to study the process of complete organelle elimination from a eukaryotic cell, our results clearly demonstrate that this organism, here formally described as the new species *Leukarachnion salinum*, is more conventional in this regard and joins a growing list of secondarily non-photosynthetic eukaryote lineages that have retained a residual plastid. This conviction is based on the very strong genomic signature for the presence of a plastid organelle in *L. salinum*, including the plastid genome and numerous nuclear genes that encode hallmark plastid proteins (components of the SELMA, TOC and TIC complexes, enzymes for SQDG synthesis, plastid division proteins), while unquestionable cytological evidence for the plastid is yet to be provided. Despite being difficult to spot, the *L. salinum* plastid is less reduced than plastids of some other non-photosynthetic taxa characterized by the absence of a genome, exemplified by the chrysophyte genus *Paraphysomonas* (Dorrell et al. 2019). Our investigations indicate that the *raison d'être* for the plastid genome in *L. salinum* (i.e. disregarding its role in the genome maintenance and gene expression itself) is to provide two specific products required for the plastid functions: the (highly simplified) ClpC protein presumably involved in protein turnover in the organelle, and tRNA(Glu) serving as an input of a haem biosynthesis pathway housed by the plastid. We identified two additional processes taking place in the organelle and potentially functionally important for the cell as a whole: biosynthesis of a tetrahydrofolate precursor and of tocotrienols. It is likely that the *L. salinum* plastid takes part in additional process, which will be uncovered by future investigations.

A prerequisite for a more comprehensive understanding of the physiological roles of the *L. salinum* cryptic plastid is to overcome limitations of the approach taken in this study, i.e. solely bioinformatic reconstructions based on a transcriptome assembly (unlikely to cover all plastid-targeted proteins and frequently not allowing for correct inference of the N-terminal sequences of the proteins, interfering thus with predictions of subcellular targeting). Thus, our future aim is to obtain a high-quality assembled and annotated genome sequence of *L. salinum* and also to employ biochemical methods to experimentally verify at least some of the bioinformatic predictions, such as the production of tocotrienols by the *L. salinum* plastid. The results of our analyses of the plastidial machineries that

mediate basic functioning of the *L. salinum* plastid as a semi-autonomous organelle highlight their various peculiar aspects presumably related to relaxation of functional constraints upon the loss of photosynthesis and further underscore the extremely divergent nature of the *L. salinum* plastid. The transcriptome assembly reported in this study provide a rich resource for further addressing the phylogenetic position of *L. salinum* and consolidating its formal classification, which is presently being done as part of an ochrophyte-wide phylotranscriptomic study.

5. Taxonomic summary

Stramenopiles

Ochrophyta *incertae sedis*

***Leukarachnion* L.Geitler 1942 emend. Barcytė & M.Eliáš**

Emended diagnosis: amoeboid cells bearing pseudopodia of different types, including anastomosing; uni- or multinucleate. Highly vacuolated. Flagellate stages present. Cysts formed. Living in freshwater and saltwater; parasitic and free-living.

Type species: *Leukarachnion batrachospermi* L. Geitler 1942, Biol Zentralb 62: 548.

Lectotype designated here: drawing d of a single cell in figure 5 in Geitler (1942).

Leukarachnion salinum Barcytė & M.Eliáš sp. nov. (figure 1A–N)

Description and diagnosis: Cells naked, amoeboid with lobopodia and filopodia. One to six reticulopodia radiating from a single cell. Uninucleate or multinucleate. Cell bodies roundish, elongated or spindle-shaped, stretching 1.5–15 µm in length. Flagellated cells of two types: biflagellate with flagella of unequal length and inserted subapically, cells ovoid to bean-shaped, 3.0 µm long and 2.5 µm wide; and uniflagellate with the single flagellum inserted apically, cells ovoid, 3.0 µm long and 2.0 µm wide. Multiple vacuoles present. Mitochondria with tubular cristae. Cysts spherical and double-walled, without protrusions, ranging from 2.5–7.0 in diameter. Morphologically differs from *L. batrachospermi* in having smaller cysts that also lack protrusions; ecologically – by living in saltwater and being exclusively bacterivorous.

Holotype designated here: Resin block of strain PRA-24 (accession number ######), deposited at the Culture Collection of Algae of Charles University in Prague, Czechia.

Type strain: ATCC® PRA-24™ in American Type Culture Collection (www.atcc.org).

Type locality: salt marsh, Virginia, USA.

Etymology: The species epithet refers to the nature of the habitat the organism was isolated from.

Data Accessibility

The assembled and annotated plastid genome sequence of *Leukarachnion salinum* was deposited at GenBank with the accession number ########. Raw RNAseq reads and transcriptome assembly from the same species are available under the NCBI BioProject ########.

Author Contributions

K.J.: data curation, formal analysis, investigation, methodology, validation, visualization, writing – review & editing; D.B.: conceptualization, data curation, investigation, validation, visualization, writing – original draft, writing – review & editing; T.P: formal analysis, investigation, resources, writing – review & editing; T.Š.: data curation, formal analysis, investigation, supervision, validation, writing – review & editing; A.E.: methodology, project administration, resources, writing – review & editing; M.E.: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, supervision, validation, writing – original draft, writing – review & editing.

Competing interests

We declare we have no competing interests.

Funding

This work was supported by the Czech Science Foundation (projects 17-21409S and 20-27648S) and the project "CePaViP", 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).

Acknowledgements
We thank Jana Salamonová (University of Ostrava) for checking the codon meaning in the *L. salinum* plastome with FACIL.

Supplementary Material
The Supplementary Material for this article can be found online at: xxxxxxxxx.

References

Almagro Armenteros JJ, Salvatore M, Emanuelsson O, Winther O, von Heijne G, Elofsson A, Nielsen H. 2019 Detecting sequence signals in targeting peptides using deep learning. *Life Sci. Alliance* **2**, e201900429. (doi:10.26508/lsa.201900429)

Altschul SF, Madden TL, Schäffer 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. (doi:10.1093/nar/25.17.3389)

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. (doi:10.1016/0005-2760(78)90054-1)

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012 SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477. (doi:10.1089/cmb.2012.0021)

Barcytė D, Eikrem W, Engesmo A, Seoane S, Wohlmann J, Horák A, Yurchenko T, Eliáš M. 2021 *Olisthodiscus* represents a new class of Ochrophyta. *J. Phycol.* **57**, 1094–1118. (doi:10.1111/jpy.13155)

Bolger AM, Lohse M, Usadel B. 2014 Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120. (doi:10.1093/bioinformatics/btu170)

Botté CY, Maréchal E. 2014 Plastids with or without galactoglycerolipids. *Trends Plant Sci.* **19**, 71–78. (doi:10.1016/j.tplants.2013.10.004)

Bringloe TT, Sauermann R, Krause-Jensen D, Olesen B, Klimova A, Klochkova TA, Verbruggen H. 2021 High-throughput sequencing of the kelp *Alaria* (Phaeophyceae) reveals epi-endobiotic associations, including a likely phaeophycean parasite. *Eur. J. Phycol.* **56**, 494 –504. (doi:10.1080/09670262.2021.1882704)

Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. 2009 trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**,1972-1973. (doi: 10.1093/bioinformatics/btp348)

Cavalier-Smith T, Scoble JM. 2013 Phylogeny of Heterokonta: *Incisomonas marina*, a uniciliate gliding opalozoan related to Solenicola (Nanomonadea), and evidence that Actinophryida evolved from raphidophytes. *Eur. J. Protistol.* **49**, 328–353. (doi:10.1016/j.ejop.2012.09.002)

Chan PP, Lowe TM. 2019 tRNAscan-SE: searching for tRNA genes in genomic sequences. *Methods Mol. Biol.* **1962**, 1–14. (doi:10.1007/978-1-4939-9173-0_1)

