European Journal of Phycology

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tejp20

First draft of an annotated genome for a lichenised strain of the green alga *Diplosphaera chodatii* (Prasiolales, Trebouxiophyceae)

Cécile Gueidan, Oliver L. Mead, Hadi Nazem-Bokaee & Sarah Mathews

To cite this article: Cécile Gueidan, Oliver L. Mead, Hadi Nazem-Bokaee & Sarah Mathews (2023) First draft of an annotated genome for a lichenised strain of the green alga *Diplosphaera chodatii* (Prasiolales, Trebouxiophyceae), European Journal of Phycology, 58:4, 427-437, DOI: 10.1080/09670262.2023.2165711

To link to this article: https://doi.org/10.1080/09670262.2023.2165711

EUROPEAN JOURNAL OF PHYCOLOGY
2023, VOL. 58, NO. 4, 427–437
https://doi.org/10.1080/09670262.2023.2165711

First draft of an annotated genome for a lichenised strain of the green alga *Diplosphaera chodatii* (Prasiolales, Trebouxiophyceae)

Cécile Gueidan [a], Oliver L. Mead [a,b], Hadi Nazem-Bokaee [a,b] and Sarah Mathews [a,c]

[a]Australian National Herbarium, National Research Collections Australia, NCMI, CSIRO, Canberra, ACT, 2601, Australia; [b]Synthetic Biology Future Science Platform, CSIRO, Canberra, ACT, 2601, Australia; [c]Department of Biological Sciences, University of Louisiana, Baton-Rouge, LA, 70803, USA

ABSTRACT

Although genome sequences of lichenized fungi are increasingly becoming available, genome sequences of microalgae involved in the lichen symbiosis are still scarce. For lichenized eukaryotic algae, genome sequencing has focused mostly on *Trebouxia* and *Asterochloris*, with little genomic data available for *Stichococcus*-like algae, such as *Diplosphaera*. The genus *Diplosphaera* is a common component of biological soil crusts, and often occurs associated with lichens of the family Verrucariaceae. It is characterized by cylindrical to spherical cells containing a plate-like chloroplast, and more specifically by a vegetative cell division that leads to the formation of typical two- to four-celled clusters. Here, we present a draft genome sequence for the algal partner of an Australian lichen specimen of *Endocarpon pusillum*. The genome was sequenced with Pac Bio long read and Illumina short read technologies, and transcriptome data were generated to inform the structural annotations. This algal strain is here identified as *Diplosphaera chodatii* based on nuSSU and ITS data. Compared with closely related lichenized and non-lichenized algae, the genome of *D. chodatii* stands out for its large size (85.6 Mb) and gene content (21,261 protein-encoding regions), as well as its high rate of duplicated genes (60% of the BUSCO genes are duplicated). These results suggest that whole genome duplication or large-scale segmental duplications may have occurred in the evolutionary history of this algal species.

HIGHLIGHTS

- Little genome data are available for lichenized algae.
- We generated the first genome for a lichenized *Diplosphaera chodatii*.
- Results suggest a possible whole genome duplication in this species.

ARTICLE HISTORY Received 18 September 2022; Revised 16 November 2022; Accepted 21 December 2022

KEYWORDS Algae; Australia; biological soil crusts; *Endocarpon pusillum*; lichen photobiont; microalgae; plastid DNA; *Stichococcus*-like algae; symbiosis; Trebouxiophyceae

Introduction

Diplosphaera (Trebouxiophyceae) is a microalgal genus which belongs to the green algal order Prasiolales (Pröschold & Darienko, 2020). The genus description was based on *Diplosphaera chodatii*, which was originally isolated from a lichen identified as *Lecanora tartarea* and collected from calcareous rocks in the Salève, France (Bialosuknia, 1909). This species-poor genus is part of a broader group of algae with unclear generic delimitations, often referred to as the *Stichococcus*-like clade (Pröschold & Darienko, 2020). These algae are found worldwide and occur in various habitats, from terrestrial to freshwater and marine (e.g. Butcher, 1952; Thüs *et al.*, 2011; Ettl & Gärtner, 2014). Morphologically, they are characterized by cylindrical or short-cylindrical cells containing a plate-like chloroplast, mostly without pyrenoid (Pröschold & Darienko, 2020; Van *et al.*, 2021*a*). Some of these algae are adapted to desiccation and others to high salinity stress (Medwed *et al.*, 2021; Van *et al.*, 2021*b*). Many *Stichococcus*-like algae colonize similar ecological niches, most often in arid habitats, and are a common component of biological soil crusts (e.g. Büdel *et al.*, 2009; Samolov *et al.*, 2020; Van *et al.*, 2021*a*, 2021*b*). Additionally, *Diplosphaera* is a common lichen photobiont, often found associated with lichenized fungi from the family Verrucariaceae (Tschermak-Woess, 1988; Thüs *et al.*, 2011; Fontaine *et al.*, 2012; Saini *et al.*, 2021), as well as part of the phycobiome of various other lichens (Moya *et al.*, 2017, 2021; Chiva *et al.*, 2022). In addition to their cylindrical to spherical cells with a plate-like chloroplast, *Diplosphaera* is characterized by reproduction by vegetative cell division, which leads to the formation of typical two- to four-celled clusters (Bialosuknia, 1909; Ettl & Gärtner, 1995; Pröschold & Darienko, 2020). At present, only one species is recognized in the genus (*D. chodatii*), as *D. mucosa* was shown to be a variety of *D. chodatii*, and *D. epiphytica* a synonym of *D. chodatii* (Darienko & Pröschold, 2018; Pröschold & Darienko, 2020).

Endocarpon pusillum Hedw. (Verrucariaceae), a widely distributed lichenized fungus from biological soil crusts (Eldridge & Tozer, 1997; McCune &

CONTACT Cécile Gueidan Cecile.Gueidan@csiro.au

Published online 21 Feb 2023

Rosentreter, 2007), associates with species of *Diplosphaera* (Thüs *et al.*, 2011). Due to its tractability in culture on artificial media, this lichenized fungus historically has been used as a model species for lichen morphological and developmental studies (Ahmadjian & Heikkilä, 1970; Ahmadjian & Jacobs, 1970; Janex-Favre & Wagner, 1986; Wagner, 1987; Wagner & Letrouit-Galinou, 1988; Stocker-Worgotter & Turk, 1988). More recently, genome data were generated for several strains of *E. pusillum* (Park *et al.*, 2014*a*; Wang *et al.*, 2014, 2015; Mead & Gueidan, 2020), with the goal of studying drought resistance (Wang *et al.*, 2014, 2015) or improving *in vitro* growth (Mead & Gueidan, 2020, 2021). These genome sequences augmented the number of genomes and metagenomes already available for lichenized fungi (e.g. Park *et al.*, 2014*b*; Meiser *et al.*, 2017; Armstrong *et al.*, 2018; Dal Grande *et al.*, 2018; Wang *et al.*, 2018; Armaleo *et al.*, 2019). In comparison, genomes of lichenized green algae remain sparse. Chloroplast and mitochondrial genomes have been published for *Trebouxia* (Martínez-Alberola *et al.*, 2019, 2020) and a full genome for *Asterochloris glomerata* (Armaleo *et al.*, 2019). Fortunately, high-quality genomes of free-living Trebouxiophycean microalgae are available (e.g. *Chlorella variabilis* in Blanc *et al.*, 2010; *Coccomyxa subellipoidea* in Blanc *et al.*, 2012; *Chlorella sorokiniana* in Orsini *et al.*, 2016 and Arriola *et al.*, 2018). However, no genome is currently available for *Diplosphaera*, nor for any species within the *Stichococcus*-like clade, apart for the chloroplast genome of *Stichococcus bacillaris* (Lemieux *et al.*, 2014).

