DR. MAREN PREUSS (Orcid ID : 0000-0002-8147-5643)

DR. HEROEN VERBRUGGEN (Orcid ID : 0000-0002-6305-4749)

DR. GIUSEPPE C. ZUCCARELLO (Orcid ID : 0000-0003-0028-7227)

Article type : Regular Article

The organelle genomes in the photosynthetic red algal parasite *Pterocladophila hemisphaerica* (Florideophyceae, Rhodophyta) have elevated substitution rates and extreme gene loss in the plastid genome ¹

Maren Preuss²

School of Biological Sciences, Victoria University of Wellington, PO Box 600, Wellington,
6140, New Zealand

Heroen Verbruggen

School of BioSciences, University of Melbourne, Parkville, VIC 3010, Australia
and

Giuseppe C. Zuccarello

School of Biological Sciences, Victoria University of Wellington, PO Box 600, Wellington,
6140, New Zealand

Running title: *Pterocladophila* organelle genomes

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as doi: [10.1111/IPY.12996-20-002](https://doi.org/10.1111/IPY.12996-20-002)

This article is protected by copyright. All rights reserved

¹ Received _____ Accepted _____

² Corresponding author: Maren142@web.de

Editorial Responsibility: M. Coleman (Associate Editor)

ABSTRACT

Comparative organelle genome studies of parasites can highlight genetic changes that occur during the transition from a free-living to a parasitic state. Our study focuses on a poorly studied group of red algal parasites, which are often closely related to their red algal hosts and from which they presumably evolved. Most of these parasites are pigmented and some show photosynthetic capacity. Here, we assembled and annotate the complete organelle genomes of the photosynthetic red algal parasite, *Pterocladophila hemisphaerica*. The plastid genome is the smallest known red algal plastid genome at 68,701 bp. The plastid genome has many genes missing, including all photosynthesis-related genes. In contrast, the mitochondrial genome is similar in architecture to that of other free-living red algae. Both organelle genomes show elevated mutation rates and significant changes in patterns of selection, measured as dN/dS ratios. This caused phylogenetic analyses, even of multiple aligned proteins, to be unresolved or give contradictory relationships. Full plastid datasets interfered by selected best gene evolution models showed the supported relationship of *P. hemisphaerica* within the Ceramiales, but the parasite was grouped with support as sister to the Gracilariales when interfered under the GHOST model. Nuclear rDNA showed a supported grouping of the parasite within a clade containing several red algal orders including the Gelidiales. This photosynthetic parasite which is unable to photosynthesize with its own plastid, due to the total loss of all photosynthesis genes, raises intriguing questions on parasite-host organelle genome capabilities and interactions.

Key index words: dN/dS; General Heterogenous Evolution On a Single Topology (GHOST) model; heterotachy; mitochondrial genome; parasitism; reduced plastid genome; positive selection; relaxed selection

Abbreviations: CTAB, cetyl trimethylammonium bromide

INTRODUCTION

Organelle (mitochondrial and plastid) genomes in parasites are useful to understand evolutionary changes occurring during the transition from free-living to parasitic modes of life. Comparative plastid genome studies in plants showed extensive gene loss after loss of photosynthesis (Wicke et al. 2013, Barrett et al. 2018, Schneider et al. 2018) and no clear convergent gene loss patterns between parasitic taxa (Wicke and Naumann 2018). Generally, genes loss is correlated with a decrease in genome size and an increase in AT content (Delannoy et al. 2011, Wicke et al. 2013, Petersen et al. 2015, Cusimano and Wicke 2016, Su et al. 2019). In contrast, changes in mitochondrial genome architecture in parasites in comparison to free-living taxa are less clear (Burger et al. 2003), as the size of mitochondrial genomes in free-living taxa can vary greatly (Sloan et al. 2012, Petersen et al. 2015, Su et al. 2019). Comparative mitochondrial studies indicate that parasite mitochondrial genomes have rather conserved architecture compared to their free-living relatives (Hancock et al. 2010, Fan et al. 2016). Few studies investigate and characterize mitochondrial and plastid genomes together and similarities between these organelle genomes in gene loss and selection might be overlooked.

One ideal group to study these organelle evolution patterns in parasites are red algal parasites, which are red algae parasitic on other red algae. These parasites are highly diverse, with over 123 species (Preuss et al. 2017, Preuss and Zuccarello 2017, 2019a) and many independent transitions to parasitism (Goff et al. 1996, Blouin and Lane 2016, Preuss et al. 2017). A majority of these parasites are at least partially pigmented (Preuss et al. 2017) and at least one species (*Pterocladophila hemisphaerica*) is able to photosynthesize independently when removed from its host (Preuss and Zuccarello 2019b). While parasitic red algae are diverse, they are poorly studied.

Red algal parasites mostly infect members of the same family (Goff et al. 1996, Preuss and Zuccarello 2014, 2017) or occasionally different families within the same order (Zuccarello et al. 2004). The close relationships between many red algal parasites and their hosts led to the proposition that red algal parasites evolved from their host (Setchell 1918, Goff et al. 1997). Phylogenetic analyses demonstrated that some parasites and their host are more closely related to each other than to other species in the same host genus (Goff et al. 1997, Zuccarello et al. 2004, Preuss and Zuccarello 2017), whereas other parasites are more distantly related to their host species, possibly due to host switching (Kurihara et al. 2010, Preuss and Zuccarello 2014, 2017).

In addition, red algal parasites exhibit a unique organelle transfer mechanism of infection by forming a connection between a parasite and a host cell, called a secondary pit connection (Goff and Coleman 1984, 1985), leading to cells containing cellular components of both species (“heterokaryons”) within the host and parasite thallus. Some studies have suggested that the heterokaryon, in some cases, transformed into a parasite cell, reduces host nuclei but keeps host plastids (Goff and Zuccarello 1994, Goff and Coleman 1995). Newly formed parasite cells would then produce reproductive structures that contained parasite nuclei but host plastids (Goff and Coleman 1995). While earlier evidence suggested that parasites only contained host plastids (Goff and Coleman 1995, Zuccarello et al. 2004); a reduced plastid (i.e., lacking all photosynthetic genes) was found in the unpigmented parasites *Choreocolax polysiphoniae* (Salomaki et al. 2015) and *Harveyella mirabilis* (Salomaki and Lane 2019). Currently, it is unknown if a similar highly reduced plastid is present in other red algal parasites, especially ones that are pigmented.

For this study, we choose the pigmented red algal parasite *Pterocladophila hemisphaerica* from New Zealand. This parasite is able to photosynthesize independently when removed from its host *Pterocladia lucida* (Gelidiales; Preuss and Zuccarello 2019b), and its taxonomic position in the order Gracilariales is still based solely on morphological similarities with two other red algal parasite genera, *Holmsella* and *Gelidiocolax* (Fredericq and Hommersand 1990) placed in this order. Our general aims of this study are to: (i) compare organelle genome architecture and characteristics between the parasite *P. hemisphaerica* and its host species, (ii) analyze selection patterns on individual genes of organelle genes in this parasite, and (iii) investigate the phylogenetic placement of this parasite.

MATERIALS AND METHODS

Algal specimens, DNA extraction and partial gene sequencing.

Red algal specimens of *Pterocladia lucida* and its parasite *Pterocladophila hemisphaerica* were collected from shore (drift) at Akitio Beach (40°37'25" S, 176°24'39" E) in November 2011 and Kairakau Beach (39°56'30" S, 176°55'50" E) in May 2013, or by SCUBA in August 2016 at Princess Bay, Wellington, New Zealand (41°20'46" S, 174°47'26" E). Drift specimens were dried in silica gel and used to amplify single genes, whereas scuba collections were freshly ground in liquid nitrogen and used for High-Throughput-Sequencing (HTS).

Parasite pustule was held with forceps and cut off with a razor blade at the base with as much distance from the host-parasite contact area as possible. Some pustules were

cystocarpic but spores were not released in culture. All samples were extracted using a modified CTAB protocol (Zuccarello and Lokhorst 2005). Extracted DNA of drift specimens were used to amplify partial *cox1*, LSU and SSU rDNA following established PCR conditions, purification and sequencing (Preuss and Zuccarello 2017). *Cox1* sequences were used to identify the clade of *Pterocladia* host in New Zealand, as three cryptic species exist (Boo et al. 2016).

Sequencing, assembly, and annotation of organelle genomes.