Chen C, MacCready JS, Ducat DC, Osteryoung KW. 2018 The molecular machinery of chloroplast division. *Plant Physiol.* **176**, 138–151. (doi:10.1104/pp.17.01272)

Cihlář J, Füssy Z, Oborník M. 2019 Evolution of tetrapyrrole pathway in eukaryotic phototrophs. *Adv. Bot. Res.* **90**, 273–309. (doi: 10.1016/bs.abr.2018.12.003)

Darimont B, Stehlin C, Szadkowski H, Kirschner K. 1998 Mutational analysis of the active site of indoleglycerol phosphate synthase from *Escherichia coli*. *Protein Sci.* **7**, 1221–32. (doi:10.1002/pro.5560070518)

Dellaporta SL, Wood J, Hicks JB. 1983 A plant DNA minipreparation: version II. *Plant Mol. Biol. Rep.* **1**, 19–21. (doi:10.1007/BF02712670)

Di Franco A, Baurain D, Glöckner G, Melkonian M, Philippe H. 2022 Lower statistical support with larger data sets: insights from the Ochrophyta radiation. *Mol. Biol. Evol.* **39**, msab300. (doi:10.1093/molbev/msab300)

Dorrell RG, Azuma T, Nomura M, Audren de Kerdrel G, Paoli L, Yang S, Bowler C, Ishii KI, Miyashita H, Gile GH, Kamikawa R. 2019 Principles of plastid reductive evolution illuminated by nonphotosynthetic chrysophytes. *Proc. Natl. Acad. Sci. U.S.A* **116**, 6914–6923. (doi:10.1073/pnas.1819976116)

Dorrell RG, Bowler C. 2017 Secondary plastids of stramenopiles. *Adv. Bot. Res.* **84**, 57–103. (doi: 10.1016/bs.abr.2017.06.003)

Dutilh BE, Jurgelenaite R, Szklarczyk R, van Hijum SA, Harhangi HR, Schmid M, de Wild B, Françoijs KJ, Stunnenberg HG, Strous M, Jetten MS, Op den Camp HJ, Huynen MA. 2011 FACIL: Fast and Accurate Genetic Code Inference and Logo. *Bioinformatics* **27**, 1929–1933. (doi:10.1093/bioinformatics/btr316)

Eddy SR. 2011 Accelerated profile HMM searches. *PLoS Comput. Biol.* **7**, e1002195. (doi:10.1371/journal.pcbi.1002195)

Emanuelsson O, Nielsen H, Brunak S, von Heijne G. 2000 Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol.* **300**, 1005–1016. (doi:10.1006/jmbi.2000.3903)

Frankovich TA, Ashworth MP, Sullivan MJ, Theriot EC, Stacy NI. 2018 Epizoic and apochlorotic *Tursiocola species* (Bacillariophyta) from the skin of florida manatees (*Trichechus manatus latirostris*). *Protist* **169**, 539–568. (doi:10.1016/j.protis.2018.04.002)

Fukasawa Y, Tsuji J, Fu SC, Tomii K, Horton P, Imai K. 2015 MitoFates: improved prediction of mitochondrial targeting sequences and their cleavage sites. *Mol. Cell. Proteomics* **14**, 1113–1126. (doi:10.1074/mcp.M114.043083)

Füssy Z, Záhonová K, Tomčala A, Krajčovič J, Yurchenko V, Oborník M, Eliáš M. 2020 The cryptic plastid of *Euglena longa* defines a new type of nonphotosynthetic plastid organelle. mSphere **5**, e00675-20. (doi:10.1128/mSphere.00675-20)

Geitler L.1942 Ein neue filarplasmodialer Organismus, *Leukarachnion batrachospermi*, und seine Lebensweise. *Biol. Zentralbl.* **62**, 541–549.

Goddard-Borger ED, Williams SJ. 2017 Sulfoquinovose in the biosphere: occurrence, metabolism and functions. *Biochem. J.* **474**, 827–849. (doi: 10.1042/BCJ20160508)

Grant J, Tekle YI, Anderson OR, Patterson DJ, Katz LA. 2009 Multigene evidence for the placement of a heterotrophic amoeboid lineage *Leukarachnion* sp. among photosynthetic stramenopiles. Protist **160**, 376–385. (doi:10.1016/j.protis.2009.01.001)

Greiner S, Lehwark P, Bock R. 2019 OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Res.* **47**, W59–W64. (doi:10.1093/nar/gkz238)

Guillou L, Chrétiennot-Dinet MJ, Boulben S, Moon-van der Staay SY, Vaulot D. 1999 *Symbiomonas scintillans* gen. et sp. nov. and *Picophagus flagellatus* gen. et sp. nov. (Heterokonta): two new heterotrophic flagellates of picoplanktonic size. *Protist* **150**, 383–398. (doi:10.1016/S1434-4610(99)70040-4)

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, Macmanes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, William T, Dewey CN, Henschel R, Leduc RD, Friedman N, Regev A. 2013 De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8**, 1494–1512. (doi:10.1038/nprot.2013.084)

Hadariová L, Vesteg M, Hampl V, Krajčovič J. 2018 Reductive evolution of chloroplasts in non-photosynthetic plants, algae and protists. *Curr. Genet.* **64**, 365–387. (doi:10.1007/s00294-017-0761-0)

Han KY, Maciszewski K, Graf L, Yang JH, Andersen RA, Karnkowska A, Yoon HS. 2019 Dictyochophyceae plastid genomes reveal unusual variability in their organization. *J. Phycol.* **55**, 1166–1180. (doi:10.1111/jpy.12904)

Hanke G, Mulo P. 2013 Plant type ferredoxins and ferredoxin-dependent metabolism. *Plant Cell Environ.* **36**, 1071–1084. (doi: 10.1111/pce.12046)

Heesch S, Peters AF, Broom JE, Hurd CL. 2008 Affiliation of the parasite *Herpodiscus durvillaeae* (Phaeophyceae) with the Sphacelariales based on DNA sequence comparisons and morphological observations. *Eur. J. Phycol.* **43**, 283–295. (doi:10.1080/09670260801911157)

Henn-Sax M, Thoma R, Schmidt S, Hennig M, Kirschner K, Sterner R. 2002 Two $(\beta\alpha)_8$-barrel enzymes of histidine and tryptophan biosynthesis have similar reaction mechanisms and common strategies for protecting their labile substrates. *Biochemistry* **41**, 12032–12042. (doi:10.1021/bi026092h)

Hiller K, Grote A, Scheer M, Münch R, Jahn D. 2004 PrediSi: prediction of signal peptides and their cleavage positions. *Nucleic Acids Res.* **32**, W375–W379. (doi:10.1093/nar/gkh378)

Horn S, Ehlers K, Fritzsch G, Gil-Rodríguez MC, Wilhelm C, Schnetter R. 2007 *Synchroma grande* spec. nov. (Synchromophyceae class. nov., Heterokontophyta): An amoeboid marine alga with unique plastid complexes. *Protist* **158**, 277–293. (doi:10.1016/j.protis.2007.02.004)

Janouškovec J, Tikhonenkov DV, Burki F, Howe AT, Kolísko M, Mylnikov AP, Keeling PJ. 2015. Factors mediating plastid dependency and the origins of parasitism in apicomplexans and their close relatives. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 10200-10207. (doi:10.1073/pnas.1423790112)

Jeffery CJ. 2019 The demise of catalysis, but new functions arise: pseudoenzymes as the phoenixes of the protein world. *Biochem. Soc. Trans.* **47**, 371–379. (doi:10.1042/BST20180473)