As part of a project to improve the outcome of lichen *in vitro* resynthesis, including through the application of metabolic network modelling (Nazem-Bokaee *et al.*, 2021), the mycobiont and photobiont of an Australian specimen of *Endocarpon pusillum* were isolated and grown axenically (Mead & Gueidan, 2021, this study) and their genomes were sequenced, assembled and annotated. In this article we present the draft genome of the algal partner *Diplosphaera chodatii*, while the fungal genome was published previously (Mead & Gueidan, 2020). This new algal genome is described and compared with available genomes of related green algae.

Materials and methods

Strain and culture conditions

A specimen of the lichen *Endocarpon pusillum* (C. Gueidan 2364, CANB913709) was collected from Black Mountain (Canberra, Australia) in 2016 and fertile squamules were used to shoot fungal spores into potato dextrose agar (PDA, Sigma-Aldrich) plates. Plates were then inspected using a stereomicroscope and single fungal spores surrounded by hymenial algae (*Diplosphaera chodatii*) were isolated and transferred to new PDA plates. Plates were then incubated for several weeks in a Climatron growth chamber (Thermoline Scientific) under the following conditions: day cycle of 12 hours of light (200 µmol m^{-2} s^{-1}) at 20°C and 12 hours of dark at 18°C. Once algal colonies were visible, algal cells were isolated from the fungal spore using a sterile inoculation needle and transferred to plates with Bold's Basal Medium (BBM) agar (Bischoff & Bold, 1963, as modified by Starr & Zeikus, 1993; see also SAG medium recipe v. 10.2008) under the same light and temperature conditions. If no bacterial or fungal contaminations were visible, these plates were then used to collect algal cells to transfer to 250 ml Erlenmeyer flasks with 100–150 ml of liquid BBM media. These flasks were grown in the Climatron chamber on an OS1 orbital shaker (Bioline) under the same light and temperature conditions.

To obtain single cell algal cultures, subsamples of these liquid cultures were run through a FACSAria sorting flow cytometer (BD Biosciences) at the Australian National University, using the alga's natural fluorescence. Fluorescence analyses were done with FACSDiva v.8 (BD Biosciences). Single cells were retrieved on 24- to 48-well cell plates with BBM agar media. These single-cell isolates were incubated in the Climatron under the same light and temperature conditions for several weeks. Once colonies were visible, they were transferred to both liquid and solid BBM media and grown as described above. These liquid and solid cultures were transferred to new media when required. Fresh liquid cultures of *Diplosphaera chodatii* strain CS-1475 were used for DNA and RNA extractions. The strain CS-1475 was deposited at the Australian National Algae Culture Collection (CSIRO, Hobart, Australia).

Genome sequencing and assembly

Both short and long read data were generated for the *Diplosphaera chodatii* strain CS-1475. For short read sequencing, algal cells were collected from liquid BBM medium using a vacuum Steritop filter (Millipore). They were added to a tube with 0.5 mm standard silica beads (Benchmark) containing 400 µl of phenol-chloroform and 400 µl of 2% SDS-1% BME lysis buffer. They were ground in a Precellys tissue grinder (Bertin Instruments) for 3 cycles of 10s at 6000 rpm. The tube was centrifuged at 13 000 g for 5 min. The supernatant was transferred and washed two times in chloroform before being incubated at 37°C for 15 min with 1 µl of RNAse A. The DNA was then extracted using an equal volume of PEG-SPRI beads suspension (NEB). The beads were then washed with 80% ethanol, then resuspended in 30 µl of nuclease free water to elute the DNA. The bead wash was repeated, eluting in 10 µl of water. A Nextera DNA Flex (Illumina) library was prepared for the genomic DNA of *Diplosphaera chodatii* and the library was sequenced together with one other Trebouxiophycean alga onto one mid output 2 × 75 bp NextSeq 500 run. Library preparation and sequencing were done at the

Ramaciotti Centre for Genomics (UNSW Sydney, Australia).

For long read sequencing, cells were collected from liquid BBM medium using a vacuum Steritop filter (Millipore). Algal cells collected on the filter were resuspended with a small volume of fresh medium, the cell suspension transferred to three 2 ml Eppendorf tubes and the tubes centrifuged to obtain pellets. Dry pellets were then resuspended in TES buffer and ground for 15s at 6000 rpm using tubes with 0.5 mm standard silica beads (Benchmark) with a Precellys tissue grinder (Bertin Instruments). Proteinase K and RNAse A treatments were done successively after the grinding, then the genomic DNA was extracted using a CTAB protocol as described by Möller *et al.* (1992). Pippin pulse gel (Sage Science), Nanodrop (Thermo Scientific) and Qubit (Invitrogen) were used by Ramaciotti to assess the quality and quantity of genomic DNA. An additional AMPure bead cleanup (Beckman Coulter) was performed by Ramaciotti to remove remaining contaminants. A 20 kb library was then prepared by Ramaciotti and the library sequenced on one SMRT cell on the Sequel platform (Pacific Biosciences).

The assembly was done by AGRF Bioinformatics (Melbourne). PacBio subreads BAM files were parsed to FASTA using SAMtools (Danecek *et al.*, 2021). The initial assembly was constructed using Flye (Kolmogorov *et al.*, 2019). Assembly contiguity and completeness were assessed using QUAST (Gurevich *et al.*, 2013) and BUSCO v. 5.2.2 using the Chlorophyta dataset (Waterhouse *et al.*, 2018; Manni *et al.*, 2021). In parallel, Illumina reads were quality trimmed using TrimGalore (https://github.com/FelixKrueger/TrimGalore). The cleaned and trimmed Illumina reads were then aligned with the initial assembly. Further, the aligned reads and assembly were subjected to two rounds of polishing with Pilon (Walker *et al.*, 2014). To verify assembly improvements, contiguity and completeness were re-assessed as described above.

Transcriptome sequencing

For RNA sequencing, algal cells were collected from liquid BBM medium using a vacuum Steritop filter (Millipore). They were added to a tube with 0.5 mm standard silica beads (Benchmark) containing 200 µl of 2% SDS and 1% beta mercaptoethanol in TE buffer. The tube was then briefly vortexed and 200 µl of acidic phenol-chloroform was added. The suspension was ground in a Precellys tissue grinder (Bertin Instruments) using three cycles of 10s at 6000 rpm, with 3 s off between cycles. The tube was then centrifuged at 13 000 g for 3 min. The supernatant was added to a tube containing 400 µl of acidic phenol-chloroform and 200 µl of 3 M potassium acetate. The tube was briefly vortexed and centrifuged at 13 000 g for 3 min. The supernatant was then transferred to a tube containing 600 µl of RNAse-free chloroform, which was vortexed and then centrifuged for 30s at 13 000 g. The supernatant was again transferred to a tube containing 600 µl of RNAse-free chloroform, which was vortexed and then centrifuged for 30s at 13 000 g. Finally, the supernatant was added to a tube with 900 µl of isopropanol and 140 µl of 5 M sodium acetate. The tube was inverted 20 times and chilled at −20°C for 20 min before being centrifuged for 15 min at 20 000 g. The supernatant was then discarded, and the pellet washed with 200 µl of 100% ethanol, then dissolved in an appropriate amount of nuclease-free water with 0.5 µl of RiboLock RNAse inhibitor. The RNA was then purified using a DNAse I treatment and cleaned using SPRI Beads (NEB). The RNA sample was submitted to the BRF (ANU) where a 75 bp PE TruSeq library (Illumina) was prepared and sequenced on the Illumina mid output NextSeq500 platform together with one other RNA sample.

Genome annotation

The structural annotation was done by AGRF Bioinformatics (Melbourne). To generate these annotations for *Diplosphaera*, Illumina RNA sequence reads were trimmed and cleaned using TrimGalore, assembled with Flye, and polished with Pilon. They were then mapped to the polished genome assembly using START (Dobin *et al.*, 2013). Gene prediction was carried out with Augustus v. 3.4.0 (Stanke *et al.*, 2008) using the assembled transcriptome. Structural annotation statistics were obtained with BRAKER2 (Brůna *et al.*, 2021) using Augustus and GeneMark-EX v.4 (Brůna *et al.*, 2021).