Library preparation and sequencing for the parasite *Pterocladophila hemisphaerica* (n≈100) and one uninfected specimen of *Pterocladia lucida* (host, n=1) of fresh material from Princess Bay were performed separately using Illumina TruSeq DNA nano by Macrogen Inc. (Seoul, Korea). Libraries of 350 bp were sequenced using a HiSeq 2500 with read lengths of 101 bp and paired ends. Sequenced reads were trimmed with the CLC Genomic Workbench 7.5.1 (CLC bio, Aarhus, Denmark) with a quality threshold of 0.05. De novo assembly in CLC and SPAdes 3.8.1 (Nurk et al. 2013) were performed using automatic k-mer size and default parameters. Plastid, mitochondrial and nuclear contigs were identified with Blast searches against a custom-build database containing known red algal genes. Long contigs identified as mtDNA and cpDNA were imported into Geneious 8.0.5 (<https://www.geneious.com>). Different assemblers gave similar results but showed slight differences in lengths and further analysis was continued with SPAdes assemblies. Organelle genome circularity was first manually checked by mapping 1000 bp of the start and end sequences of the SPAdes contigs against the CLC scaffold and then against the raw data. Gene prediction was carried out in MFannot (<http://megasun.bch.umontreal.ca/cgi-bin/mfannot/mfannotInterface.pl>) and tRNA prediction in ARAGORN (<http://130.235.46.10./ARAGORN/>), manually checked and annotated in Geneious. Open reading frames (ORF) were used to identify missing genes and were manually annotated. No pseudogenes were observed. The circular mitochondrial and plastid genomes were visualised in OGDRAW (Greiner et al. 2019). Biological functions of protein coding genes were determined in UniProt (<http://uniprot.org>) and conserved domains blasted against the NCBI site (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>).

Annotation of nuclear rDNA.

Previously partial amplified nuclear genes of *Pterocladophila hemisphaerica* and *Pterocladia lucida* were blast searched against contigs identifying whole SSU rDNA and LSU

rDNA sequences and confirming identical overlapping sequences. RNAmmer prediction server (Lagesen et al. 2007) was used to predict the beginning and end of genes.

Analysis of mitochondrial and plastid gene selection.

Mitochondrial and plastid genes of the parasite were further analyzed to determine selection patterns. First, non-synonymous (dN) and synonymous (dS) substitution between four species for each organelle genome dataset (parasite: *Pterocladophila hemisphaerica*, host: *Pterocladia lucida*, species in host genus: *Pterocladia capillacea* [mt] or *Gelidium vagum* [cp] and outgroup: *Corallina officinalis* [mt] or *Calliarthron tuberculosum* [cp]) were calculated in SNAP (<https://www.hiv.lanl.gov/content/sequence/SNAP/SNAP.html>) and used to calculate dN/dS for each gene. Comparison between parasite and free-living, and free-living to free-living red algae were used to show if genes are under increased selection in the parasite. Second, aBSREL (adaptive branch-site random effect likelihood; Smith et al. 2015) in HyPhy (<http://datamonkey.org/>) was used to determine if the branch of *P. hemisphaerica* was under positive selection. For every gene, the branch of *P. hemisphaerica* (parasite) was tested against all other branches included in the phylogeny. Maximum number of included taxa were 83 for plastid alignments and 45 taxa for mitochondrial alignments. Genetic code 4 (Mold/Protozoan mtDNA) were used for mitochondrial genes and genetic code 1 (universal code) was used for plastid genes as genetic code 11 (bacterial code) is not available. RELAX (Wertheim et al. 2014) in HyPhy was used to determine if relaxed selection pressure occurred in mitochondrial and plastid genes of *P. hemisphaerica*. Analyses was identical to aBSREL.

Comparative plastid genome analysis.

Progressive Mauve alignment was used to compare plastid genomes using the full alignment option with default seed weight and automatically determined locally collinear blocks (LCB) score (Darling et al. 2004). Compared plastid genomes were the *Pterocladophila hemisphaerica* (newly determined) and its host *Pterocladia lucida* (newly determined), the two complete parasite plastid genomes at present (*Choreocolax polysiphoniae*, *Harveyella mirabilis*) and representatives plastid genomes within the Gelidiales, Gracilariales and other red algal orders (Table S1 in the Supporting Information).

Phylogenetic analysis to determine host specificity.

Host specificity of *Pterocladophila hemisphaerica* was determined using *cox1* and alignment was partitioned by codon and tested for the best codon model using ModelFinder

(Kalyaanamoorthy et al. 2017) within IQ-Tree (Trifinopoulos et al. 2016). For 1st and 2nd codon positions the TNe model and for 3rd codon HKY+F+I model were selected. IQ-Tree was used to construct maximum-likelihood (ML) trees with 100 bootstrap plus 1000 SH-aLRT (Guindon et al. 2010) and 1000 ultrafast bootstrap analyses (Minh et al. 2013).

Alignment for phylogenetic analysis.

PRANK alignments in wasabi (<http://wasabiapp.org>; Veidenberg et al. 2016) were used for nuclear rRNA genes. Models selected by ModelFinder in IQ-tree for nuclear genes were TvMe+I+G4 for LSU and GTR+F+I+G4 for SSU.

Taxon selection for phylogenetic analysis was based on all available data for mitochondrial and plastid genomes in the subclass Rhodymeniophycidae with the Corallinales as outgroup, following Verbruggen et al. (2010; Table S1). All protein coding genes were translated into amino acid sequences and aligned using MAFFT (Katoh and Standley 2013) in Geneious. All mitochondrial and plastid genes were trimmed using the best automated method in TrimAL (<http://trimal.cgenomics.org>). Only shared plastid genes between the parasites *Pterocladophila hemisphaerica* and *Choreocolax polysiphoniae* were included in the phylogenies. Plastid and mitochondrial data (cp: 67 genes, mt: 24 genes) set were portioned by genes and best models selected using ModelFinder (Tables S2-S3 in the Supporting Information).

Phylogenetic analyses were performed for nuclear, plastid and mitochondrial datasets as described above. The phylogenetic relationship of the parasite was mostly unresolved, or showed contradictory relationships between all three datasets. MrBayes interference was not performed because it is known to overestimate support (Simmons et al. 2004). In addition, the branch of *P. hemisphaerica* was noticeably long in the mitochondrial and plastid phylogenies. The inconsistent phylogenetic position of *P. hemisphaerica* and the observation of long branches led to further analyses of the mitochondrial and plastid datasets to resolve the phylogenetic relationship and decrease branch length of the parasite.

1. Removal of genes with elevated rates in mitochondrial and plastid alignments.

Genes with elevated rates of evolution in *Pterocladophila hemisphaerica* were removed in the mitochondrial and plastid alignment. Elevated gene rates in *P. hemisphaerica* were calculated by using the ratio of uncorrected distances of nucleotides between the outgroup (*Calliarthron* sp.) and *P. hemisphaerica* and between an ingroup (mitochondria: *Schimmelmanna*, plastid: *Caloglossa*) and *P. hemisphaerica*. Only 25% of the mitochondrial

and plastid genes with the lowest elevated rates were retained (cp: 17 genes, mt: 6 genes, Table S2-S3). Evolutionary models were selected for the retained mitochondrial and plastid genes and phylogenetic analyses were performed as described above (Tables S2-S3).

2. Model selection for heterotachous-evolved sequences.

Untrimmed nucleotide alignments of the full mitochondria and plastid datasets (cp: 67, mt: 24 genes) were performed under GTR, four class mixture model (GTR+FO*H4) following General Heterogeneous Evolution On a Single Topology (GHOST) model in IQ-tree (Crotty et al. 2019) with 100 bootstrap, plus 1000 SH-aLRT and 1000 ultrafast bootstrap support.

RESULTS

Parasite plastid genome.

The circular mapping plastid genome of *Pterocladia hemisphaerica* was generated with a total of 91,423 reads with a mean coverage of 134X. The genome is highly reduced, consisting of only 68,701 bp containing only 70 genes (Fig. 1, Table 1) and lacks any photosynthesis and ATP synthesis genes but many genes for genetic systems, metabolism, ribosomal proteins and transport are still present (Table S4 in the Supporting Information). The plastid genome is densely packed with only 13% non-coding regions and all protein coding genes (67.5-85.7% AT content), and rRNAs (63.8-67.5%) and tRNAs (47.3-73.0%) show an A-T bias. The rRNA 5S gene (*rrn5*) is also missing from the parasite plastid genome (Table S4). The parasite plastid contains several ORFs not found in the host or other red algae: *orf114* has a ribosomal protein L22 conserved domain, *orf151* with a N-terminal reserve transcriptase domain and *orf407* is without any conserved domains. The host, *Pterocladia lucida*, has a standard red algal plastid genome size (176,635 bp) and organization (Figs. S1-S2 in the Supporting Information), and shares many genes with other free-living red algae (Fig. S2). The plastid genome of the host was generated with a total 629,867 reads with a mean coverage of 360X.