Jeong M, Kim JI, Nam SW, Shin W. 2021 Molecular phylogeny and taxonomy of the genus *Spumella* (Chrysophyceae) based on morphological and molecular evidence. *Front. Plant Sci.* **12**, 758067. (doi:10.3389/fpls.2021.758067)

Kamikawa R, Azuma T, Ishii KI, Matsuno Y, Miyashita H. 2018 Diversity of organellar genomes in non-photosynthetic diatoms. *Protist* **169**, 351–361. (doi:10.1016/j.protis.2018.04.009)

Kamikawa R, Moog D, Zauner S, Tanifuji G, Ishida KI, Miyashita H, Mayama S, Hashimoto T, Maier UG, Archibald JM, Inagaki Y. 2017 A non-photosynthetic diatom reveals early steps of reductive evolution in plastids. *Mol. Biol. Evol.* **34**, 2355–2366. (doi:10.1093/molbev/msx172)

Kamikawa R, Tanifuji G, Ishikawa SA, Ishii K, Matsuno Y, Onodera NT, Ishida K, Hashimoto T, Miyashita H, Mayama S, Inagaki Y. 2015 Proposal of a twin arginine translocator system-mediated constraint against loss of ATP synthase genes from nonphotosynthetic plastid genomes. *Mol. Biol. Evol.* **32**, 2598–2604. (doi:10.1093/molbev/msv134)

Katoh K, Standley DM. 2013 MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780. (doi:10.1093/molbev/mst010)

Kayama M, Maciszewski K, Yabuki A, Miyashita H, Karnkowska A, Kamikawa R. 2020 Highly reduced plastid genomes of the non-photosynthetic dictyochophyceans *Pteridomonas* spp. (Ochrophyta, SAR) are retained for tRNA-Glu-based organellar heme biosynthesis. *Front. Plant. Sci.* **11**, 602455. (doi:10.3389/fpls.2020.602455)

Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. 2015 The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858. (doi:10.1038/nprot.2015.053)

Kim JI, Jeong M, Archibald JM, Shin W. 2020 Comparative plastid genomics of non-photosynthetic chrysophytes: genome reduction and compaction. *Front. Plant Sci.* **11**, 572703. (doi:10.3389/fpls.2020.572703)

Kück P, Meusemann K. 2010 FASconCAT: convenient handling of data matrices. *Mol. Phylogenet. Evol.* **56**, 1115–11188. (doi:10.1016/j.ympev.2010.04024)

Lemgruber L, Kudryashev M, Dekiwadia C, Riglar DT, Baum J, Stahlberg H, Ralph SA, Frischknecht F. 2013 Cryo-electron tomography reveals four-membrane architecture of the *Plasmodium* apicoplast. *Malar. J.* **12**, 25. (doi:10.1186/1475-2875-12-25)

Li H, Durbin R. 2010 Fast and accurate long-read alignment with Burrows-Wheeler Transform. *Bioinformatics* **26**, 589–595. (doi:10.1093/bioinformatics/btp698)

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R,1000 Genome Project Data Processing Subgroup. 2009 The sequence alignment/map (SAM) format and SAMtools. *Bioinformatics* **25**, 2078–2079. (doi:10.1093/bioinformatics/btp352)

Maier UG, Zauner S, Hempel F. 2015 Protein import into complex plastids: cellular organization of higher complexity. *Eur. J. Cell Biol.* **94**, 340–248. (doi:10.1016/ejcbdoi.2015.05.008)

Mathur V, Kolísko M, Hehenberger E, Irwin NAT, Leander BS, Kristmundsson Á, Freeman MA, Keeling PJ. 2019 Multiple independent origins of apicomplexan-like parasites. *Curr. Biol.* **29**, 2936–2941.e5 (doi:10.1016/j.cub.2019.07.019)

Matsumoto T, Ishikawa SA, Hashimoto T, Inagaki Y. 2011 A deviant genetic code in the green alga-derived plastid in the dinoflagellate *Lepidodinium chlorophorum*. *Mol. Phylogenet. Evol.* **60**, 68–72. (doi:10.1016/j.ympev. 2011.04.010)
Michaud M, Prinz WA, Jouhet J. 2017. Glycerolipid synthesis and lipid trafficking in plant mitochondria. *FEBS J.* **284**, 376–390. (doi:10.1111/febs.13812)
Milne I, Bayer M, Stephen G, Cardle L, Marshall D. 2016 Tablet: Visualizing Next-Generation Sequence Assemblies and Mappings. *Methods Mol. Biol.* **1374**, 253–268. (doi:10.1007/978-1-4939-3167-5_14)
Miyagishima SY. 2011 Mechanism of plastid division: from a bacterium to an organelle. *Plant Physiol.* **155**, 1533–1544. (doi:10.1104/pp.110.170688)
Miyagishima SY, Nozaki H, Nishida K, Nishida K, Matsuzaki M, Kuroiwa T. 2004 Two types of FtsZ proteins in mitochondria and red-lineage chloroplasts: the duplication of FtsZ is implicated in endosymbiosis. *J. Mol. Evol.* **58**, 291–303. (doi:10.1007/s00239-003-2551-1)
Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, Tosatto SCE, Paladin L, Raj S, Richardson LJ, Finn RD, Bateman A. 2021 Pfam: The protein families database in 2021. *Nucleic Acids Res.* **49**, D412–D419. (doi:10.1093/nar/gkaa913)
Muñoz-Gómez SA, Susko E, Williamson K, Eme L, Slamovits CH, Moreira D, López-García P, Roger AJ. 2022 Site-and-branch-heterogeneous analyses of an expanded dataset favour mitochondria as sister to known Alphaproteobacteria. *Nat. Ecol. Evol.* **6**, 253–262. (doi:10.1038/s41559-021-01638-2)
Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015 IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274. (doi:10.1093/molbev/msu300)
Novák Vanclová AMG, Zoltner M, Kelly S, Soukal P, Záhonová K, Füssy Z, Ebenezer TE, Lacová Dobáková E, Eliáš M, Lukeš J, Field MC, Hampl V. 2020 Metabolic quirks and the colourful history of the *Euglena gracilis* secondary plastid. *New Phytol.* **225**, 1578–1592. (doi:10.1111/nph.16237)
Onyshchenko A, Ruck EC, Nakov T, Alverson AJ. 2019 A single loss of photosynthesis in the diatom order Bacillariales (Bacillariophyta). *Am. J. Bot.* **106**, 560–572 (doi:10.1002/ajb2.1267)
Palmer JD. 1985 Comparative organization of chloroplast genomes. *Annu. Rev. Genet.* **19**: 325-354. (doi:10.1146/annurev.ge.19.120185.001545)
Pánek T, Barcytė D, Treitli SC, Záhonová K, Sokol M, Ševčíková T, Zadrobílková E, Jaške K, Yubuki N, Čepička I, Eliáš M. 2022 A new lineage of non-photosynthetic green algae with extreme organellar genomes. *BMC Biol.* **20,** 66. (doi:10.1186/s12915-022-01263-w)
Patron NJ, Waller RF. 2007 Transit peptide diversity and divergence: a global analysis of plastid targeting signals. *Bioessays* **29**, 1048–1058. (doi:10.1002/bies.20638)
Pearlmutter NL, Timpano P. 1984 The biology of *Chlamydomyxa montana*: Ultrastructure of the cyst. *Protoplasma* **122**, 68–74. (doi: 10.1007/BF01279438)
Petsalaki EI, Bagos PG, Litou ZI, Hamodrakas SJ. 2006 PredSL: a tool for the N-terminal sequence-based prediction of protein subcellular localization. *Genomics Proteomics Bioinformatics* **4**, 48–55. (doi:10.1016/S1672-0229(06)60016-8)
Pombert JF, Blouin NA, Lane C, Boucias D, Keeling PJ. 2014. A lack of parasitic reduction in the obligate parasitic green alga *Helicosporidium*. *PLoS Genet.* **10**, e1004355. (doi:10.1371/journal.pgen.1004355)
Preuss M, Verbruggen H, Zuccarello GC. 2020 The organelle genomes in the photosynthetic red algal parasite *Pterocladiophila hemisphaerica* (Florideophyceae, Rhodophyta) have elevated substitution rates and extreme gene loss in the plastid. *J. Phycol.* **56**, 1006–1018. (doi:10.1111/jpy.12996)
Przybyla-Toscano J, Couturier J, Remacle C, Rouhier N. 2021. Occurrence, Evolution and Specificities of Iron-Sulfur Proteins and Maturation Factors in Chloroplasts from Algae. *Int. J. Mol. Sci.* **22**, 3175. (doi: 10.3390/ijms22063175)
Rambaut A. 2009 FigTree version 1.4.4 [computer program] http://tree.bio.ed.ac.uk.
Richards TA, Dacks JB, Campbell SA, Blanchard JL, Foster PG, McLeod R, Roberts CW. 2006 Evolutionary origins of the eukaryotic shikimate pathway: gene fusions, horizontal gene