The functional annotation was performed using OmicsBox v. 2.1.14 (BioBam Bioinformatics). The predicted protein sequences were loaded in FASTA format, and a BLAST search carried out using CloudBlast. The non-redundant protein sequence database nr v. 5 was queried with Chlorophyta as the taxonomy filter. The search was performed with a maximum e-value of $1.0e^{-3}$ and allowed a maximum of 20 BLAST hits. For identifying GO terms, protein sequences with BLAST hits were mapped against the GOA database v. 2021.11 (Gotz *et al.*, 2008). Next, the functional labels were retrieved using an annotation cut-off value of 55 and GO weight of five. Finally, 'Combined Analysis' was performed on the mapped sequences to identify enzyme codes by searching both the KEGG and Reactome databases.

Visual representations of the annotations for the chloroplast and mitochondrial genomes were generated with GeSeq v. 2.03 (Tillich *et al.*, 2017) and OGDRAW v 1.1.1 (Greiner *et al.*, 2019) on the Chlorobox website (chlorobox.mpimp-golm.mpg.de), using Trebouxiophyceae organellar genomes available in RefSeq (NCBI) and the *Trebouxia lynnae* organellar genomes (Martínez-Alberola *et al.*, 2019, 2020; Barreno *et al.*, 2022). The resulting graphs were edited using Illustrator v. 25.4.1 (Adobe).

Long genome sequence reads, short genome sequence reads and transcriptome sequence reads have been deposited into the NCBI Sequence Read Archive (PRJNA606981). Assembly and annotations are also available from NCBI under the same bioProject (PRJNA606981). Mitochondrial and chloroplast genome sequences and annotations were deposited to NCBI under the accession numbers OP846047 and OP846046, respectively.

Strain phylogenetic placement

As both GenBank BLAST comparisons and a previous phylogenetic framework at the Trebouxiophyceae level (Thüs *et al.*, 2011) suggested that *Endocarpon* photobionts belong to *Diplosphaera*, the following phylogenetic analysis was restricted to a clade including this genus and its sister taxon *Tetratostichococcus*, as shown in Pröschold & Darienko (2020). Sequences of three strains of *Deuterostichococcus epilithicus* were used as an outgroup (Pröschold & Darienko, 2020). The sequences included the nuclear ribosomal SSU and ITS regions. To find the same regions in the genome of our strain of *Diplosphaera chodatii*, a sequence from GenBank (MT078181) was mapped onto our assembled genome using Geneious Prime v. 2022.1.1 (Geneious) and the corresponding sequence saved as FASTA. All sequences (Supplementary table S1) were aligned using Mesquite v. 3.61 (Maddison & Maddison, 2011). Ambiguous regions delimited as described in Lutzoni *et al.* (2000) were excluded from the alignment. The dataset was analysed using a Maximum likelihood (ML) criterion with RAxML v. 8.2.12 (Stamatakis *et al.*, 2005, 2008), as implemented on the CIPRES Web Portal (http://www.phylo.org; Miller *et al.*, 2010). The model used for the tree search was GTRCAT and 1000 pseudo-replicates were carried out for the bootstrap analysis. The resulting tree was visualized with FigTree v. 1.4.4 (https://github.com/rambaut/figtree) and edited with Illustrator v. 25.4.1 (Adobe).

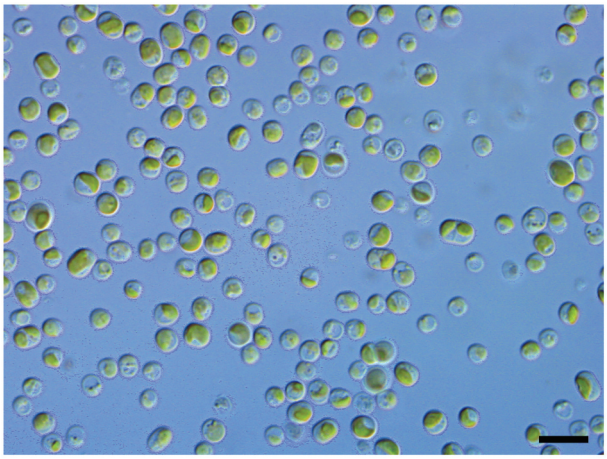

Fig. 1. Morphology of the Australian strain of *Diplosphaera chodatii* isolated from the lichen *Endocarpon pusillum* and grown in axenic BBM liquid culture. Scale bar: 20 µm. Photo: C. Gueidan.

Results and discussion

Strain phylogenetic placement

The isolated algal strain displayed the characteristics of the algal genus *Diplosphaera*, with short cylindrical to spherical cells, which sometimes occurred as two-celled clusters, and a plate-like chloroplast in each cell (Fig. 1). The nuclear ribosomal SSU and ITS alignment included

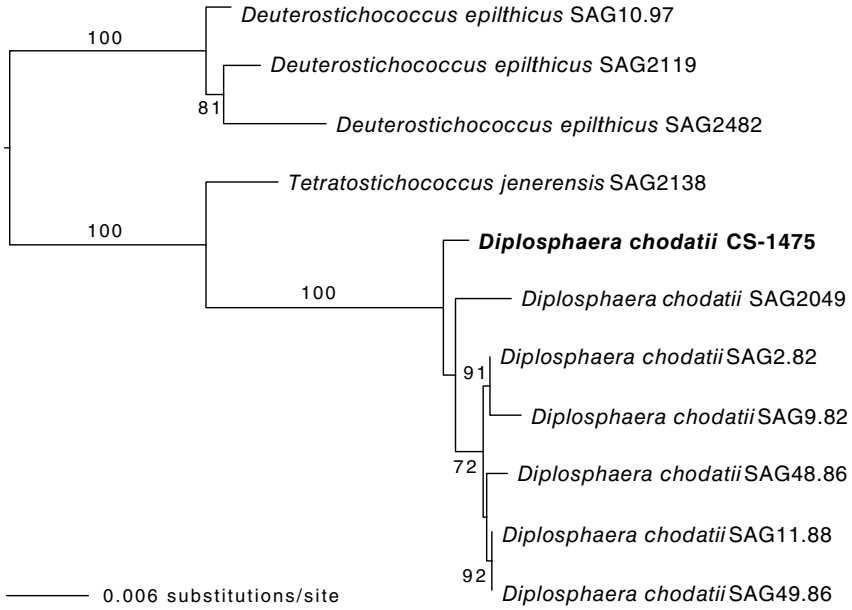

Fig. 2. Phylogenetic position of the lichenized Australian strain CS-1475 of *Diplosphaera chodatii* based on the nuclear ribosomal SSU-ITS region and a Maximum likelihood criterion. Support values are shown above or below the branches.

Table 1. Assembly metrics for the draft genome of *Diplosphaera chodatii* CS-1475.

Assembly	Primary	Polished
# contigs (≥ 1000 bp)	61	61
# contigs (≥ 5000 bp)	58	58
# contigs (≥ 10 000 bp)	56	55
# contigs (≥ 25 000 bp)	52	52
# contigs (≥ 50 000 bp)	47	47
Total length (≥ 0 bp)	87 054 862	85 619 781
Total length (≥ 1000 bp)	87 053 077	85 618 345
Total length (≥ 5000 bp)	87 046 337	85 611 608
Total length (≥ 10 000 bp)	87 032 241	85 588 384
Total length (≥ 25 000 bp)	86 968 555	85 534 733
Total length (≥ 50 000 bp)	86 792 440	85 358 783
# contigs	63	62
Largest contig	4 438 261	4 337 559
Total length	87 054 862	85 619 288
GC (%)	60	60
N50	2 671 111	2 626 102
N75	1 705 131	1 687 480
L50	14	14
L75	24	24
# N's per 100 kbp	0	0

Table 2. Assembly completeness of the draft genome of *Diplosphaera chodatii* CS-1475 as estimated using BUSCO and the Chlorophyta dataset.