The parasites *Pterocladia hemisphaerica*, *Choreocolax polysiphoniae* and *Harveyella mirabilis* have in common a highly reduced plastid genome, but it is more reduced in *P. hemisphaerica* (68,701 versus ~90k bp) with a core of shared genes (n=54) and a few unique genes (Fig. S3 in the Supporting Information). Among the 8 genes (*accA*, *accB*, *accD*, *rpl32*, *rpl33*, *rnz*, *ycf19*, *ycf21*) that *P. hemisphaerica* shares with the parasite *H. mirabilis* are genes for fatty acid biosynthesis whereas the 3 genes (*acpP*, *dnaB*, *fabH*) shared with the parasite *C. polysiphoniae* are genes for fatty acid biosynthesis and DNA replication. *C.*

polysiphoniae and *H. mirabilis* share 6 genes (*ilvB*, *ilvH*, *rne*, *rp1*, *rpl22*, *rpl29*) for biosynthetic processes and replication that are not found in *P. hemisphaerica* (Fig S3). In comparison to the plastid genome of free-living red algae, the reduced plastid genome of *Harveyella mirabilis* has no inversions or rearrangements, whereas different rearrangements are found in the plastid genomes of *C. polysiphoniae* and *P. hemisphaerica* compared to free-living red algae (Fig. S2). Gene arrangements and inversions are different between *P. hemisphaerica* and *C. polysiphoniae*. (Fig. S2). All inversions of the plastid genome of *C. polysiphoniae* are not found in any compared free-living red algae, whereas one inversion was only observed in the genome of *P. hemisphaerica* and two inversions are shared with the order Gracilariales and Gelidiales (Fig. S2).

Selection in parasite plastid genes.

Parasite plastid genes compared to three free-living red algae show that the majority (54 out of 66 genes) have higher dN/dS ratios than the comparison between only the three free-living taxa. dN/dS ratios of all compared plastid genes were below 1 with the exception of *rpl2* (Table 2).

The branches of seven parasite plastid genes (*accB*, *odpA*, *rpl3*, *rpl33*, *rps3*, *rps13*, *ycf62*) showed significant positive selection in comparison to all other branches in the phylogeny (Table 2, Table S5 in the Supporting Information). Increased selection intensity ($k > 1$) was significantly different in 10 plastid parasite genes (*dnaB*, *fabH*, *rpl20*, *rpl32*, *rpl35*, *rpoC1*, *rps11*, *rps16*, *rps16*, *ycf19*) compared to all other taxa, and with no obvious gene selection (e.g., gene function; Table 2, Table S6 in the Supporting Information).

Mitochondrial genome.

The circular mapping mitochondrial genome of *Pterocladia hemisphaerica* was generated by a total of 174,214 reads with a mean coverage of 690X and is 25,486 bp long (Fig. S4 in the Supporting Information). The mitochondrial genome contains 24 protein coding genes and has similar architecture to the mitochondrial genome of its host *Pterocladia lucida* (cryptic species Group II; Tables 3, S7, Figs. S5-S6 in the Supporting Information). The mitochondrial genome of the host was generated by a total of 83,093 reads with a mean coverage of 332X. The parasite mitochondrial genome is extremely densely packed with less than 8% non-coding regions and a high A-T content in all protein coding (70.7-89.1% AT content), rRNAs (59.5-73.7%) and tRNAs (71.7-79.1%) genes (Table S7).

Selection in parasite mitochondrial genes.

Parasite mitochondrial genes compared to three free-living red algae show that the majority (17 of 23 genes) have higher dN/dS ratios compared to ratios between the three free-living taxa. dN/dS ratios of all compared mitochondrial genes were below 1 (Table 4). The branches of three mitochondrial genes (*rps11*, *sdh2*, *tatC*) showed significant positive selection in comparison to all other branches in the phylogeny (Table 4, Table S8 in the Supporting Information). Increased selection intensity ($k > 1$) was significant different in two mitochondrial parasite genes (*nad5*, *rps11*) compared to all other taxa (Table S9 in the Supporting Information).

Phylogeny

Nuclear DNA

The concatenated nuclear dataset (LSU and SSU rDNA) contained 202 taxa and was 3,611 bp long. ML topology showed a strong UFbootstrap and SH-aLRT support of the parasite *P. hemisphaerica* grouping with the Gelidiales and other red algal orders. This grouping excludes the Gracilariales and Ceramiales. The position of the parasite amongst the other red algal orders is unsupported (Fig. S7 in the Supporting Information).

cpDNA

The concatenated plastid dataset contained 67 genes, 81 taxa and was 16,981 amino acids long. ML topology showed a supported relationship of *Pterocladophila hemisphaerica* on an extreme long branch with the parasite *Choreocolax polysiphoniae* within the Ceramiales (Fig. S8 in the Supporting Information). After removal of plastid genes with elevated rates, the remaining dataset consisted of 17 genes (*acpP*, *dnaB*, *infC*, *odpB*, *rpl4*, *rpl11*, *rpl21*, *rpl23*, *rpl36*, *rpoB*, *rps12*, *rps16*, *rps19*, *syh*, *trpA*, *ycf19*, *ycf62*) and was 4,130 amino acids long. ML topology grouped *P. hemisphaerica* on a long branch in a moderately supported clade containing the orders Gelidiales and Gracilariales. The parasite was sister to the Gracilariales but without support (Fig. S9 in the Supporting Information). The concatenated plastid dataset with a length of 56,679 nucleotides, including 67 genes and 81 taxa, produced a ML topology under the GHOST model showing a supported relationship of the parasites as sister to the Gracilariales (Fig. 2). While the parasite branch is long it is reduced in length using this evolutionary model compared to other topologies (i.e., Figs S8-S9).

mtDNA

The full concatenated mitochondrial dataset contained 24 genes, 45 taxa and was 5,616 amino acids long. ML topology grouped *P. hemisphaerica* on a long branch as sister to the Ceramiales without support (Fig. S10 in the Supporting Information). After removal of mitochondrial genes with elevated rates of evolution, the remaining dataset consisted of 6 genes (*atp6*, *cox1*, *cox3*, *rps3*, *rps12*, *secY*) and was 1,605 amino acids long. ML topology showed an unsupported relationship of *P. hemisphaerica* as sister to the Plocamiales (Fig. S11 in the Supporting Information). The full concatenated mitochondrial dataset with a length of 18,981 nucleotides including 24 genes and 45 taxa and ML topology interfered under the GHOST model grouped the parasite on a long branch as sister to the Ceramiales without support (Fig. S12 in the Supporting Information).

Host organelle genomes in parasite dataset.

Host mtDNA, nDNA and cpDNA were identified within the HTS data of the parasite tissue. The number of reads were high enough to assemble and annotate whole plastid (total reads: 717,546; mean coverage: 410X) and mitochondrial genomes (total reads 131,147; mean coverage: 524X) of the host from this tissue. Host organelle genomes sequenced separately, were almost identical to host contigs derived from parasite HTS data (3 bp differences each).

DISCUSSION

The main finding of this study is the reduced plastid genome of *Pterocladophila hemisphaerica*, which is the smallest known plastid genome in red algae. The parasite's plastid genome is compact, 5S rRNA (*rrn5*) and many genes are lost, and genes are under higher selection than free-living algae. In contrast, the mitochondrial genome of the parasite is similar in gene content and genome size to free-living red algae but also shows increased sequence divergence and selection. For the first time, phylogenetic relationships of a red algal parasite and their host were studied using complete organelle datasets, but we were still not able to resolve the phylogenetic placement of *P. hemisphaerica*.

The evolution of *Pterocladophila hemisphaerica* from a free-living ancestor to a parasite has led to a highly reduced plastid genome. Genome reduction was first seen in the red algal parasite, *Choreocolax polysiphoniae* (Salomaki et al. 2015) and later in *Harveyella mirabilis* (Salomaki and Lane 2019). Extensive plastid gene reduction is a characteristic of many unpigmented parasitic plants (Bungard 2004, Krause 2008, Bellot et al. 2016). While plastid genomes rearrangements, compared to free-living red alga, appear to be independent of one another in *P. hemisphaerica* and *C. polysiphoniae*, they share the majority of protein

coding genes, with a similar complement of genes lost, mostly photosynthesis-related genes. Studies of parasitic plants with different degrees of photosynthetic ability showed that rearrangements and gene deletion similarities and difference can be traced between taxa (Wicke et al. 2013, Ravin et al. 2016, Frailey et al. 2018). Increased taxon sampling of red algal parasites, with different relationships to their hosts, will show if there are any general patterns in gene loss/rearrangement in their evolution.