transfer, and endosymbiotic replacements. *Eukaryot. Cell* **9**, 1517–1531. (doi:10.1128/EC.00106-06)
Salomaki ED, Nickles KR, Lane CE. 2015 The ghost plastid of *Choreocolax polysiphoniae*. *J. Phycol.* **51**, 217–221. (doi:10.1111/jpy.12283)
Sato S. 2011. The apicomplexan plastid and its evolution. *Cell Mol Life Sci.* **68**, 1285–1296. (doi:10.1007/s00018-011-0646-1)
Schmidt M, Horn S, Ehlers K, Wilhelm C, Schnetter R. 2015 *Guanchochroma wildpretii* gen. et spec. nov. (Ochrophyta) provides new insights into the diversification and evolution of the algal class synchromophyceae. *PLoS One* **10**, e0131821. (doi:10.1371/journal.pone.0131821)
Sekiguchi H, Moriya M, Nakayama T, Inouye I. 2002 Vestigial chloroplasts in heterotrophic stramenopiles *Pteridomonas danica* and *Ciliophrys infusionum* (Dictyochophyceae). *Protist* **153,** 157–167. (doi:10.1078/1434-4610-00094)
Ševčíková T, Yurchenko T, Fawley KP, Amaral R, Strnad H, Santos LMA, Fawley MW, Eliáš M. 2019 Plastid Genomes and Proteins Illuminate the Evolution of Eustigmatophyte Algae and Their Bacterial Endosymbionts. *Genome Biol. Evol.* **11**, 362-379. (doi:10.1093/gbe/evz004)
Song, L., Florea, L. 2015 Rcorrector: efficient and accurate error correction for Illumina RNA-seq reads. *GigaSci.* **4**, 48. (https://doi.org/10.1186/s13742-015-0089-y)
Tanifuji G, Kamikawa R, Moore CE, Mills T, Onodera NT, Kashiyama Y, Archibald JM, Inagaki Y, Hashimoto T. 2020 Comparative plastid genomics of *Cryptomonas* species reveals fine-scale genomic responses to loss of photosynthesis. *Genome Biol. Evol.* **12**, 3926–3937. (doi:10.1093/gbe/evaa001)
Turmel M, Lopes Dos Santos A, Otis C, Sergerie R, Lemieux C. 2019 Tracing the Evolution of the Plastome and Mitogenome in the Chloropicophyceae Uncovered Convergent tRNA Gene Losses and a Variant Plastid Genetic Code. *Genome Biol. Evol.* **11**, 1275–1292. (doi:10.1093/gbe/evz074)
Tzin V, Galili G, Aharoni A. 2012 Shikimate Pathway and Aromatic Amino Acid Biosynthesis. In: *eLS*. John Wiley & Sons, Ltd: Chichester. (doi:10.1002/9780470015902.a0001315.pub2)
Valach M, Burger G, Gray MW, Lang BF. 2014 Widespread occurrence of organelle genome-encoded 5S rRNAs including permuted molecules. *Nucleic Acids Res.* **42**, 13764–13777. (doi:10.1093/nar/gku1266)
Valasatava Y, Rosato A, Banci L, Andreini C. 2016 MetalPredator: a web server to predict iron-sulfur cluster binding proteomes. *Bioinformatics* **32,** 2850–2852. (doi:10.1093/bioinformatics/btw238)
Wenderoth K, Marquardt J, Fraunholz M, Van de Peer Y, Wastl J, Maier UG. 1999 The taxonomic position of *Chlamydomyxa labyrinthuloides*. *Eur. J. Phycol.* **34**, 97–108. (doi: 10.1080/09670269910001736152)
Yoshihara A, Kobayashi K. 2022 Lipids in photosynthetic protein complexes in the thylakoid membrane of plants, algae, and cyanobacteria. *J. Exp. Bot.* erac017. (doi:10.1093/jxb/erac017)
Záhonová K, Füssy Z, Oborník M, Eliáš M, Yurchenko V. 2016 RuBisCO in Non-Photosynthetic Alga Euglena longa: Divergent Features, Transcriptomic Analysis and Regulation of Complex Formation. *PLoS One* **11**, e0158790. (doi:10.1371/journal.pone.0158790)
Zimmermann L, Stephens A, Nam SZ, Rau D, Kübler J, Lozajic M, Gabler F, Söding J, Lupas AN, Alva V. 2018 A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. *J. Mol. Biol.* **430**, 2237–2243. (doi:10.1016/j.jmb.2017.12.007)

Figure legends

Figure 1. Light microscopy of different life stages of *Leukarachnion salinum* sp. nov. (*Leukarachnion* sp. PRA-24). (A–B) Uninucleate amoebae with lobopodia (lp). (C–D) Multinucleate plasmodia with different number of filopodia (fp). (F–H) Resting stages with prominent nuclei (n) and vacuoles (v). (I) Cell with a single huge vacuole. (J–K) Cysts with a doubled-layered thick covering: inner (icw) and outer (ocw) cell wall. (L) Biflagellate cell with subapical flagella (f). (M) Uniflagellate cell with an apical flagellum. (N) Interaction of two cells. Scale bars: 10 µm.