Metrics	Number	%
Complete BUSCOs (C)	1344	88.4
Complete and single-copy BUSCOs (S)	432	28.4
Complete and duplicated BUSCOs (D)	912	60.0
Fragmented BUSCOs (F)	44	2.9
Missing BUSCOs (M)	131	8.7
Total BUSCO groups searched	1519	

a total of 4003 nucleotide positions, 2776 of which could be unambiguously aligned. This corresponded to 118 distinct alignment patterns, and no gaps were present. The resulting ML tree (Fig. 2) shows high support (100% bootstrap) for the placement of the *Diplosphaera* strain CS-1475 within a clade of *Diplosphaera chodatii* strains. It, therefore, confirms the identification of the *Diplosphaera* associated with the lichenized fungus *Endocarpon pusillum* as *Diplosphaera chodatii*, as recently circumscribed by Pröschold & Darienko (2020).

Whole genome assembly and annotations

PacBio long and Illumina short reads were assembled into 62 contigs (Table 1). Most contigs (47 out of 62) were longer than 50 000 bp. The largest contig was 4.3 Mb, and the N50 value was 2.6 Mb. The total length of the assembly was 85.6 Mb and the estimated coverage about 200×. Based on a BUSCO search on the Chlorophyta dataset (1519 genes), the genome is estimated to be 88.4% complete (Table 2). Among the complete set of BUSCO genes, 60% were duplicated. The GC content was of 60.2% for the nuclear genome, 27.1% for the chloroplast genome and 26.5% for the mitochondrial genome (Table 3).

A total of 21 261 protein-encoding regions were structurally annotated using Augustus. The mean gene length was 3149 bp and the mean number of exons per CDS was 10 (Supplementary table S2). From the total of 21 261 protein sequences structurally annotated, 17 680 sequences (83%) were identified with LAST hits, 16 374 (77%) were mapped to GO terms, and 11 197 (53%) annotated with functional labels. There were 3581 protein sequences (17%) with no LAST hits. The average length of the protein sequences was 508 aa and 99.4% of all protein sequences had a length smaller than 2500 aa (Supplementary fig. S1).

Figure 3 shows the counts of protein sequences (for 75% of total mapped sequences) based on the assigned GO terms for each of the main GO categories (i.e. biological processes, cellular components and molecular functions). Under the 'biological processes' GO category (Fig. 3a), the main protein functions are metabolic, translation, transport and phosphorylation. The top 'molecular functions' were predicted to be those acting as transferases, hydrolases, oxidoreductases as well as metal ion and

Table 3. Comparison of genome statistics of *Diplosphaera chodatii* CS-1475 and other Trebouxiophycean algae, either lichenized or free-living.

	Diplosphaera chodatii	*Asterochloris glomerata*[b]	*Chlorella variabilis*[b]	*Coccomyxa subellipoidea*[b]	*Trebouxia lynnae*[b]	*Stichococcus bacillaris*[b]
Nuclear Genome						
Coverage	~200×	24.8×	9×	12×	na[a]	na[a]
Number of scaffolds	62	151	413	29	na[a]	na[a]
Genome size (Mb)	85.6	56	46.2	48.8	na[a]	na[a]
Number of predicted genes	21 261	10 025	9791	9851	na[a]	na[a]
Nuclear GC content (%)	60.2	na	67	53	na[a]	na[a]
Average gene length (bp)	3149	4240	2928	3503	na[a]	na[a]
Average number of introns per gene	9	9	7.3	7	na[a]	na[a]
Organellar genomes						
Mitochondrion – genome length (bp)	83 919	110 932	na[a]	65 497	70 070	na[a]
Mitochondrion – number of proteins	29	32	na[a]	na[a]	33	na[a]
Mitochondrion – GC content (%)	26.5	na[a]	na[a]	53.2	32.7	na[a]
Chloroplast – genome length (bp)	174 073	217 546	na[a]	175 731	303 163	116 952
Chloroplast – number of proteins	75	73	na[a]	na[a]	77	na[a]
Chloroplast – GC content (%)	27.1	na[a]	na[a]	50.7	31.9	31.9

[a]Entries labelled 'na' were entries for which data were not available. [b]Available data were found in Armaleo *et al.* (2019, *Asterochloris glomerata*), Blanc *et al.* (2010, *Chlorella variabilis*), Blanc *et al.* (2012, *Coccomyxa subellipsoidea*), Martínez-Alberola *et al.* (2019, 2020, *Trebouxia lynnae*) and Lemieux *et al.* (2014, *Stichococcus bacillaris*).

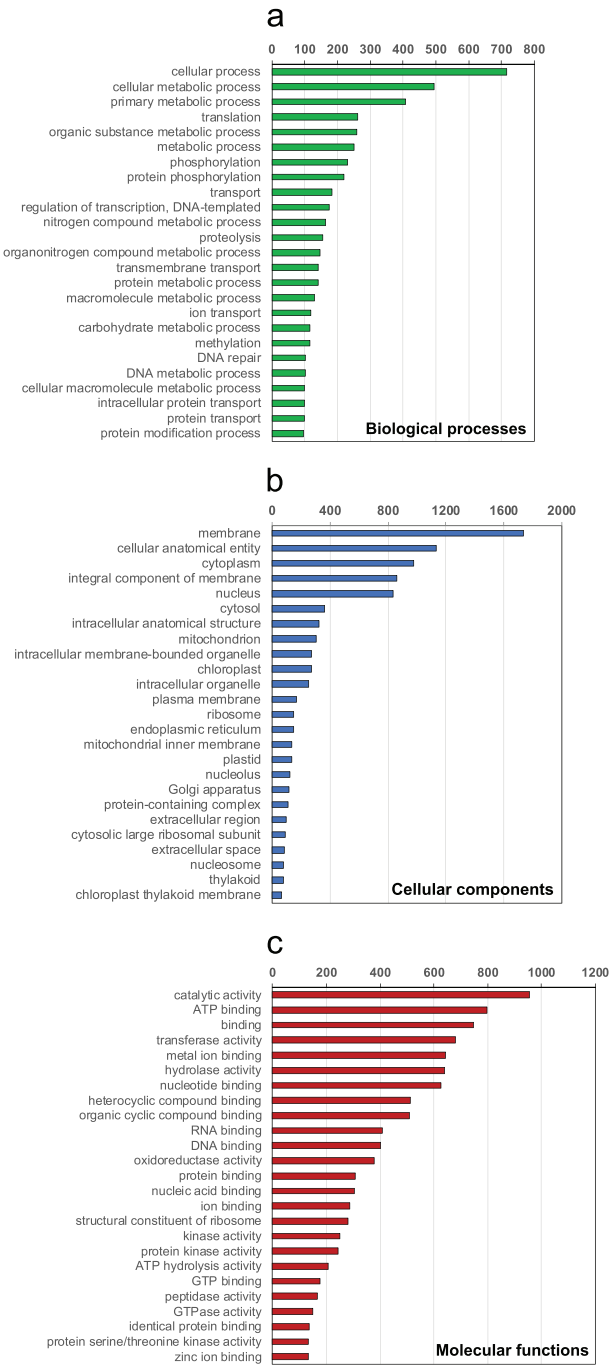

Fig. 3. Counts of protein sequences in *Diplosphaera chodatii* based on their GO annotations: **a**) biological processes, **b**) cellular components, **c**) molecular functions. The sequence counts were only shown for the first 25 GO terms (including repeated sequences) in each GO category and at all GO levels (up to level 5).

nucleotide binding activities (Fig. 3b). With respect to the 'cellular components' GO category (Fig. 3c), the counts for protein sequences in membranes were almost double the counts of sequences in other major compartments such as the cytosol or nucleus.