Pterocladophila hemisphaerica while missing plastid encoded photosynthetic genes is able to photosynthesize independently when removed from its host (Preuss and Zuccarello 2019b). All genes involved in photosynthesis (*ndh*, *atp*, *pet*, *psa*, *psb*, *rbc*) with the exception of *petF* are lost in the plastid genome of *P. hemisphaerica*. The question then arises, do the parasite cells rely on host plastids for photosynthesis? All photosynthesis-related genes may have been relocated to the nuclear genome of the parasite, but more likely the parasite uses host plastids, with photosynthetic function, residing in heterokaryotic cells, cells containing both parasite and host organelles (Zuccarello and West 1994, Goff and Coleman 1995, Blouin and Lane 2016). This is further supported by the ease in which we were able to reconstruct host organelles genomes from a ‘parasite’ tissue dataset, in which we would expect few, if any, embedded host cells. This is the first example of a photosynthetic parasite with almost complete loss of photosynthetic genes in its plastid genome, and stands in contrast to all other reduced plastids in plants where the loss of the majority of photosynthetic genes is linked to a non-photosynthetic life mode of parasitic plants (Wicke et al. 2013, Naumann et al. 2016, Roquet et al. 2016). The two previous red algal parasite plastid genomes (Salomaki et al. 2015, Salomaki and Lane 2019) are from non-pigmented, and presumably non-photosynthetic, parasites.

Higher A-T content and bias in plastid genomes is a common feature of parasitic organisms in comparison to their free-living relatives. Every gene in *Pterocladophila hemisphaerica* has a higher A-T content than its host, *Pterocladia lucida*. Overall, the full plastid genome of *P. hemisphaerica* has 5% more A-T content than its host, and *Choreocolax polysiphoniae* has 13% more than its host *Vertebrata lanosa*, supporting the pattern of A-T bias in other parasitic plastid genomes (Wicke and Naumann 2018, Su et al. 2019).

Our study was not able to annotate *rrn5* (5S ribosomal RNA subunit), and it is assumed to be lost in the plastid genome of *Pterocladophila hemisphaerica*. Only non-photosynthetic plastid-containing Haemosporidia and Piroplasmida (Apicomplexa) are thought to have lost *rrn5* in their plastid genome, but it is also possible that *rrn5* is too degenerate to be recognized (Valach et al. 2014). All other non-photosynthetic and

photosynthetic plant and algal parasites still contain plastid-encoded *rrn5* (Salomaki et al. 2015, Naumann et al. 2016, Roquet et al. 2016, Suzuki et al. 2018). This pattern of loss might be observed in the future with increased sequencing from other parasite plastid genomes.

The mitochondrial genome of *Pterocladophila hemisphaerica* is highly conserved in size, architecture and gene number, similar to its host and other Florideophyceae (Yang et al. 2015, Salomaki and Lane 2017). This pattern of mitochondrial genome similarity between parasite and host has been shown in most parasitic plants (Bellot et al. 2016, Fan et al. 2016). The conservation of mitochondrial genome architecture between *P. hemisphaerica* and other red algae would indicate that they are under similar evolutionary constraints. Yet our analysis of dN/dS showed the majority of mitochondria parasite genes are under higher selection than free-living red algae. In addition, some gene are under positive or relaxed selection. Both, the parasites divergent sequences and selection might indicate early genome evolution in parasites before the mitochondrial genome starts losing genes. One of the few examples of mitochondrial genomes studied in plant parasites indicated gene loss and decreased genome size in mitochondria (Skippington et al. 2015).

Our study demonstrates for the very first time that elevated substitution rates are found in the organelle genomes of red algal parasites in comparison to free-living red algae, whether or not the genomes have lost genes. Highly divergent plastid genomes can be also found in parasitic plants and the elevated substitution rates within the plastid genomes is reflected by the parasites position on long branches in phylogenetic reconstructions (Naumann et al. 2016, Su et al. 2019). Some studies propose that the elevated substitution rates are caused by the decrease of selective pressures on the plastid genome during the transition from free-living to parasitic life style (Wicke et al. 2016, Su et al. 2019). Another possibility is that the decrease in selective pressure is influenced by nitrogen limitation, where building A-T pairs in DNA replication requires less nitrogen and energy (ATP) in comparison to G-C pairs, as shown in parasitic microorganisms (Seward and Kelly 2016). Another possibility is a combination of mutation and genetic drift (Naumann et al. 2016, Su et al. 2019). Our current knowledge of red algal parasites does not allow us to select or exclude any of these possibilities. As the number of organelle genomes in red algal parasites increases, we may be able to identify evolutionary patterns that will address these possibilities by comparing parasites with different nutrient host dependencies and nitrogen availability.

Currently, the parasite *Pterocladophila hemisphaerica* is placed taxonomically with two other parasite genera (*Holmsella* and *Gelidiocolax*) in the Gracilariales as these parasites share the morphological characteristics of a 2-celled carpogonial branch, straight

spermatangial chains and transverse divisions of spermatangial parent cells (Fredericq and Hommersand 1990). The placement of *Holmsella* spp. within the Gracilariales was confirmed with a nuclear DNA marker (Zuccarello et al. 2004). Our nuclear rRNA analysis grouped *P. hemisphaerica* in a clade with red algal orders, but not in any particular order, including the Gelidiales, but not the Gracilariales. In contrast, full amino acid phylogenies of plastid data grouped *P. hemisphaerica* as sister to *Choreocolax polysiphoniae* within the Ceramiales, on an extremely long branch, and full plastid nucleotide phylogenies interfered under the GHOST model grouped the parasite as sister to the Gracilariales. Mitochondrial datasets were not able to resolve, with support, any phylogenetic relationships of the parasite. These incongruent results for *P. hemisphaerica* demonstrate that even genome-size datasets cannot always provide reliable placement of these parasites, especially if gene sequences are highly divergent between the parasite and the non-parasitic species. Our data does not support the current placement of the parasite in the parasitic family Pterocladophilaceae within the Gracilariales. The unique morphological characters of *P. hemisphaerica* have always caused uncertainty in its taxonomic placement (Fan and Papenfuss 1959), and its placement in the Gracilariales, and family Pterocladophilaceae, was mostly due to general characters (Fredericq and Hommersand 1990). Therefore, we propose to place *Pterocladophila hemisphaerica* in *incertae sedis* until further investigations can determine the origin the parasites (including *Gelidiocolax* spp.). *Holmsella* might be either placed in the genus of its closest relative or in the resurrected parasitic genus *Holmsella* within the Gracilariales after re-analysing the dataset with current available data following previous phylogenetic studies in red algal parasites (Ng et al. 2013, Preuss and Zuccarello 2014, 2017).

One reason for incongruent results can be long branch attraction (LBA), where fast-evolving taxa group together without reflecting their phylogenetic relationship (Felsenstein 1978, Bergsten 2005), for example, between parasitic taxa and their closest free-living relatives (Morin 2000, Evans et al. 2008). Our plastid datasets show that the phylogenetic placement of a parasite can be strongly influenced by gene selection (e.g., removal of genes with elevated rates) and model selection (individual gene partition vs. GHOST model). Therefore, mutation rates and selection patterns should be carefully investigated when studying phylogenetic relationships of other organelle genomes in red algal parasites.

ACKNOWLEDGEMENTS

468 We thank Dan Crossett and Andrea Glockner-Fagetti for diving support. We also thank
469 Chiela Cremen and Christopher Jackson for technical support.

Author Manuscript

REFERENCES

- Barrett, C.F., Wicke, S. & Sass, C. 2018. Dense infraspecific sampling reveals rapid and independent trajectories of plastome degradation in a heterotrophic orchid complex. *New Phytol.* 218:1192–204.
- Bellot, S., Cusimano, N., Luo, S., Sun, G., Zarre, S., Gröger, A., Temsch, E. & Renner, S.S. 2016. Assembled plastid and mitochondrial genomes, as well as nuclear genes, place the parasite Family Cynomoriaceae in the Saxifragales. *Genome Biol. Evol.* 8:2214–30.
- Bergsten, J. 2005. A review of long-branch attraction. *Cladistics* 21:163–93.
- Blouin, N.A. & Lane, C.E. 2016. Red algae provide fertile ground for exploring parasite evolution. *Perspect. Phycol.* 3:11–9.
- Boo, G.H., Nelson, W.A., Preuss, M., Kim, J.Y. & Boo, S.M. 2016. Genetic segregation and differentiation of a common subtidal alga *Pterocladia lucida* (Gelidiales, Rhodophyta) between Australia and New Zealand. *J. Appl. Phycol.* 28:2027–34.
- Bungard, R.A. 2004. Photosynthetic evolution in parasitic plants: insight from the chloroplast genome. *BioEssays* 26:235–47.
- Burger, G., Gray, M.W. & Lang, B.F. 2003. Mitochondrial genomes: anything goes. *Trends Genet.* 19:709–16.
- Crotty, S.M., Minh, B.Q., Bean, N.G., Holland, B.R., Tuke, J., Jermin, L.S. & von Haeseler, A. 2019. GHOST: Recovering historical signal from heterotachously-evolved sequence alignments. *Syst. Biol.* 2:249–64.
- Cusimano, N. & Wicke, S. 2016. Massive intracellular gene transfer during plastid genome reduction in nongreen Orobanchaceae. *New Phytol.* 210:680–93.
- Darling, A.C.E., Mau, B., Blattner, F.R. & Perna, N.T. 2004. Mauve: Multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* 14:1394–403.
- Delannoy, E., Fujii, S., Colas des Francs-Small, C., Brundrett, M. & Small, I. 2011. Rampant gene loss in the underground orchid *Rhizanthella gardneri* highlights evolutionary constraints on plastid genomes. *Mol. Biol. Evol.* 28:2077–86.
- Evans, N.M., Lindner, A., Raikova, E. V., Collins, A.G. & Cartwright, P. 2008. Phylogenetic placement of the enigmatic parasite, *Polypodium hydriforme*, within the Phylum Cnidaria. *BMC Evol. Biol.* 8:139.
- Fan, K.C. & Papenfuss, G.F. 1959. Red algal parasites occurring on members of the Gelidiales. *Madroño*. 15:33–8.
- Fan, W., Zhu, A., Kozaczek, M., Shah, N., Pabón-Mora, N., González, F. & Mower, J.P. 2016. Limited mitogenomic degradation in response to a parasitic lifestyle in