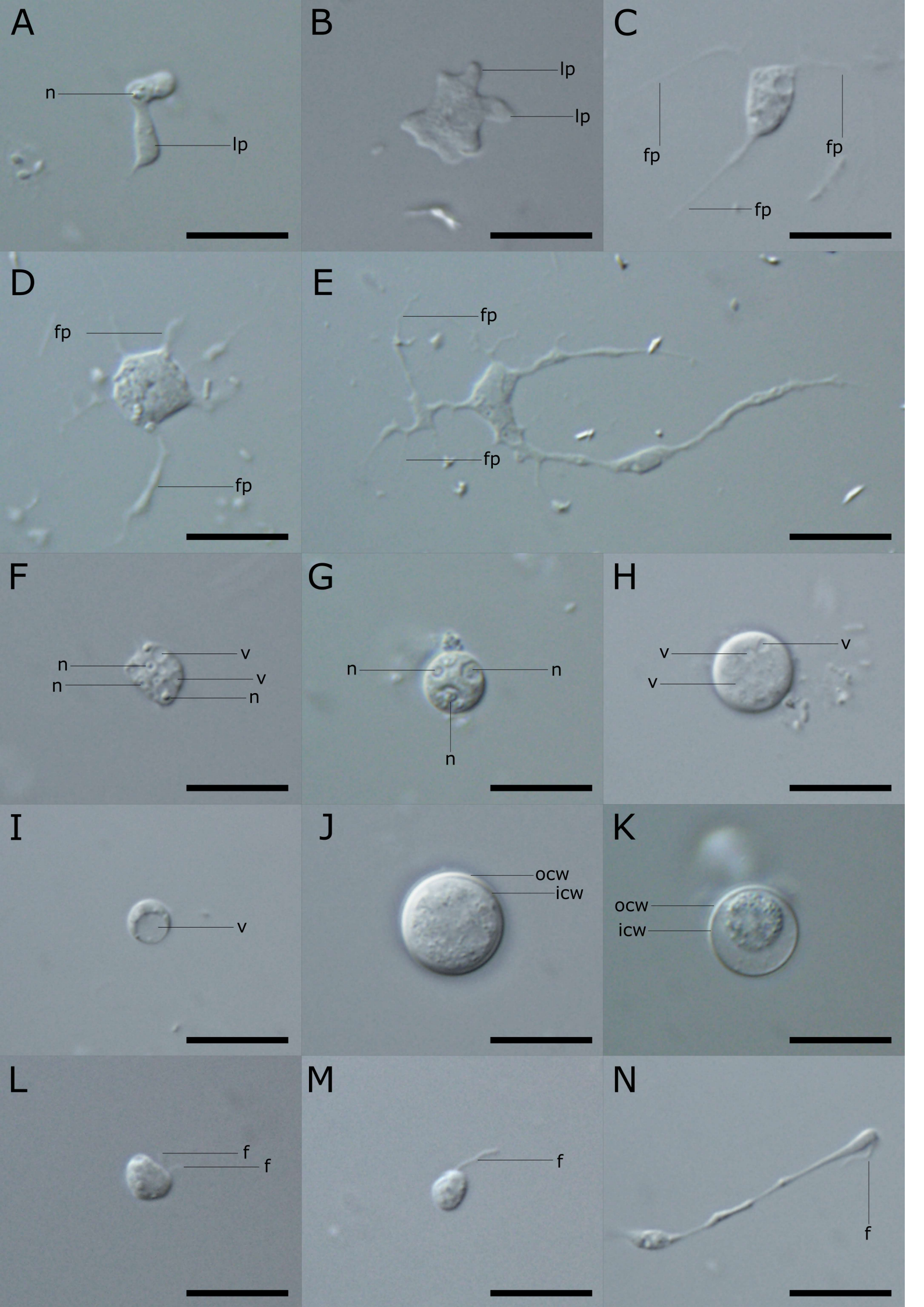
Figure 2. Plastid genome of *Leukarachnion salinum*. (A) Map of the genome. Genes are shown as boxes (coloured according to the functional category they belong to; see the graphical legend) facing inward or outward of the circle depending on the direction in which they are transcribed (clockwise or counter-clockwise, respectively; se the grey arrows). The thickenings of the outer circle delimit regions representing inverted repeats. The internal circle in grey indicates the variation in the GC content along the plastid genome. LSC – long single-copy region; SSC – short single-copy region; IRA and IRB – inverted repeats. (B) Functional categories represented in the gene sets of plastid genomes of diverse non-photosynthetic algae and plants, including *L. salinum*. (C) Occurrence of protein-coding genes in plastid genomes of non-photosynthetic ochrophytes and selected other eukaryotes with a colourless plastid of a red-algal origin. Parts B and C were adopted with modifications (addition and reordering of taxa) from Kayama et al. (2020).

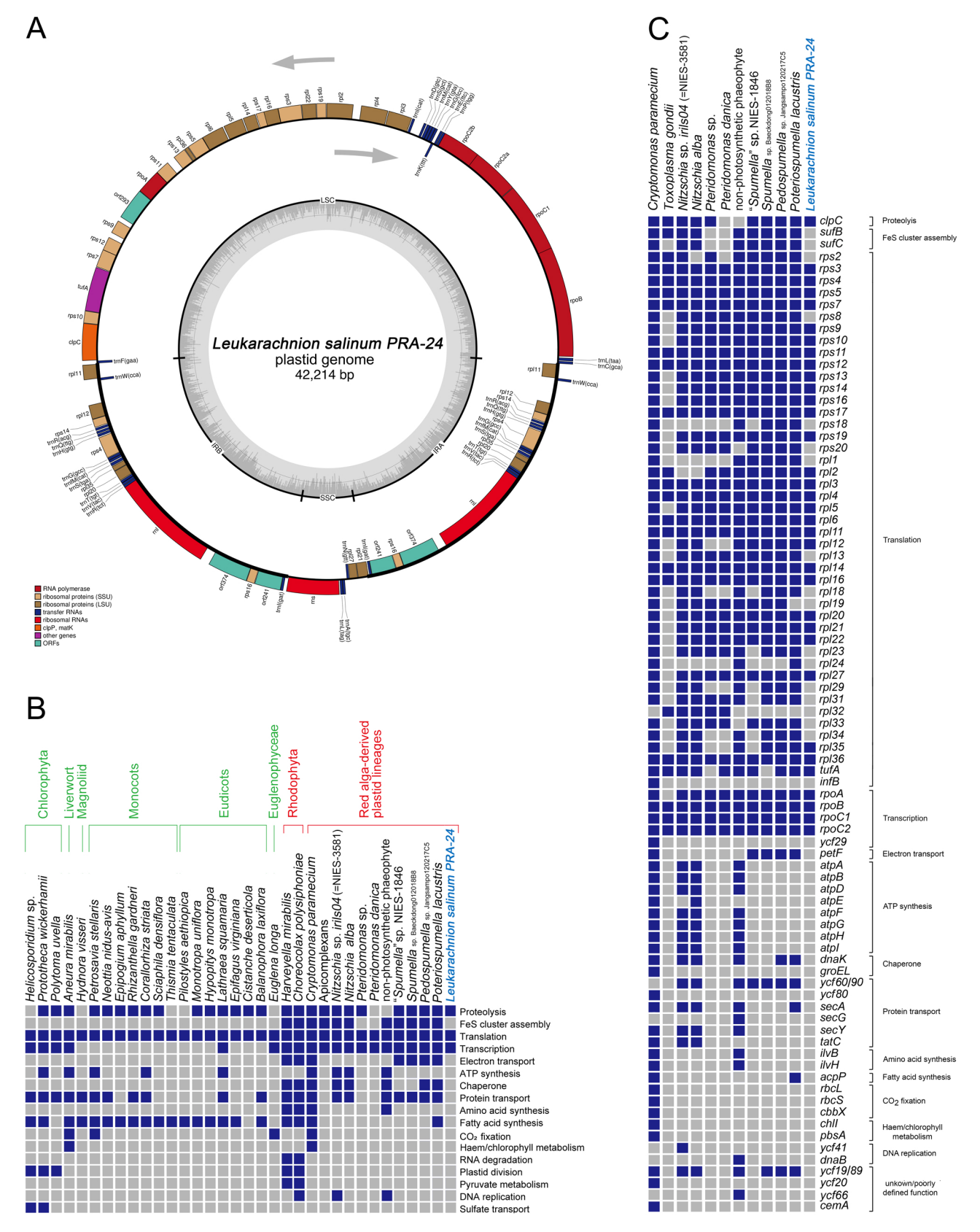
Figure 3. Ultrastructure of *Leukarachnion salinum*. (A) General organization of the cell (B) Cyst with a thick double-layered cell wall. (C) Fuzzy surface (black arrowheads) of the outer cell wall layer. (D) Mitochondrion with tubular cristae. (E) Lysosomes. (F) Biflagellate cell. (G) Basal bodies of anterior and posterior flagella. (H) Transition plate (white arrowhead) with four gyres (black arrows) of a transition helix of the anterior flagellum. White arrow points to a dense band connecting axonemal doublets to an unidentified structure. Abbreviations: abb – anterior basal body; af – anterior flagellum; bb – basal body; icw – inner cell wall; ly – lysosome; m – mitochondrion; MT – microtubules; mvb – multivesicular body; N – nucleus; ocw – outer cell wall; pbb – posterior basal body; pf – posterior flagellum; v – vacuole. Scale bars: 0.1 µm (D,H), 0.2 µm (E,I), 0.5 µm (A,C,F,G), 1 µm (B).