Organellar genomes

The chloroplast genome was 174 073 bp long and included 75 protein-encoding genes (Fig. 4). Compared with the complete chloroplast genome of *Trebouxia lynnae* (Martínez-Alberola *et al.*, 2020), 71 of the 77 protein-encoding genes annotated in *Trebouxia* were also found in *Diplosphaera*. Genes found in *Trebouxia* but not annotated in *Diplosphaera* were magnesium chelatase I (*chl*I), photosystem II protein N (*psb*N), tRNA-lysidine synthetase S (*til*S), photosystem I assembly protein 3 (*ycf*3), photosystem I assembly protein 4 (*ycf*4) and photosystem I assembly protein 12 (*ycf*12). Conversely, the ribosomal protein S4 (*rps*4), the ATP synthase C (*atp*C), photosystem I assembly protein 62 (*ycf*62) and photosystem I assembly protein 78 (*ycf*78) were annotated in *Diplosphaera* but not in *Trebouxia*. This contrasts with *Trebouxia*, for which *rps*4 was found in the nuclear genome, but not in the chloroplast genome (Martínez-Alberola *et al.*, 2020).

The mitochondrial genome was 83 919 bp long and included 29 protein-encoding genes (Fig. 5). Compared with the complete mitochondrial genome of *Trebouxia lynnae* (Martínez-Alberola *et al.*, 2019), 29 of the 33 protein-encoding genes annotated in *Trebouxia* were also found in *Diplosphaera*. As in other available Trebouxiophyceae mitochondrial genomes, NADH dehydrogenase 10 (*nad*10), ribosomal protein L14 (*rpl*14) and ribosomal protein S8 (*rps*8) seemed absent in *Diplosphaera*. Additionally, ribosomal protein L10 (*rpl*10), ribosomal protein S4 (*rps*4), ribosomal protein S7 (*rps*7) and ribosomal protein S11 (*rps*11) were not apparent in the *Diplosphaera* genome. Further work is necessary to confirm if they are absent.

Comparison with other algal nuclear and organellar genomes and possible nuclear genome duplication

The genome of *Diplosphaera chodatii* CS-1475 was compared with the published genomes of two lichenized algae, *Asterochloris glomerata* (Armaleo *et al.*, 2019) and *Trebouxia lynnae* (Martínez-Alberola *et al.*, 2019, 2020; organellar genomes only), and three free-living algae, *Chlorella variabilis* (Blanc *et al.*, 2010), *Coccomyxa subellipsoidea* (Blanc *et al.*, 2012) and *Stichococcus bacillaris* (Lemieux *et al.*, 2014). All these taxa belong to the Trebouxiophyceae. The organellar genomes of *Diplosphaera chodatii* appear as standard in comparison to these other species (Table 3). The length of the mitochondrial genome of *D. chodatii* is 83 919 bp and within the range (65 497–110 932 bp) of the three other algal species with available data (*A. glomerata*, *C. subellipsoidea* and *T. lynnae*). The number of mitochondrial protein-encoding genes are also similar among these algal species, with 29 genes in *D. chodatii*, 32 in *A. glomerata*

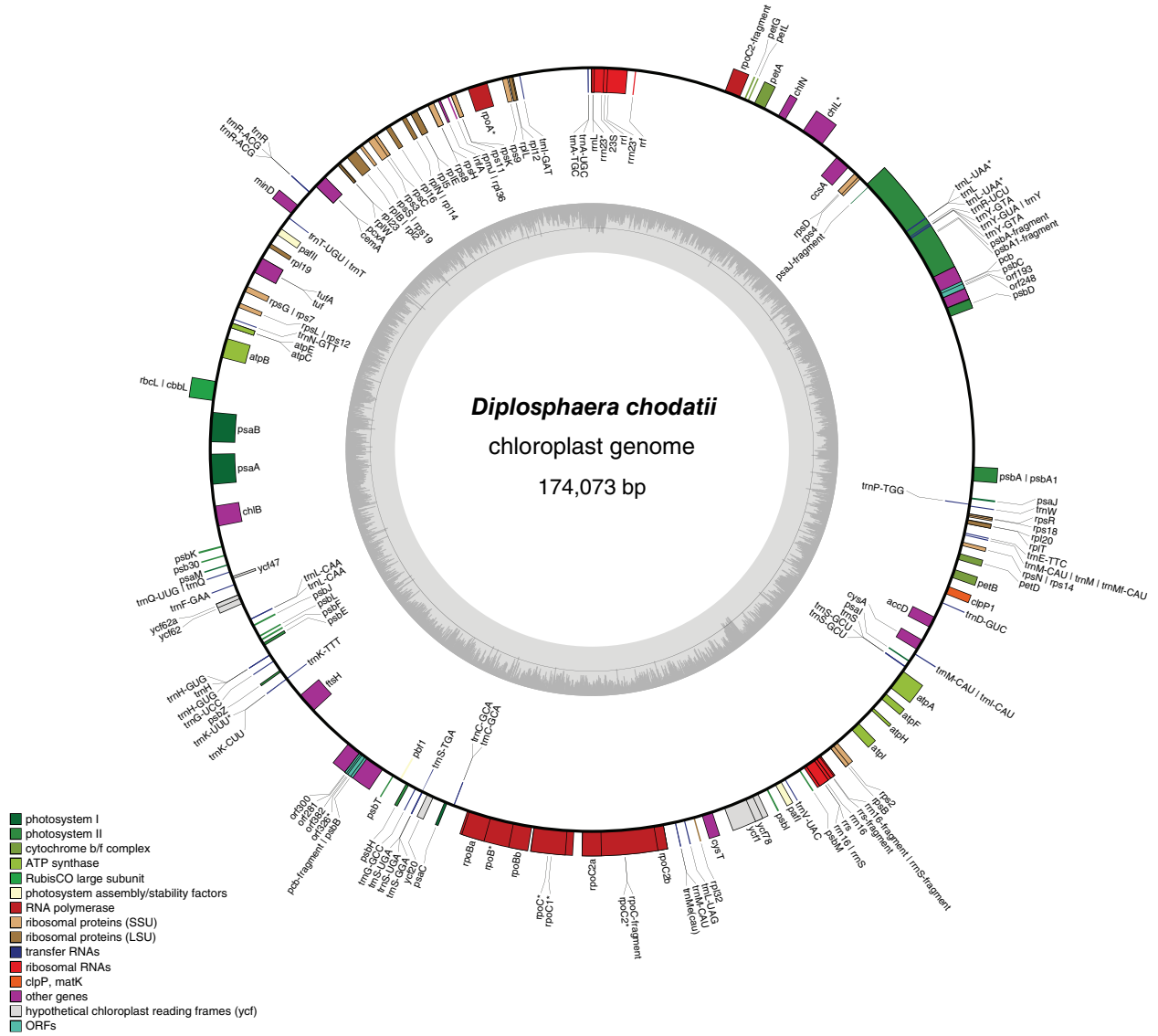

Fig. 4. Gene map of the draft chloroplast genome of the green alga *Diplosphaera chodatii* strain CS-1475. The inner graph represents the GC content. Transcription directions are as follows: clockwise if inside the circle and counterclockwise if outside the circle. An asterisk indicates the presence of at least one intron.

and 33 in *T. lynnae*. The length of the chloroplast genome of *D. chodatii* (174 073 bp) also falls within the range (116 952–303 163 bp) found in other algae it was compared with (Table 3), and the number of chloroplast protein-encoding genes was similar among these algal species (75 genes for *D. chodatii*, 73 for *A. glomerata* and 77 for *T. lynnae*).