- Orobanchaceae. *Sci. Rep.* 6:36285.
- Felsenstein, J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. *Syst. Zool.* 27:401.
- Frailey, D.C., Chaluvadi, S.R., Vaughn, J.N., Coatney, C.G. & Bennetzen, J.L. 2018. Gene loss and genome rearrangement in the plastids of five hemiparasites in the Family Orobanchaceae. *BMC Plant Biol.* 18:30.
- Fredericq, S. & Hommersand, M. 1990. Morphology and systematics of *Holmsella pachyderma* (Pterocladophilaceae, Gracilariales). *Br. Phycol. J.* 25:39–51.
- Goff, L.J., Ashen, J. & Moon, D. 1997. The evolution of parasites from their hosts: a case study in the parasitic red algae. *Evol.* 51:1068–78.
- Goff, L.J. & Coleman, A.W. 1984. Transfer of nuclei from a parasite to its host. *Proc. Natl. Acad. Sci. USA* 81:5420–4.
- Goff, L.J. & Coleman, A.W. 1985. The role of secondary pit connections in red algal parasitism. *J. Phycol.* 21:483–508.
- Goff, L.J. & Coleman, A.W. 1995. Fate of parasite and host organelle DNA during cellular-transformation of red algae by their parasites. *Plant Cell.* 7:1899–911.
- Goff, L.J., Moon, D.A., Nyvall, P., Stache, B., Mangin, K. & Zuccarello, G. 1996. The evolution of parasitism in the red algae: molecular comparisons of adelphoparasites and their hosts. *J. Phycol.* 32:297–312.
- Goff, L.J. & Zuccarello, G.C. 1994. The evolution of parasitism in red algae - cellular interactions of adelphoparasites and their hosts. *J. Phycol.* 30:695–720.
- 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:W59–64.
- Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W. & Gascuel, O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* 59:307–21.
- Hancock, L., Goff, L. & Lane, C. 2010. Red algae lose key mitochondrial genes in response to becoming parasitic. *Genome Biol. Evol.* 2:897–910.
- Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., Von Haeseler, A. & Jermini, L.S. 2017. ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14:587–9.
- Katoh, K. & Standley, D.M. 2013. MAFFT Multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* 30:772–80.

- Krause, K. 2008. From chloroplasts to “cryptic” plastids: Evolution of plastid genomes in parasitic plants. *Curr. Genet.* 54:111–21.
- Kurihara, A., Abe, T., Tani, M. & Sherwood, A.R. 2010. Molecular phylogeny and evolution of red algal parasites: a case study of *Benzaitenia*, *Janczewskia*, and *Uhulania* (Ceramiales). *J. Phycol.* 46:580–90.
- Lagesen, K., Hallin, P., Rødland, E.A., Stærfeldt, H.H., Rognes, T. & Ussery, D.W. 2007. RNAmmer: Consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35:3100–8.
- Minh, B.Q., Nguyen, M.A.T. & Von Haeseler, A. 2013. Ultrafast approximation for phylogenetic bootstrap. *Mol. Biol. Evol.* 30:1188–95.
- Morin, L. 2000. Long branch attraction effects and the status of “basal eukaryotes”: Phylogeny and structural analysis of the ribosomal RNA gene cluster of the free-living diplomonad *Trepomonas agilis*. *J. Eukaryot. Microbiol.* 47:167–77.
- Naumann, J., Der, J.P., Wafula, E.K., Jones, S.S., Wagner, S.T., Honaas, L.A., Ralph, P.E., Bolin, J.F., Maass, E., Neinhuis, C., Wanke, S. & de Pamphilis, C.W. 2016. Detecting and characterizing the highly divergent plastid genome of the nonphotosynthetic parasitic plant *Hydnora visseri* (Hydnoraceae). *Genome Biol. Evol.* 8:345–63.
- Ng, P. K., Lim, P. E., Kato, A. & Phang, S.M. 2013. Molecular evidence confirms the parasite *Congracilaria babae* (Gracilariaceae, Rhodophyta) from Malaysia. *J. Appl. Phycol.* 26:1287–300.
- Nurk, S., Bankevich, A., Antipov, D., Gurevich, A., Korobeynikov, A., Lapidus, A., Prjibelsky, A., Pyshkin, A., Sirotkin, A., Sirotkin, Y., Stepanauskas, R., McLean, J., Lasken, R., Clingenpeel, S.R., Woyke, T., Tesler, G., Alekseyev, M.A. & Pevzner, P.A. 2013. Assembling genomes and mini-metagenomes from highly chimeric reads. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, pp. 158–70.
- Petersen, G., Cuenca, A. & Seberg, O. 2015. Plastome evolution in hemiparasitic mistletoes. *Genome Biol. Evol.* 7:2520–32.
- Preuss, M., Nelson, W.A. & Zuccarello, G.C. 2017. Red algal parasites: a synopsis of described species, their hosts, distinguishing characters and areas for continued research. *Bot. Mar.* 60:13–25.
- Preuss, M. & Zuccarello, G.C. 2014. What’s in a name? Monophyly of genera in the red algae: *Rhodophyllis parasitica* sp. nov. (Gigartinales, Rhodophyta); a new red algal parasite from New Zealand. *Algae.* 29:279–88.

- Preuss, M. & Zuccarello, G.C. 2017. Three new red algal parasites from New Zealand: *Cladhymenia oblongifoliaphila* sp. nov. (Rhodomelaceae), *Phycodrys novae-zelandiae* sp. nov. (Delesseriaceae) and *Judithia parasitica* sp. nov. (Kallymeniaceae). *Phycologia*. 57:9–19.
- Preuss, M. & Zuccarello, G.C. 2019a. Development of the red algal parasite *Vertebrata aterrimophila* sp. nov. (Rhodomelaceae, Ceramiales) from New Zealand. *Eur. J. Phycol.* 54:175–83.
- Preuss, M. & Zuccarello, G.C. 2019b. Comparative studies of photosynthetic capacity in three pigmented red algal parasites: Chlorophyll *a* concentrations and PAM fluorometry measurements. *Phycol. Res.* 67:89–93.
- Ravin, N. V., Gruzdev, E. V., Beletsky, A. V., Mazur, A.M., Prokhortchouk, E.B., Filyushin, M.A., Kochieva, E.Z., Kadnikov, V.V., Mardanoy, A.V. & Skryabin, K.G. 2016. The loss of photosynthetic pathways in the plastid and nuclear genomes of the non-photosynthetic mycoheterotrophic eudicot *Monotropa hypopitys*. *BMC Plant Biol.* 16:153–61.
- Roquet, C., Coissac, É., Cruaud, C., Boleda, M., Boyer, F., Alberti, A., Gielly, L., Taberlet, P., Thuiller, W., Van Es, J., & Lavergne, S. 2016. Understanding the evolution of holoparasitic plants: The complete plastid genome of the holoparasite *Cytinus hypocistis* (Cytinaceae). *Ann. Bot.* 118:885–96.
- Salomaki, E.D. & Lane, C.E. 2017. Red algal mitochondrial genomes are more complete than previously reported. *Genome Biol. Evol.* 9:48–63.
- Salomaki, E.D. & Lane, C.E. 2019. Molecular phylogenetics supports a clade of red algal parasites retaining native plastids: taxonomy and terminology revised. *J. Phycol.* 55:279–88.
- Salomaki, E.D., Nickles, K.R. & Lane, C.E. 2015. The ghost plastid of *Choreocolax polysiphoniae*. *J. Phycol.* 51:217–21.
- Schneider, A.C., Chun, H., Stefanović, S. & Baldwin, B.G. 2018. Punctuated plastome reduction and host–parasite horizontal gene transfer in the holoparasitic plant genus *Aphyllon*. *Proc. R. Soc. B Biol. Sci.* 285:20181535.
- Setchell, W.A. 1918. Parasitism among the red algae. *Proc. Am. Philos. Soc.* 57:155–72.
- Seward, E.A. & Kelly, S. 2016. Dietary nitrogen alters codon bias and genome composition in parasitic microorganisms. *Genome Biol.* 17:226.
- Simmons, M.P., Pickett, K.M. & Miya, M. 2004. How meaningful are bayesian support