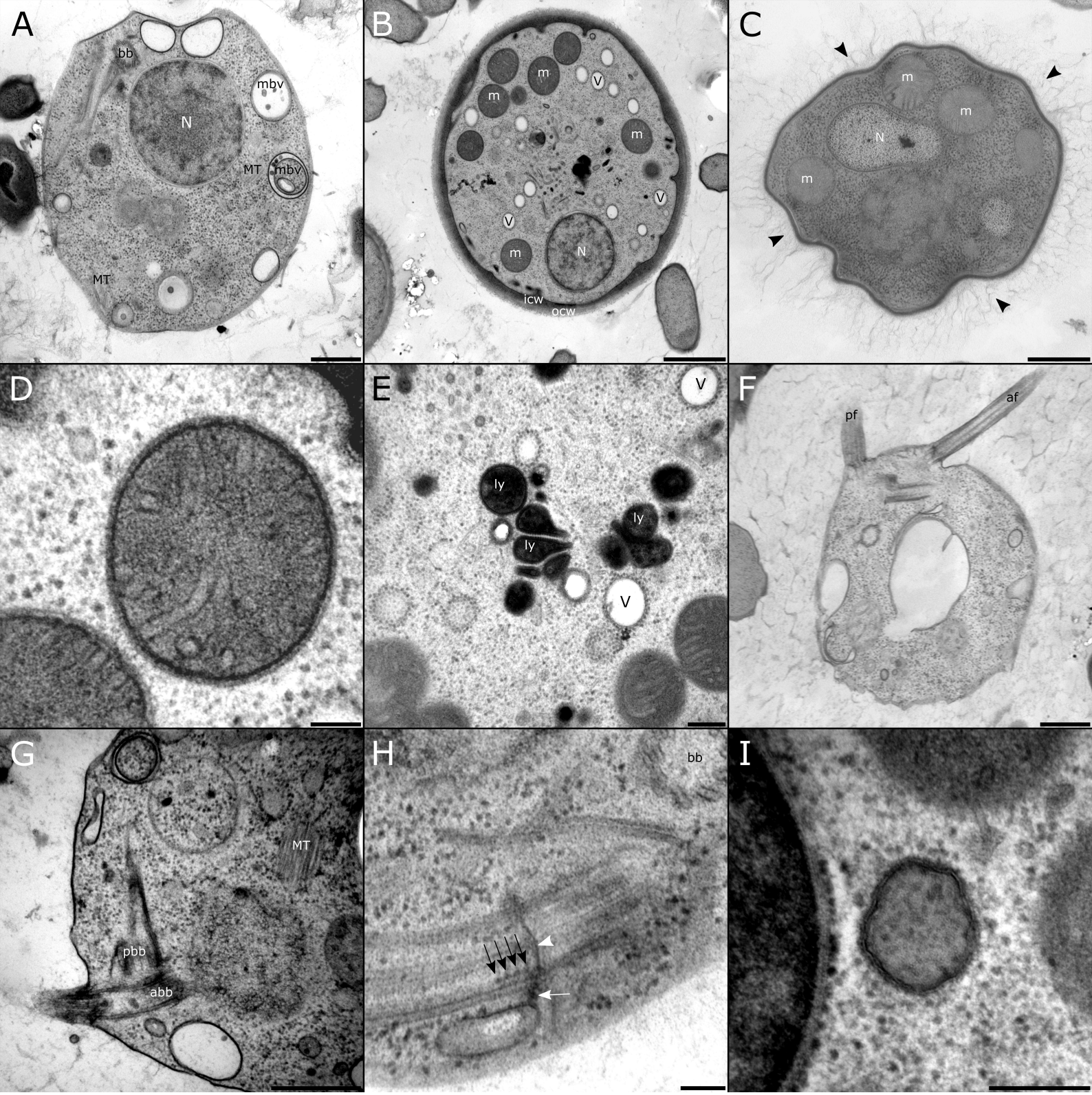
Figure 4. Bioinformatic reconstruction of the *Leukarachnion salinum* plastid and its main functional pathways. The pathways are distinguished by different colours of the intermediates and final products. Arrows with solid lines indicate a chemical transformation (reaction), with the respective enzymes catalysing the reaction indicated in ovals adjacent to the arrows (for the meaning of the abbreviations used see below). The crossed arrow in grey to L-Tryptophan highlights the apparent lack of this biosynthetic branch in *L. salinum*. Arrows with dashed lines indicate transport of the particular compound between cellular compartments. Enzyme abbreviations: Haem synthesis – HemA = glutamyl-tRNA reductase; HemL = glutamate-1-semialdehyde 2,1-aminomutase; HemB = porphobilinogen synthase; HemC = hydroxymethylbilane synthase; HemD = uroporphyrinogen-III synthase; HemE = uroporphyrinogen decarboxylase; HemF = coproporphyrinogen III oxidase; HemY = protoporphyrinogen/coproporphyrinogen III oxidase; HemH = protoporphyrin/coproporphyrin ferrochelatase. Tocotrienol synthesis – TAT = tyrosine aminotransferase; HPD = 4-hydroxyphenylpyruvate dioxygenase; VTE2 = homogentisate phytyltransferase / homogentisate geranylgeranyltransferase; VTE1 = tocopherol cyclase; VTE3 = MPBQ/MSBQ methytransferase. SQDG (sulphoquinovosyldiacylglycerol) synthesis – UGP2/USP = UTP---glucose-1-phosphate uridylyltransferase / UDP-sugar pyrophosphorylase; SQD1 = UDP-sulfoquinovose synthase; SQD2 = sulfoquinovosyltransferase. Shikimate pathway – DAHP = 3-deoxy-D-arabino-heptulosonate 7-phosphate synthetase; DHQS = 3-dehydroquinate synthase; DHQD = 3-dehydroquinate dehydratase; SDH = shikimate dehydrogenase; SK = shikimate kinase); EPSP = 5-enolpyruvylshikimate-3-phosphate synthetase; CS = chorismate synthase. Folate synthesis –ADCS = 4-amino-4-deoxychorismate synthase; ADCL = 4-amino-4-deoxychorismate lyase; DHPS = dihydropteroate synthase (the activity conferred by the protein FolKP, which may in fact localise to the mitochondrion in *L. salinum*), DHFS = dihydrofolate synthase (the activity conferred by the protein FolC1). GluRS =