However, the nuclear genome of *D. chodatii* is significantly larger in size (85.6 Mb vs. 48.8–56 Mb) and in number of predicted genes (21 261 genes for *D. chodatii* vs. 10 025 genes for *A. glomerata*) than other algae it was compared with. In eukaryotes, several processes can lead to increased nuclear genome size, including whole genome duplication and the accumulation of transposable elements (Ayala-Usma *et al.*, 2021). Here, in addition to the larger genome size and higher number of genes, the BUSCO analysis reveals an important duplication of core orthologous genes, hinting more towards whole genome duplication or, at least, the duplication of large-scale fragments of the genome. This genome duplication could be the result of autoploidy or hybridization. Although further work will be necessary to determine the cause of the large genome size and high gene duplication rate in this strain of *D. chodatii*, this work highlights the importance of genome studies of lichenized algae. Thanks to their association with fungi, these poorly studied microalgae may have evolved unique genome characteristics and biological functions and processes.

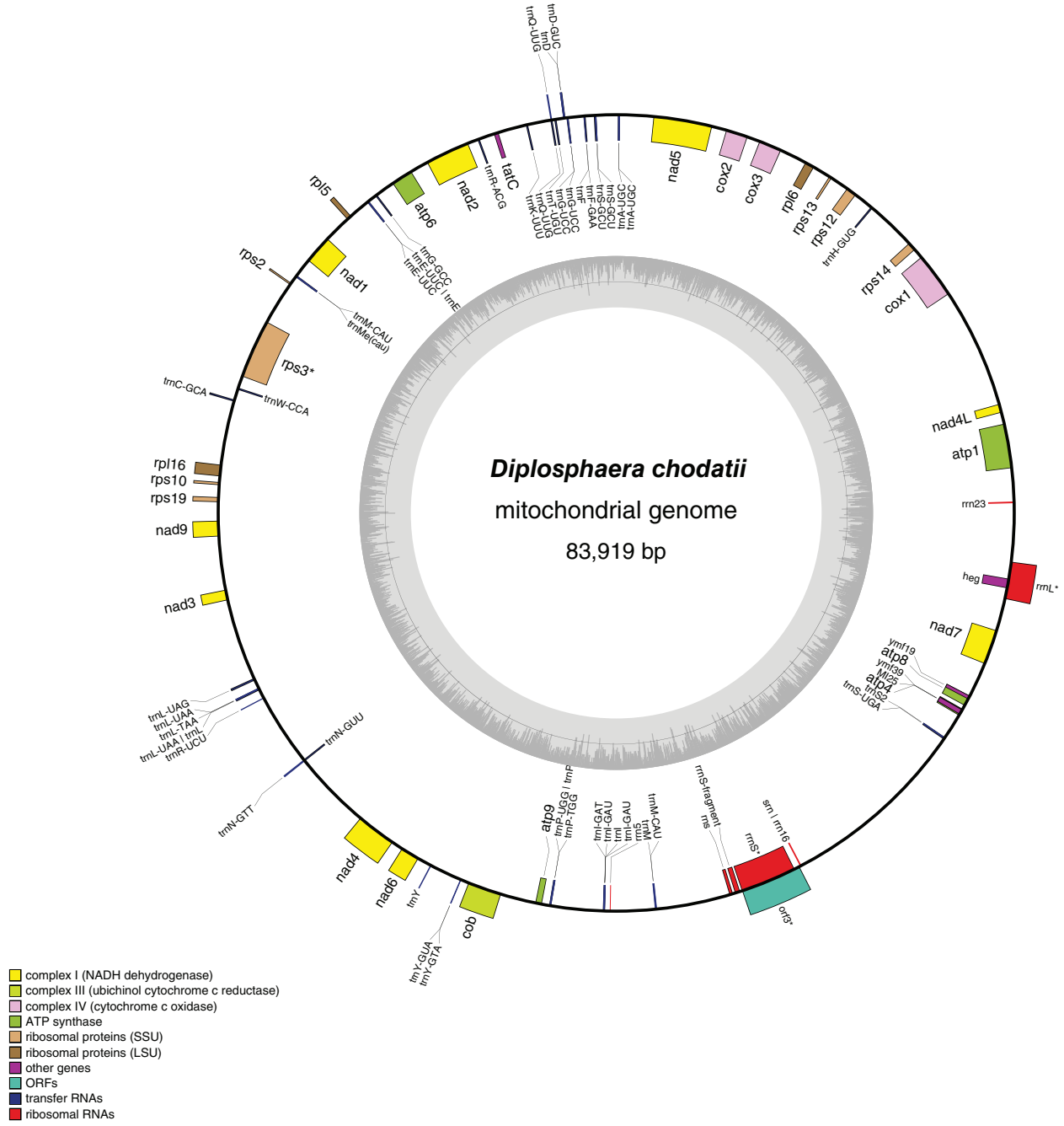

Fig. 5. Gene map of the draft mitochondrial genome of the green alga *Diplosphaera chodatii* strain CS-1475. The inner graph represents the GC content. Transcription directions are as follows: clockwise if inside the circle and counterclockwise if outside the circle. An asterisk indicates the presence of at least one intron.

Acknowledgements

The authors would like to thank Lan Li (CSIRO) for her technical support, Harpreet Vohra (ANU) for sharing her expertise in flow cytometry, and Naga Kasinadhuni (AGRF) for his bioinformatic support. We also thank the staff at ANACC (CSIRO) for their assistance.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was funded by a grant from the CSIRO Synthetic Biology Future Science Platform (OD-206013).

Supplementary material

The following supplementary material is accessible via the Supplementary Content tab on the article's online page at https://doi.org/10.1080/09670262.2023.2165711

Supplementary figure S1. Distribution of the length of protein sequences in the draft genome of *Diplosphaera chodatii* CS-1475.

Supplementary table S1. Species and strains used for the phylogenetic analysis.
Supplementary table S2. Structural annotation metrics for the draft genome of *Diplosphaera chodatii* CS-1475.

Author contributions

C. Gueidan: original concept, culture and molecular work, drafting and editing manuscript; O. Mead: culture and molecular work and editing manuscript; H. Nazem-Bokaee: data analysis and editing manuscript; S. Mathews: original concept and editing manuscript.

ORCID

Cécile Gueidan http://orcid.org/0000-0001-9295-9996
Oliver L. Mead http://orcid.org/0000-0001-7919-9515
Hadi Nazem-Bokaee http://orcid.org/0000-0002-4246-1804
Sarah Mathews http://orcid.org/0000-0002-5518-7541