- values? *Mol. Biol. Evol.* 21:188–99.
- Skippington, E., Barkmanb, T.J., Ricea, D.W. & Palmera, J.D. 2015. Miniaturized mitogenome of the parasitic plant *Viscum scurruloideum* is extremely divergent and dynamic and has lost all nad genes. *Proc. Natl. Acad. Sci. USA* 112:E3515–24.
- Sloan, D.B., Alverson, A.J., Chuckalovcak, J.P., Wu, M., McCauley, D.E., Palmer, J.D. & Taylor, D.R. 2012. Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. *PLoS Biol.* 10:e1001241.
- Smith, M.D., Wertheim, J.O., Weaver, S., Murrell, B., Scheffler, K. & Kosakovsky Pond, S.L. 2015. Less is more: An adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Mol. Biol. Evol.* 32:1342–53.
- Su, H. J., Barkman, T.J., Hao, W., Jones, S.S., Naumann, J., Skippington, E., Wafula, E.K., Hu, J.M., Palmer, J.D. & de Pamphilis, C.W. 2019. Novel genetic code and record-setting AT-richness in the highly reduced plastid genome of the holoparasitic plant *Balanophora*. *Proc. Natl. Acad. Sci. USA* 116:934–43.
- Suzuki, S., Endoh, R., Manabe, R.I., Ohkuma, M. & Hirakawa, Y. 2018. Multiple losses of photosynthesis and convergent reductive genome evolution in the colourless green algae *Prototheca*. *Sci. Rep.* 8:940.
- Trifinopoulos, J., Nguyen, L.T., von Haeseler, A. & Minh, B.Q. 2016. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* 44:W232–5.
- Valach, M., Burger, G., Gray, M.W. & Lang, B.F. 2014. Widespread occurrence of organelle genome-encoded 5S rRNAs including permuted molecules. *Nucleic Acids Res.* 42:13764–77.
- Veidenberg, A., Medlar, A. & Loytynoja, A. 2016. Wasabi: An integrated platform for evolutionary sequence analysis and data visualization. *Mol. Biol. Evol.* 33:1126–30.
- Verbruggen, H., Maggs, C.A., Saunders, G.W., Le Gall, L., Yoon, H.S. & De Clerck, O. 2010. Data mining approach identifies research priorities and data requirements for resolving the red algal tree of life. *BMC Evol. Biol.* 10:16.
- Wertheim, J.O., Murrell, B., Smith, M.D., Kosakovsky Pond, S.L. & Scheffler, K. 2014. RELAX: Detecting relaxed selection in a phylogenetic framework. *Mol. Biol. Evol.* 32:820–32.
- Wicke, S., Müller, K.F., de Pamphilis, C.W., Quandt, D., Wickett, N.J., Zhang, Y., Renner, S.S., Scheeweiss, G.M. 2013. Mechanisms of functional and physical genome reduction

- in photosynthetic and nonphotosynthetic parasitic plants of the broomrape family. *Plant Cell*. 25:3711–25.
- Wicke, S., Müller, K.F., DePamphilis, C.W., Quandt, D., Bellot, S. & Schneeweiss, G.M. 2016. Mechanistic model of evolutionary rate variation en route to a nonphotosynthetic lifestyle in plants. *Proc. Natl. Acad. Sci. USA* 113:9045–50.
- Wicke, S. & Naumann, J. 2018. Molecular evolution of plastid genomes in parasitic flowering plants. *Adv. Bot. Res.* 85:315–47.
- Yang, E.C., Kim, K.M., Kim, S.Y., Lee, J.M., Boo, G.H., Lee, J.H., Nelson, W.A., Yi, G., Schmidt, W.E., Fredericq, S., Boo, S.M., Bhattacharya, D. & Yoon, H.S. 2015. Highly conserved mitochondrial genomes among multicellular red algae of the Florideophyceae. *Genome Biol. Evol.* 7:2394–406.
- Zuccarello, G.C. & Lokhorst, G.M. 2005. Molecular phylogeny of the genus *Tribonema* (Xanthophyceae) using *rbcL* gene sequence data: Monophyly of morphologically simple algal species. *Phycologia* 44:384–92.
- Zuccarello, G.C., Moon, D. & Goff, L.J. 2004. A phylogenetic study of parasitic genera placed in the family Choreocolacaceae (Rhodophyta). *J. Phycol.* 40:937–45.
- Zuccarello, G.C. & West, J.A. 1994. Comparative development of the red algal parasites *Bostrychiocolax australis* gen. et sp. nov. and *Dawsoniocolax bostrychia* (Choreocolacaceae, Rhodophyta). *J. Phycol.* 30:137–46.

Table 1. Whole plastid genome characteristics in the parasite *Pterocladophila hemisphaerica* and its host *Pterocladia lucida*, plus the parasite *Choreocolax polysiphoniae* and its host *Vertebrata lanosa* (Salomaki et al. 2015) and the parasite *Harveyella mirabilis* (Salomaki and Lane 2019). *Harveyella mirabilis* was found on *Odonthalia washingtoniensis* but no further information of the plastid genomes of the host species are available. Length, A-T content, number of protein coding genes, tRNAs and rRNAs.

	cpDNA length (bp)	A-T content (%)	Protein coding genes	tRNAs	rRNAs
<i>Pterocladophila hemisphaerica</i>	68,701	74.2	70	26	2

<i>Pterocladia lucida</i>	176,635	70.3	200	30	3
<i>Choreocolax polysiphoniae</i>	90,243	79.5	71	24	3
<i>Vertebrata lanosa</i>	167,158	70.0	192	27	3
<i>Harveyella mirabilis</i>	90,654	76.5	84	23	3

Table 2. Plastid genes present in the parasite *Pterocladophila hemisphaerica*, in alphabetical order of their function. dN/dS ratios between parasite and free-living taxa; and only between free living-taxa; parasite branches under significant positive selection; and parasite genes under significant relaxation. Bold numbers indicate higher dN/dS ratio in parasite/free-living taxa comparison.

		range of dN/dS between parasite and free-living taxa	range of dN/dS between only free- living taxa
Acyl carrier protein	<i>acpP</i>	0.25-0.42	0.08-0.09
Cell division	<i>ftsH</i>	0.49-0.52	0.17-0.19
DNA replication	<i>dnaB</i>	0.69-0.87	0.51-0.81
Fatty acid biosynthesis	<i>accA</i>	0.25-0.32	0.11-0.20
	<i>accB</i>	0.44-0.59	0.13-0.30
	<i>accD</i>	0.23-0.27	0.07-0.15
	<i>fabH</i>	0.25-0.29	0.14-0.24
Glycolytic processes	<i>odpA</i>	0.23-0.32	0.11-0.20
	<i>odpB</i>	0.12-0.18	0.07-0.10
Histidyl-tRNA aminiacylation	<i>syh</i>	0.35-0.37	0.16-0.26
Iron-sulfur	<i>petF</i>	0.35-0.48	0.20-0.40

cluster transfer			
	<i>sufB</i>	0.18-0.23	0.07-0.09
	<i>sufC</i>	0.23-0.44	0.17-0.28
Metabolic processes	<i>clpC</i>	0.13-0.16	0.01-0.03
	<i>trpG</i>	0.28-0.39	0.15-0.25
Protein biosynthesis	<i>tufA</i>	0.16-0.19	0.05-0.06
Protein folding	<i>dnaK</i>	0.18-0.22	0.07-0.14
Protein transport	<i>secY</i>	0.51-0.68	0.09-0.16
Ribonuclease	<i>rnz</i>	-	-
Translation	<i>infC</i>	0.41-0.61	0.10-0.20
	<i>rpl2</i>	1.13-1.23	1.01-1.28
	<i>rpl3</i>	0.31-0.34	0.10-24
	<i>rpl4</i>	0.45-0.58	0.16-0.29
	<i>rpl5</i>	0.36-0.40	0.06-0.10
	<i>rpl6</i>	0.26-0.37	0.12-0.18
	<i>rpl11</i>	0.31-0.44	0.09-0.15
	<i>rpl12</i>	0.41-0.56	0.05-0.24
	<i>rpl13</i>	0.46-0.61	0.18-0.25
	<i>rpl14</i>	0.12-0.39	0.03-0.11
	<i>rpl16</i>	0.33-0.41	0.06-0.09
	<i>rpl18</i>	0.33-0.46	0.17-0.20
	<i>rpl19</i>	0.85-1.05	0.10-0.28
	<i>rpl20</i>	0.22-0.37	0.11-0.15
	<i>rpl21</i>	0.31	0.13-0.43
	<i>rpl23</i>	0.63-0.75	0.17-0.30
	<i>rpl27</i>	0.22-0.27	0.09-0.12
	<i>rpl31</i>	0.41-0.71	0.06-0.09
	<i>rpl32</i>	0.36-0.69	0.25-0.37
	<i>rpl33</i>	0.22-0.27	0.03-0.13
	<i>rpl35</i>	0.20-0.48	0.13-0.41
	<i>rpl36</i>	0.18-0.36	0.10-0.12