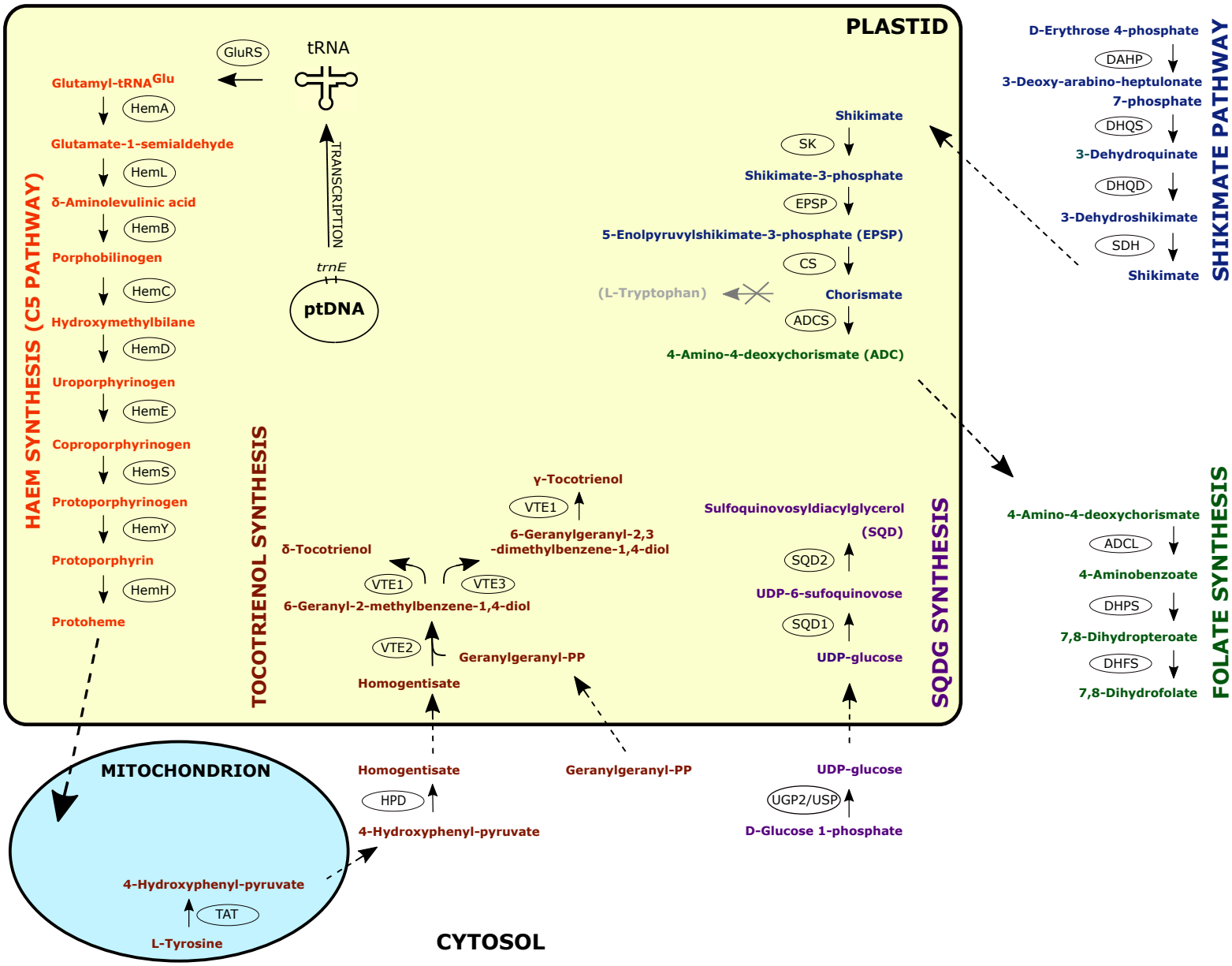
glutamyl-tRNA synthetase. Protein sequences corresponding to the enzymes and additional details are provided in see supplementary electronic material, table S3.

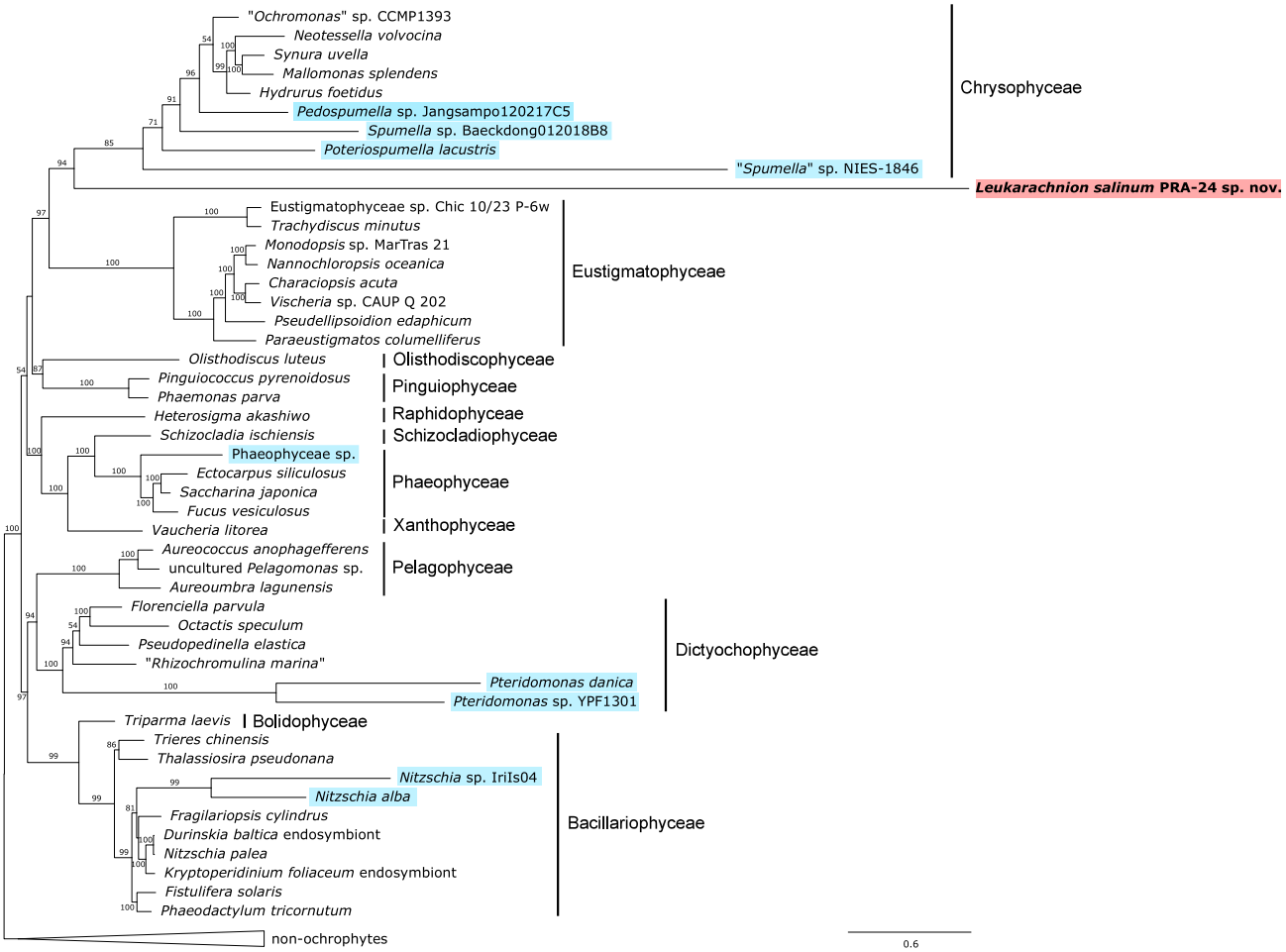
Figure 5. Phylogenetic relationships in the Ochrophyta based on plastid genome-encoded proteins. Displayed is the maximum likelihood phylogenetic tree (IQ-TREE, substitution model LG+C60+F+G) inferred from a concatenated alignment of 30 conserved proteins (6,627 aligned amino acid position). Bootstrap support values are shown when ≥50. Non-photosynthetic taxa are highlighted with a coloured background.