References

Ahmadjian, V. & Heikkilä, H. (1970). The culture and synthesis of *Endocarpon pusillum* and *Staurothele clopima*. *The Lichenologist*, **4**: 259–267.
Ahmadjian, V. & Jacobs, J.B. (1970). The ultrastructure of lichens. III. *Endocarpon pusillum*. *The Lichenologist*, **4**: 268–270.
Armaleo, D., Müller, O., Lutzoni, F., Andrésson, Ó.S., Blanc, G., Bode, H.B., Collart, F.R., Dal Grande, F., Dietrich, F., Grigoriev, I.V., Joneson, S., Kuo, A., Larsen, P.E., Logsdon Jr., J.M., Lopez, D., Martin, F., May, S.P., McDonald, T.R., Merchant, S.S., Miao, V., Morin, E., Oono, R., Pellegrini, M., Rubinstein, N., Sanchez-Puerta, M.V., Savelkoul, E., Schmitt, I., Slot, J. C., Soanes, D., Szövényi, P., Talbot, N.J., Veneault-Fourrey, C. & Xavier, B.B. (2019). The lichen symbiosis re-viewed through the genomes of *Cladonia grayi* and its algal partner *Asterochloris glomerata*. *BMC Genomics*, **20**: 605.
Armstrong, E.E., Prost, S., Ertz, D., Westberg, M., Frisch, A. & Bendiksby, M. (2018). Draft genome sequence and annotation of the lichen-forming fungus *Arthonia radiata*. *Genome Announcements*, **6**: e00281–18.
Arriola, M.B., Velmurugan, N., Zhang, Y., Plunkett, M.H., Hondzo, H. & Barney, B.M. (2018). Genome sequences of *Chlorella sorokiniana* UTEX 1602 and *Micractinium conductrix* SAG 241.80: implications to maltose excretion by a green alga. *Plant Journal*, **93**: 566–586.
Ayala-Usma, D.A., Cárdenas, M., Guyot, R., Chaib De Mares, M., Bernal, A., Reyes Muñoz A. & Restrepo, S. (2021). A whole genome duplication drives the genome evolution of *Phytophthora betacei*, a closely related species to *Phytophthora infestans*. *BMC Genomics*, **22**: 795.
Barreno, E., Muggia, L., Chiva, S., Molins, A., Bordenave, C., García-Breijo, F. & Moya, P. (2022). *Trebouxia lynnae* sp. nov. (former *Trebouxia* sp. TR9): biology and biogeography of an epitome lichen symbiotic microalga. *Biology*, **11**: 1196.
Bialosuknia, M.W. (1909). Sur un nouveau genre de Pleurococcacées. *Bulletin de la Société Botanique de Genève, Série 2*, **1**: 101–104.
Bischoff, H.W. & Bold, H.C. (1963). Phycological studies. IV. Some soil algae from Enchanted Rock and related algal species. *University of Texas Publications*, **6318**: 1–95.
Blanc, G., Duncan, G., Agarkova, I., Borodovsky, M., Gurnon, J., Kuo, A., Lindquist, E., Lucas, S., Pangilinan, J., Polle, J., Salamov, A., Terry, A., Yamada, T., Dunigan, D.D., Grigoriev, I.V., Claverie, J. M. & Van Etten, J.L. (2010). The *Chlorella variabilis* NC64A genome reveals adaptation to photosymbiosis, coevolution with viruses, and cryptic sex. *Plant Cell*, **22**: 2943–2955.
Blanc, G., Agarkova, I., Grimwood, J., Kuo, A., Brueggeman, A., Dunigan, D.D., Gurnon, J., Ladunga, I., Lindquist, E., Lucas, S., Pangilinan, J., Pröschold, T., Salamov, A., Schmutz, J., Weeks, D., Yamada, T., Lomsadze, A., Borodovsky, M., Claverie, J. M., Grigoriev, I.V., & Van Etten, J.L. (2012). The genome of the polar eukaryotic microalga *Coccomyxa subellipsoidea* reveals traits of cold adaptation. *Genome Biology*, **13**: R39.
Brůna, T., Hoff, K.J., Lomsadze, A., Stanke, M. & Borodovsky, M. (2021). BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. *NAR Genomics and Bioinformatics*, **3**: 1–11.
Büdel, B., Darienko, T., Deutschewitz, K., Dojani, S., Friedl, T., Mohr, K.I., Salisch, M., Reisser, W. & Weber, B. (2009). Southern african biological soil crusts are ubiquitous and highly diverse in drylands, being restricted by rainfall frequency. *Microbial Ecology*, **57**: 229–247.
Butcher, R.W. (1952). Contributions to our knowledge of the smaller marine algae. *Journal of the Marine Biological Association of the United Kingdom*, **31**: 175–191.
Chiva, S., Moya, P. & Barreno E. (2022). Lichen phycobiomes as source of biodiversity for microalgae of the *Stichococcus*-like genera. *Biologia*, **77**: 1–9.
Dal Grande, F., Meiser, A., Greshake Tzovaras, B., Otte, J., Ebersberger, I. & Schmitt, I. (2018). The draft genome of the lichen-forming fungus *Lasallia hispanica* (Frey) Sancho & A. Crespo. *The Lichenologist*, **50**: 329–340.
Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M. & Li, H. (2021). Twelve years of SAMtools and BCFtools. *Gigascience*, **10**: giab008.
Darienko, T. & Pröschold, T. (2018). The establishment of *Diplosphaera epiphytica* sp. nov. and its distinction from *Chlorella sphaerica* Tschermak-Woess. *Notulae Algarum*, **54**: 1–3.
Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. & Gingeras TR. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**: 15–21.
Eldridge, D.J. & Tozer, M.E. (1997). *A practical guide to soil lichens and bryophytes of Australia's dry country*. Department of Land and water Conservation, Sydney.
Ettl, H. & Gärtner, G. (1995). *Syllabus der Boden-, Luft- und Flechtenalgen*. Gustav Fischer Verlag, Stuttgart.
Ettl, H. & Gärtner, G. (2014). *Syllabus der Boden-, Luft- und Flechtenalgen*. Springer, Berlin and Heidelberg.
Fontaine, K.M., Beck, A., Stocker-Wörgötter, E. & Piercey-Normore, M.D. (2012). Photobiont relationships and phylogenetic history of *Dermatocarpon luridum* var. *luridum* and related *Dermatocarpon* species. *Plants*, **1**: 39–60.

Gotz, S., Garcia-Gomez, J.M., Terol, J., Williams, T.D., Nagaraj, S.H., Nueda, M.J., Robles, M., Talon, M., Dopazo, J. & Conesa, A. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Research*, **36**: 3420–3435.

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

Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. (2013). QUAST: Quality assessment tool for genome assemblies. *Bioinformatics*, **29**: 1072–1075.

Janex-Favre, M.C. & Wagner, J. (1986). Recherches ontogéniques et structurales sur les pycnides des Dermatocarpacées (pyrénolichens). *Bulletin de la Société Mycologique de France*, **102**: 161–182.

Kolmogorov, M., Yuan, J., Lin, Y. & Pevzner, P.A. (2019). Assembly of long, error-prone reads using repeat graphs. *Nature Biotechnology*, **37**: 540–546.

Lemieux, C., Otis, C. & Turmel, M. (2014). Chloroplast phylogenomic analysis resolves deep-level relationships within the green algal class Trebouxiophyceae. *BMC Evolutionary Biology*, **14**: 211.

Lutzoni, F., Wagner, P., Reeb, V. & Zoller, S. (2000). Integrating ambiguously aligned regions of DNA sequences in phylogenetic analyses without violating positional homology. *Systematic Biology*, **49**: 628–651.

Maddison, W.P. & Maddison, D.R. (2011). *Mesquite: a modular system for evolutionary analysis*. version 3.61.

Martínez-Alberola, F., Barreno, E., Casano, L.M., Gasulla, F., Molins, A. & del Campo, E.M. (2019). Dynamic evolution of mitochondrial genomes in Trebouxiophyceae, including the first completely assembled mtDNA from a lichen-symbiont microalga (*Trebouxia* sp. TR9). *Scientific Reports*, **9**: 8209.

Martínez-Alberola, F., Barreno, E., Casano, L.M., Gasulla, F., Molins, A., Moya, P., Gonzàlez-Hourcade, M. & del Campo, E.M. (2020). The chloroplast genome of the lichen-symbiont microalga *Trebouxia* sp. TR9 (Trebouxiophyceae, Chlorophyta) shows short inverted repeats with a single gene and loss of the *RPS*4 gene, which is encoded by the nucleus. *Journal of Phycology*, **56**: 170–184.

Manni, M., Berkeley, M.R., Seppey, M., Simão, F.A. & Zdobnov, E.M. (2021). BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Molecular Biology and Evolution*, **38**: 4647–4654.

McCune, B. & Rosentreter, R. (2007). Biotic soil crust lichens of the Columbia basin. *Monographs in North American Lichenology*, **1**: 1–105.

Mead, O.L. & Gueidan, C. (2020). Complete genome sequence of an Australian strain of the lichen-forming fungus *Endocarpon pusillum* Hedw. *Microbiology Resource Announcements*, **9**: e01079–20.

Mead, O.L. & Gueidan C. (2021). Testing carbon and nitrogen sources for the *in vitro* growth of the model lichenized fungus *Endocarpon pusillum* Hedw. *The Lichenologist*, **53**: 257–264.