	<i>rps2</i>	0.43-0.50	0.03-0.12
	<i>rps3</i>	0.61-0.63	0.11-0.16
	<i>rps4</i>	0.26-0.38	0.05-0.11
	<i>rps5</i>	0.30-0.41	0.08-0.15
	<i>rps6</i>	0.31-0.43	0.04-0.35
	<i>rps7</i>	0.30-0.31	0.02-0.13
	<i>rps8</i>	0.35-0.44	0.08-0.11
	<i>rps9</i>	0.40-0.44	0.08-0.11
	<i>rps10</i>	0.14-0.42	0.08-0.17
	<i>rps11</i>	0.15-0.34	0.05-0.12
	<i>rps12</i>	0.09-0.13	0.05-0.08
	<i>rps13</i>	0.38-0.66	0.12-0.25
	<i>rps14</i>	0.29-0.53	0.09-0.20
	<i>rps16</i>	0.29-0.73	0.24-0.24
	<i>rps19</i>	0.05-0.20	0.07-0.15
Transcription	<i>rpoA</i>	0.34-0.44	0.04-0.10
	<i>rpoB</i>	0.23-0.29	0.06-0.14
	<i>rpoC1</i>	0.21-0.28	0.06-0.09
	<i>rpoC2</i>	0.42-0.49	0.14-0.27
Transport	<i>secA</i>	0.71-0.88	0.57-0.78
	<i>ycf38</i>	0.22-0.38	0.08-0.14
tRNA	<i>ycf62</i>	0.54-0.60	0.14-0.34
processing			
Tryptophan	<i>trpA</i>	0.33-0.42	0.16-0.34
synthase			
Uncharacterized	<i>tsf</i>	0.40-0.53	0.14-0.21
protein			
	<i>ycf19</i>	0.37-0.44	0.02-0.03
	<i>ycf21</i>	0.22-0.38	0.08-0.14

Table 3. Mitochondrial genomes of *Pterocladophila hemisphaerica* and *Pterocladia lucida*. Length; A-T content; number of protein coding genes, tRNAs and rRNAs.

	mtDNA length (bp)	A-T content (%)	Protein coding genes	tRNAs	rRNAs
<i>Pterocladophila hemisphaerica</i>	25,486	77.5%	24	24	3
<i>Pterocladia lucida</i>	25,257	70.4%	24	23	3

Table 4. Mitochondrial genes present in the parasite *Pterocladophila hemisphaerica* in alphabetical order of their function. dN/dS ratios between parasite and free-living taxa; and only between free living-taxa; parasite branches under significant positive selection; and genes under significant relaxation. Bold numbers indicate higher dN/dS ratio in parasite/free-living taxa comparison.

		range of dN/dS between parasite and free-living taxa	range of dN/dS within only free- living taxa
ATP synthesis	<i>atp4</i>	0.75-1.01	0.24-0.45
	<i>atp6</i>	0.16-0.24	0.05-0.06
	<i>atp8</i>	0.69-0.87	0.22-0.63
	<i>atp9</i>	0.01-0.02	0.00-0.01
	<i>nad1</i>	0.14-0.18	0.04-0.08
	<i>nad2</i>	0.42-0.51	0.14-0.22
	<i>nad3</i>	0.19-0.20	0.08-0.10
	<i>nad4</i>	0.20-0.27	0.06-0.13
	<i>nad4L</i>	0.25-0.39	0.05-0.12
	<i>nad5</i>	0.30-0.38	0.07-0.14
	<i>nad6</i>	0.37-0.42	0.22-0.25
Electron transport chain	<i>cob</i>	0.15-0.16	0.04-0.09
	<i>cox1</i>	0.07	0.02-0.04
	<i>cox2</i>	0.10-0.13	0.08
	<i>cox3</i>	0.16-0.24	0.04-0.07
Translation	<i>rpl16</i>	0.43-0.80	0.17-0.37

	<i>rps3</i>	0.74-0.92	0.37-0.64
	<i>rps11</i>	0.37-0.59	0.23-0.69
	<i>rps12</i>	0.17-0.24	0.09-0.18
Tricarboxylic acid cycle	<i>sdh2</i>	0.17-0.23	0.12-0.20
	<i>sdh3</i>	0.49-0.73	0.31-0.78
	<i>sdh4</i>	0.83-1.18	0.24-0.41
Uncharacterized protein	<i>tatC</i>	0.52-0.75	0.38-0.70

Fig. 1. The circular plastid genome of the parasite *Pterocladophila hemisphaerica* with protein coding genes, rRNA's and tRNA's colored by functional groups (legend). Genes are being transcribed clockwise when inside the circle and counterclockwise when outside the circle.

Fig. 2. ML topology of 67 concatenated plastid gene protein sequences inferred under the General Heterogeneous Evolution On a Single Topology (GHOST) model for the parasite *Pterocladophila hemisphaerica* (bold), its host *Pterocladia lucida* (bold) plus representatives of the Gelidiales, Ceramiales (including the red algal parasites *Choreocolax polysiphoniae*, *Harveyella mirabilis* and *Leachiella pacifica*), Gracilariales and other related taxa from Genbank (Table S1). Outgroup *Calliarthron tuberculosum* was removed to facilitate presentation. Branch support values are given as 100 bootstrap/ SH-aLRT test/ 1000 UFbootstrap. Values <85% bootstrap, <80% SH-aLRT test and <95% UFbootstrap are not shown. Asterisk represents 100% support. Scale bar indicates substitutions per site. The parasite *P. hemisphaerica* groups with support as sister to the Gracilariales.

Table S1. Sequences and whole organelle genomes, plus associated Genbank Accession numbers, used for molecular analysis of parasite, host species and representatives within the Florideophytes. Sequences and whole organelle genomes newly obtained during this study are highlighted in bold.

Table S2. Best fit models determined by ModelFinder in IQ-tree for the mitochondrial data set contain all 24 genes and the alignment where genes with elevated rates were removed. Calculated ratio of uncorrected distances between parasite *Pterocladophila hemisphaerica*

and outgroup and parasite and ingroup. Ratio in bold indicates 25% of all genes with the lowest elevated rates selected.

Table S3. Best fit models determined by ModelFinder in IQ-tree for each individual gene of the all genes alignment, and the alignment where genes with elevated rates (filtered) were removed. Calculated ratio of uncorrected distances between parasite *Pterocladophila hemisphaerica* and outgroup and parasite and ingroup. Ratio in bold indicates 25% of all genes with the lowest elevated rates selected.

Table S4. Plastid protein coding genes, tRNA and rRNA in alphabetical order by functional group with gene length in bp and A-T content in percentage in *Pterocladophila hemisphaerica* and *Pterocladia lucida*. - = missing in the plastid genome.

Table S5. Test for positive selection on the branch of *Pterocladophila hemisphaerica* in comparison to all other taxa included in the phylogeny. Number of ω tree rate classes estimates the % of branches and % of tree length are explained by either simple dN/dS models ($\omega = 1$) or show evidence of multiple dN/dS rate classes ($\omega = 2-3$). ω tree rate class are not significant different (0) or are significant different (1). Akaike information criterion fits the baseline model (MG94xREV) with a single dN/dS class per branch, and no site-to-site variation and improves parameter estimations (full model) of the adaptive rate class model.

Table S6. Test for relaxed selection of plastid genes in parasite *Pterocladophila hemisphaerica* compared to all other branches in the phylogeny. The relaxation coefficient (k) indicates if genes are under relaxed selection ($k > 1$) or intensified selection ($k < 1$). P-value in bold indicates significant differences. Likelihood-ratio test (LRT) compared the null and alternative model and Akaike information criterion measures the fit of the null model and alternative model.

Table S7. Mitochondrial protein coding genes, tRNA and rRNA in alphabetical order by functional group with gene length in bp and A-T content in percentage in *Pterocladophila hemisphaerica* and *Pterocladia lucida*.

Table S8. Test for positive selection on the branch of *Pterocladophila hemisphaerica* in comparison to all other taxa included in the phylogeny. Number of ω tree rate classes

estimates the % of branches and % of tree length are explained by either simple dN/dS models ($\omega = 1$), or show evidence of multiple dN/dS rate classes ($\omega = 2-3$). ω tree rate class are not significant different (0) or are significant different (1). Akaike information criterion fits the baseline model (MG94xREV) with a single dN/dS class per branch, and no site-to-site variation and improves parameter estimations (full model) of the adaptive rate class model.