A
n
lp
B
lp
lp
C
fp
fp
fp
D
fp
fp
E
fp
fp
F
n
n
v
v
n
G
n
n
n
H
v
v
v
I
v
J
ocw
icw
K
ocw
icw
L
f
f
M
f
N
f


A
Leukarachnion salinum PRA-24
plastid genome
42,214 bp
LSC
SSC
IRA
IRB
RNA polymerase
ribosomal proteins (SSU)
ribosomal proteins (LSU)
transfer RNAs
ribosomal RNAs
clpP, matK
other genes
ORFs
B
Chlorophyta
Liverwort
Magnoliid
Monocots
Eudicots
Euglenophyceae
Rhodophyta
Red alga-derived plastid lineages
Helicosporidium sp.
Prototheca wickerhamii
Polytoma uvella
Aneura mirabilis
Hydnora visseri
Petrosavia stellaris
Neottia nidus-avis
Epipogium aphyllum
Rhizanthella gardneri
Corallorhiza striata
Sciaphila densiflora
Thismia tentaculata
Pilostyles aethiopica
Monotropa uniflora
Hypopitys monotropa
Lathraea squamaria
Epifagus virginiana
Cistanche deserticola
Balanophora laxiflora
Euglena longa
Harveyella mirabilis
Choreocolax polysiphoniae
Cryptomonas paramecium
Apicomplexans
Nitzschia sp. irils04 (=NIES-3581)
Nitzschia alba
Pteridomonas sp.
Pteridomonas danica
non-photosynthetic phaeophyte
"Spumella" sp. NIES-1846
Spumella sp. Baeckdong012018B8
Pedospumella sp. Jangsampo120217C5
Poteriospumella lacustris
Leukarachnion salinum PRA-24
Proteolysis
FeS cluster assembly
Translation
Transcription
Electron transport
ATP synthesis
Chaperone
Protein transport
Amino acid synthesis
Fatty acid synthesis
CO₂ fixation
Haem/chlorophyll metabolism
RNA degradation
Plastid division
Pyruvate metabolism
DNA replication
Sulfate transport
C
Cryptomonas paramecium
Toxoplasma gondii
Nitzschia sp. irils04 (=NIES-3581)
Nitzschia alba
Pteridomonas sp.
Pteridomonas danica
non-photosynthetic phaeophyte
"Spumella" sp. NIES-1846
Spumella sp. Baeckdong012018B8
Pedospumella sp. Jangsampo120217C5
Poteriospumella lacustris
Leukarachnion salinum PRA-24
clpC
sufB
sufC
rps2
rps3
rps4
rps5
rps7
rps8
rps9
rps10
rps11
rps12
rps13
rps14
rps16
rps17
rps18
rps19
rps20
rpl1
rpl2
rpl3
rpl4
rpl5
rpl6
rpl11
rpl12
rpl13
rpl14
rpl16
rpl18
rpl19
rpl20
rpl21
rpl22
rpl23
rpl24
rpl27
rpl29
rpl31
rpl32
rpl33
rpl34
rpl35
rpl36
tufA
infB
rpoA
rpoB
rpoC1
rpoC2
ycf29
petF
atpA
atpB
atpD
atpE
atpF
atpG
atpH
atpI
dnaK
groEL
ycf60/90
ycf80
secA
secG
secY
tatC
ilvB
ilvH
acpP
rbcL
rbcS
cbbX
chlI
pbsA
ycf41
dnaB
ycf19/89
ycf20
ycf66
cemA
Proteolysis
FeS cluster assembly
Translation
Transcription
Electron transport
ATP synthesis
Chaperone
Protein transport
Amino acid synthesis
Fatty acid synthesis
CO2 fixation
Haem/chlorophyll metabolism
DNA replication
unkown/poorly defined function


A
bb
mbv
N
MT
mbv
MT
B
m
m
V
m
V
V
m
N
icw
ocw
C
m
m
N
m
D
E
V
ly
ly
ly
V
F
pf
af
G
MT
pbb
abb
H
bb
I


PLASTID
HAEM SYNTHESIS (C5 PATHWAY)
Glutamyl-tRNA Glu
GluRS
tRNA
HemA
Glutamate-1-semialdehyde
HemL
δ-Aminolevulinic acid
HemB
Porphobilinogen
HemC
Hydroxymethylbilane
HemD
Uroporphyrinogen
HemE
Coproporphyrinogen
HemS
Protoporphyrinogen
HemY
Protoporphyrin
HemH
Protoheme
TRANSCRIPTION
trnE
ptDNA
TOCOTRIENOL SYNTHESIS
γ-Tocotrienol
VTE1
6-Geranylgeranyl-2,3
-dimethylbenzene-1,4-diol
δ-Tocotrienol
VTE1
VTE3
6-Geranyl-2-methylbenzene-1,4-diol
VTE2
Geranylgeranyl-PP
Homogentisate
Shikimate
SK
Shikimate-3-phosphate
EPSP
5-Enolpyruvylshikimate-3-phosphate (EPSP)
CS
(L-Tryptophan)
Chorismate
ADCS
4-Amino-4-deoxychorismate (ADC)
SQDG SYNTHESIS
Sulfoquinovosyldiacylglycerol
(SQD)
SQD2
UDP-6-sufoquinovose
SQD1
UDP-glucose
D-Erythrose 4-phosphate
DAHP
3-Deoxy-arabino-heptulonate
7-phosphate
DHQS
3-Dehydroquinate
DHQD
3-Dehydroshikimate
SDH
Shikimate
SHIKIMATE PATHWAY
4-Amino-4-deoxychorismate
ADCL
4-Aminobenzoate
DHPS
7,8-Dihydropteroate
DHFS
7,8-Dihydrofolate
FOLATE SYNTHESIS
MITOCHONDRION
4-Hydroxyphenyl-pyruvate
TAT
L-Tyrosine
Homogentisate
HPD
4-Hydroxyphenyl-pyruvate
Geranylgeranyl-PP
UDP-glucose
UGP2/USP
D-Glucose 1-phosphate
CYTOSOL


"Ochromonas" sp. CCMP1393
Neotessella volvocina
Synura uvella
Mallomonas splendens
Hydrurus foetidus
Pedospumella sp. Jangsampo120217C5
Spumella sp. Baeckdong012018B8
Poteriospumella lacustris
"Spumella" sp. NIES-1846
Chrysophyceae
Leukarachnion salinum PRA-24 sp. nov.
Eustigmatophyceae sp. Chic 10/23 P-6w
Trachydiscus minutus
Monodopsis sp. MarTras 21
Nannochloropsis oceanica
Characiopsis acuta
Vischeria sp. CAUP Q 202
Pseudellipsoidion edaphicum
Paraeustigmatos columelliferus
Eustigmatophyceae
Olisthodiscus luteus
Olisthodiscophyceae
Pinguiococcus pyrenoidosus
Phaemonas parva
Pinguiophyceae
Heterosigma akashiwo
Raphidophyceae
Schizocladia ischiensis
Schizocladiophyceae
Phaeophyceae sp.
Ectocarpus siliculosus
Saccharina japonica
Fucus vesiculosus
Phaeophyceae
Vaucheria litorea
Xanthophyceae
Aureococcus anophagefferens
uncultured Pelagomonas sp.
Aureoumbra lagunensis
Pelagophyceae
Florenciella parvula
Octactis speculum
Pseudopedinella elastica
"Rhizochromulina marina"
Pteridomonas danica
Pteridomonas sp. YPF1301
Dictyochophyceae
Triparma laevis
Bolidophyceae
Trieres chinensis
Thalassiosira pseudonana
Nitzschia sp. IriIs04
Nitzschia alba
Fragilariopsis cylindrus
Durinskia baltica endosymbiont
Nitzschia palea
Kryptoperidinium foliaceum endosymbiont
Fistulifera solaris
Phaeodactylum tricornutum
Bacillariophyceae
non-ochrophytes
0.6