Medwed, C., Holzinger, A., Hofer, S., Hartmann, A., Michalik, D., Glaser, K. & Karsten, U. (2021). Ecophysiological, morphological, and biochemical traits of free living *Diplosphaera chodatii* (Trebouxiophyceae) reveal adaptation to harsh environmental conditions. *Protoplasma*, **258**: 1187–1199.

Meiser, A., Otte, J., Schmitt, I. & Grande, F.D. (2017). Sequencing genomes from mixed DNA samples - evaluating the metagenome skimming approach in lichenized fungi. *Scientific Reports*, **7**: 14881.

Miller, M.A., Pfeiffer, W. & Schwartz, T. (2010). *Creating the CIPRES Science Gateway for inference of large phylogenetic trees*. 2010 Gateway Computing Environments Workshop (GCE), New Orleans, LA.

Möller, E.M., Bahnweg, G., Sandermann, H. & Geiger, H. H. (1992). A simple and efficient protocol for isolation of high molecular weight DNA from filamentous fungi, fruit bodies, and infected plant tissues. *Nucleic Acids Research*, **20**: 6115–6116.

Moya, P., Molins, A., Martínez-Alberola, F., Muggia, L. & Barreno, E. (2017). Unexpected associated microalgal diversity in the lichen *Ramalina farinacea* is uncovered by pyrosequencing analyses. *PLoS ONE*, **12**: e0175091.

Moya, P., Chiva, S., Molins, A., Garrido-Benavent, I. & Barreno, E. (2021). Unravelling the symbiotic microalgal diversity in *Buellia zoharyi* (lichenized ascomycota) from the Iberian Peninsula and Balearic Islands using DNA metabarcoding. *Diversity*, **13**: 220.

Nazem-Bokaee, H., Hom, E.F.Y., Warden, A.C., Mathews, S., Gueidan, C. (2021). Towards a systems biology approach to understanding the lichen symbiosis: opportunities and challenges of implementing network modelling. *Frontiers in Microbiology*, **12**: 667864.

Orsini, M., Cusano, R., Costelli, C., Malavasi, V., Concas, A., Angius, A. & Cao, G. (2016). Complete genome sequence of chloroplast DNA (cpDNA) of *Chlorella sorokiniana*. *Mitochondrial DNA Part A*, **27**: 838–839.

Park, S.Y., Choi, J., Lee, G.W., Park, C.H., Jung, A.K., Oh, S. O., Lee, Y.H. & Hur, J.S. (2014a). Draft genome sequence of *Endocarpon pusillum* strain KoLRILF000583. *Genome Announcements*, **2**: e00452–14.

Park, S.Y., Choi, J., Lee, G.W., Jeong, M.H., Kim, J.A., Oh, S.O., Lee, Y.H. & Hur, J.S. (2014b). Draft genome sequence of *Umbilicaria muehlenbergii* KoLRILF000956, a lichen-forming fungus amenable to genetic manipulation. *Genome Announcements*, **2**: e00357–14.

Pröschold, T. & Darienko, T. (2020). The green puzzle *Stichococcus* (Trebouxiophyceae, Chlorophyta): New generic and species concept among this widely distributed genus. *Phytotaxa*, **441**: 113–142.

Saini, K.C., Bast, F., Nayaka, S. & Gautam, A.K. (2021). Morpho-molecular characterization of rock-inhabiting lichen *Dermatocarpon miniatum* (Verrucariaceae, Ascomycota) and its symbiont in Indian Himalayas. *The Nucleus*, **64**: 193–202.

Samolov, E., Baumann, K., Büdel, B., Jung, P., Leinweber, P., Mikhailyuk, T., Karsten, U., & Glaser, K. (2020). Biodiversity of algae and cyanobacteria in biological soil crusts collected along a climatic gradient in Chile using an integrative approach. *Microorganisms*, **8**: 1047.

Stamatakis, A., Ludwig, T. & Meier, H. (2005). RAxML-III: a fast program for maximum likelihood-based inference of large phylogenetic trees. *Bioinformatics*, **21**: 456–463.

Stamatakis, A., Hoover, P. & Rougemont, J. (2008). A rapid bootstrap algorithm for the RAxML web-servers. *Systematic Biology*, **75**: 758–771.

Stanke, M., Diekhans, M., Baertsch, R. & Haussler, D. (2008). Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics*, **24**: 637–644.

Starr, R.C. & Zeikus, J.A. (1993). UTEX—the culture collection of algae at the University of Texas at Austin 1993 list of cultures. *Journal of Phycology*, **29**: 1–106.

Stocker-Worgotter, E. & Turk, R. (1988). Licht- und elektronenmikroskopische Untersuchungen von Entwicklungsstadien der Flechte *Endocarpon pusillum* unter Kulturbedingungen. *Plant Systematics and Evolution*, **158**: 313–328.

Thüs, H., Muggia, L., Pérez-Ortega, S., Favero-Longo S.E., Joneson S., O'Brien H., Nelsen M.P., Duque-Thüs R., Grube M., Friedl T., Brodie J., Andrew C.J., Lücking R., Lutzoni F. & Gueidan, C. (2011). Revisiting photobiont diversity in the lichen family Verrucariaceae (Ascomycota). *European Journal of Phycology*, **46**: 399–415.

Tillich, M., Lehwark, P., Pellizzer, T., Ulbricht-Jones, E.S., Fischer, A., Bock, R. & Greiner, S. (2017). GeSeq – versatile and accurate annotation of organelle genomes. *Nucleic Acids Research*, **45**: W6–W11.

Tschermak-Woess, E. (1988). The algal partner. In *CRC Handbook of Lichenology*, Vol. I (Galun, M., editor), 39–92. CRC Press, Boca Raton.

Van, T.A., Karsten, U. & Glaser, K. (2021a). A chemosystematic investigation of selected *Stichococcus*-like organisms (Trebouxiophyta). *Algae*, **36**: 123–135.

Van, T.A., Sommer, V. & Glaser, K. (2021b). The eco-physiological performance and traits of genera within the *Stichococcus*-like clade (Trebouxiophyceae) under matric and osmotic Stress. *Microorganisms*, **9**: 1816.

Walker, B.J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C.A., Zeng, Q., Wortman, J., Young, S.K. & Earl, A.M. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One*, **9**: e112963.

Wagner, J. (1987). Ontogenie des périthèces d'*Endocarpon pusillum* (Pyrénolichen, Verrucariacées). *Canadian Journal of Botany*, **65**: 2441–2449.

Wagner, J. & Letrouit-Galinou, M.A. (1988). Structure et ontogénèse du thalle squamuleux de lichen *Endocarpon pusillum* (Pyrénolichens, Verrucariacées). *Canadian Journal of Botany*, **66**: 2118–2127.

Wang, Y., Liu, B., Zhang, X., Zhou, Q., Zhang, T., Li, H., Yu, Y., Zhang, X., Hao, X., Wang, M, Wang, L. & Wei, J. C. (2014). Genome characteristics reveal the impact of lichenization on lichen-forming fungus *Endocarpon pusillum* Hedwig (Verrucariales, Ascomycota). *BMC Genomics*, **15**: 34.

Wang, Y.Y., Zhang, X.Y., Zhou, Q.M., Zhang, X.L. & Wei, J.C. (2015). Comparative transcriptome analysis of the lichen-forming fungus *Endocarpon pusillum* elucidates its drought adaptation mechanisms. *Science China Life Sciences*, **58**: 89–s100.

Wang, Y., Yuan, X., Chen, L., Wang, X., & Li, C. (2018). Draft genome sequence of the lichen-forming fungus *Ramalina intermedia* strain YAF0013. *Genome Announcements*, **6**: e00478–18.

Waterhouse, R.M., Seppey, M., Simão, F.A., Manni, M., Ioannidis, P., Klioutchnikov, G., Kriventseva, E.V. & Zdobnov, E.M. (2018). BUSCO applications from quality assessments to gene prediction and phylogenomics. *Molecular Biology & Evolution*, **35**: 543–548.