Table S9. Test for relaxed selection of genes in the parasite *Pterocladia hemisphaerica* compared with all other taxa included in the phylogeny. The relaxation coefficient (k) indicates if genes are under relaxed selection ($k > 1$) or intensified selection ($k < 1$). P-value in bold indicates significant differences. Likelihood-ratio test (LRT) compared the null and alternative model and Akaike information criterion measures the fit of the null model and alternative model.

Fig. S1. The circular plastid genome of *Pterocladia lucida* with protein coding genes, rRNA's and tRNA's colored by functional groups (legend). Genes are being transcript clockwise when inside the circle and counterclockwise when outside the circle.

Fig. S2. Progressive Mauve alignment of 15 red algal plastid genomes including the three parasites: *Chorecolax polysiphoniae*, *Harveyella mirabilis* and *Pterocladia hemisphaerica*. Inversions and rearrangement, compared to free-living red algae, are only found in the plastid genome of *C. polysiphoniae* and *P. hemisphaerica* and not in *H. mirabilis*. In addition, *P. hemisphaerica* has two of the same inversions (indicated by a star in the parasites and highlighted by a black arrow) as found in representatives of the Gracilariales and Gelidiales.

Fig. S3. Venn diagram of shared and unique plastid genes in three red algal parasites: *Chorecolax polysiphoniae* (66 genes), *Pterocladia hemisphaerica* (70 genes), *Harveyella mirabilis* (83 genes).

Fig. S4. The circular mitochondrial genome of the parasite *Pterocladia hemisphaerica* with protein coding genes, rRNA's and tRNA's colored by functional groups (legend). Genes are being transcript clockwise when inside the circle and counterclockwise when outside the circle.

Fig. S5. The circular mitochondrial genome of the parasite *Pterocladia lucida* with protein coding genes, rRNA's and tRNA's colored by functional groups (legend). Genes are being transcript clockwise when inside the circle and counterclockwise when outside the circle.

Fig. S6. ML topology of partial *cox1* of two *Pterocladia lucida* samples (bold) infected with *Pterocladophila hemisphaerica*, plus representatives of the three cryptic species of *Pterocladia lucida*, and *Pterocladia capillacea*, from Genbank (Table S1). Outgroups *Gelidiella acerosa* and *Gelidium pacificum* were removed to facilitate presentation. Branch support values are given as bootstrap/ SH-aLRT test/ UFbootstrap. Asterisk represents 100% support. Values <85% bootstrap, <80% SH-aLRT test and <95% UFbootstrap are not shown.

Fig. S7. ML topology of concatenated LSU and SSU rDNA sequence data for the parasite *Pterocladophila hemisphaerica* and its host *Pterocladia lucida* (in bold), plus representatives of Gelidiales, Ceramiales, Gracilariales and other related taxa from Genbank (Table S1). *Jania sagittata* and *Chiharaea bodegensis* were used as outgroups. Branch support values are given as bootstrap/ SH-aLRT test/ UFbootstrap. Asterisk represents 100% support. Values <85% bootstrap, <80% SH-aLRT test and <95% UFbootstrap are not shown. Number in parenthesis behind orders indicate the number of taxa included. Scale bar indicates substitutions per site. The parasite *P. hemisphaerica* groups with strong UFbootstrap support with the Gelidiales (order of hosts) and other red algal orders, and excluding the Gracilariales, but the position of the parasite amongst those red algal orders is unsupported.

Fig. S8. ML topology of 67 concatenated plastid gene protein sequences for the parasite *Pterocladophila hemisphaerica* (bold) and its host *Pterocladia lucida* (bold) plus representatives of the Gelidiales, Ceramiales (including the red algal parasites *Choreocolax polysiphoniae*, *Harveyella mirabilis* and *Leachiella pacifica*), Gracilariales and other related taxa from Genbank (Table S1). *Calliarthron tuberculosum* was used as outgroup but removed for clarity. Branch support values are given as bootstrap/ SH-aLRT test/ UFbootstrap. One asterisk represents 100% support of all three tests. Values <85% bootstrap, <80% SH-aLRT test and <95%UFbootstrap are not shown. Scale bar indicates substitutions per site. The parasite *P. hemisphaerica* groups with support as sister to the red algal parasite *Choreocolax polysiphoniae* within the Ceramiales on an extremely long branch.

Fig. S9. ML topology of 17 concatenated plastid gene protein sequences (*acpP*, *dnaB*, *infC*, *odpB*, *rpl4*, *rpl11*, *rpl21*, *rpl23*, *rpl36*, *rpoB*, *rps12*, *rps16*, *rps19*, *syh*, *trpA*, *ycf19*, *ycf62*) for the parasite *Pterocladophila hemisphaerica* (bold) and its host *Pterocladia lucida* (bold) plus representatives of Gelidiales, Ceramiales (including the red algal parasites *Choreocolax polysiphoniae*, *Harveyella mirabilis* and *Leachiella pacifica*), Gracilariales and other related taxa from Genbank (Supplementary Table S1). *Calliarthron tuberculosum* was used as outgroup but removed for clarity. Branch support values are given as bootstrap/ SH-aLRT test/ UFbootstrap. Asterisk represents 100% support. Values <85% bootstrap, <80% SH-aLRT test and <95% UFbootstrap are not shown. Scale bar indicates substitutions per site. The parasite *P. hemisphaerica* groups in a clade with the orders Gracilariales and Gelidiales with UFbootstrap and SH-aLRT support, but the relationship of the parasite among these orders is unsupported.

Fig. S10. ML topology of 24 concatenated mitochondrial gene protein sequences for the parasite *Pterocladophila hemisphaerica* (bold) and its host *Pterocladia lucida* (bold) plus representatives of Gelidiales, Ceramiales (including the red algal parasite *Choreocolax polysiphoniae*), Gracilariales and other related taxa from Genbank (Supplementary Table S1). *Corallina officinalis* and *Calliarthron tuberculosum* were used as outgroups but removed for clarity. Branch support values are given as bootstrap/ SH-aLRT test/ UFbootstrap. One asterisk represents 100% support. Values <85% bootstrap and <80% SH-aLRT test and <95% UFbootstrap are not shown. The parasite *P. hemisphaerica* groups without support as sister to the Ceramiales.

Fig. S11. ML topology of 6 concatenated mitochondrial gene protein sequences (*atp6*, *cox1*, *cox3*, *rps3*, *rps12*, *secY*) for the parasite *Pterocladophila hemisphaerica* (bold) and its host *Pterocladia lucida* (bold) plus representatives of Gelidiales, Ceramiales (including the red algal parasites *Choreocolax polysiphoniae*), Gracilariales and other related taxa from Genbank (Table S1). *Corallina officinalis* and *Calliarthron tuberculosum* were used as outgroups but removed for clarity. Branch support values are given as bootstrap/ SH-aLRT test/ UFbootstrap. Asterisk represents 100% support. Values <85% bootstrap, <80% SH-aLRT test and <95% UFbootstrap are not shown. The parasite *P. hemisphaerica* groups without support as sister to the Plocamiales.

Fig. S12. ML topology of 24 mitochondrial genes inferred under General Heterogeneous Evolution On a Single Topology (GHOST) model for the parasite *Pterocladophila hemisphaerica* (bold) and its host *Pterocladia lucida* (bold) plus representatives of Gelidiales, Ceramiales (including the red algal parasite *Choreocolax polysiphoniae*), Gracilariales and other related taxa from Genbank (Table S1). *Corallina officinalis* and *Calliarthron tuberosum* were used as outgroups but removed for clarity. Branch support values are given as bootstrap/ SH-aLRT test/ UFbootstrap. Asterisk represents 100% support. Values <85% bootstrap, <80% SH-aLRT test and <95% UFbootstrap are not shown. The parasite *P. hemisphaerica* groups without support as sister to the Ceramiales.

jpy_12996-20-002_f1.jpg

22

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Preuss, M;Verbruggen, H;Zuccarello, GC

Title:

The Organelle Genomes in the Photosynthetic Red Algal Parasite *Pterocladophila hemisphaerica* (Florideophyceae, Rhodophyta) Have Elevated Substitution Rates and Extreme Gene Loss in the Plastid Genome

Date:

2020-08-01

Citation:

Preuss, M., Verbruggen, H. & Zuccarello, G. C. (2020). The Organelle Genomes in the Photosynthetic Red Algal Parasite *Pterocladophila hemisphaerica* (Florideophyceae, Rhodophyta) Have Elevated Substitution Rates and Extreme Gene Loss in the Plastid Genome. JOURNAL OF PHYCOLOGY, 56 (4), pp.1006-1018. <https://doi.org/10.1111/jpy.12996>.

Persistent Link:

<http://hdl.handle.net/11343/275665>