

2017

Organellar Genome Evolution in Red Algal Parasites: Differences in Adelpho- and Alloparasites

Eric Salomaki
University of Rhode Island, eric.salomaki@gmail.com

Follow this and additional works at: https://digitalcommons.uri.edu/oa_diss

Terms of Use

All rights reserved under copyright.

Recommended Citation

Salomaki, Eric, "Organellar Genome Evolution in Red Algal Parasites: Differences in Adelpho- and Alloparasites" (2017). *Open Access Dissertations*. Paper 614.
https://digitalcommons.uri.edu/oa_diss/614

This Dissertation is brought to you by the University of Rhode Island. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons-group@uri.edu. For permission to reuse copyrighted content, contact the author directly.

ORGANELLAR GENOME EVOLUTION IN RED ALGAL
PARASITES: DIFFERENCES IN ADELPHO- AND

ALLOPARASITES

BY

ERIC SALOMAKI

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

BIOLOGICAL SCIENCES

UNIVERSITY OF RHODE ISLAND

2017

DOCTOR OF PHILOSOPHY DISSERTATION

OF

ERIC SALOMAKI

APPROVED:

Dissertation Committee:

Major Professor Christopher E. Lane

Jason Kolbe

Tatiana Rynearson

Nasser H. Zawia

DEAN OF THE GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND

2017

ABSTRACT

Parasitism is a common life strategy throughout the eukaryotic tree of life. Many devastating human pathogens, including the causative agents of malaria and toxoplasmosis, have evolved from a photosynthetic ancestor. However, how an organism transitions from a photosynthetic to a parasitic life history strategy remains mostly unknown. Parasites have independently evolved dozens of times throughout the Florideophyceae (Rhodophyta), and often infect close relatives. This framework enables direct comparisons between autotrophs and parasites to investigate the early stages of parasite evolution. Parasitic red algae have traditionally been defined as either ‘adelphoparasites’, which infect hosts within their same family or tribe, or ‘alloparasites’ which infect hosts in other families. Prior to this research, investigations have primarily focused on understanding the development and evolution of ‘adelphoparasites’. All adelphoparasites studied to date have been shown to have lost their native plastid and instead incorporate a host plastid when packaging their spores. Additionally, previously published ‘adelphoparasite’ mitochondrion genomes have reduced coding capacity. The goal of this research was to investigate 1) the evolutionary impact on the plastid of the ‘alloparasite’ *Choreocolax polysiphoniae*, 2) the coding capacity of the *C. polysiphoniae* mitochondrion, and 3) the evolutionary origins of Rhodomelaceae ‘alloparasites’.

A combination of Sanger-sequencing of targeted PCR products and next-generation sequencing of genomic DNA and total RNA was used to investigate the organelles of the ‘alloparasites’ *Choreocolax polysiphoniae*, *Harveyella mirabilis*, *Leachiella pacifica*, and a previously undescribed species of *Leachiella*, the

‘adelphoparasites’ *Gracilariophila oryzoides* and *Gonimophyllum skottsbergii*, and the host of *C. polysiphoniae*, *Vertebrata lanosa*. Organellar genomes were assembled using CLC genomics workbench and Geneious Pro and subsequently manually annotated using BLAST and Pfam. Comparative analyses of organellar genomes were completed using MAUVE genome alignment software. Total RNA was assembled using the Trinity based pipeline, Agalma and annotated using InterProScan. Analyses of transcriptomic data were completed using Silix and HiFix.

This research generated plastid genomes for *C. polysiphoniae* and its host *V. lanosa*. The *C. polysiphoniae* plastid represents the first plastid genome sequenced for a red algal parasite. Interestingly, this plastid has reduced coding capacity and has lost genes involved with photosynthetic processes and its presence challenges the previously proposed paradigm of red algal parasite evolution. Investigations of red algal parasite mitochondria demonstrated that parasites retain fully functional and typical Florideophyceae mitochondria. Finally, an investigation of parasites typically considered as alloparasites supports a monophyletic clade of parasites, which all retain their native plastid genomes, arose and radiated to infect different hosts within the Rhodomelaceae. Data generated here supports previous findings that ‘alloparasites’ rarely infect hosts in different families. Therefore the terms ‘adelphoparasite’ and ‘alloparasite’, which are based on evolutionary relationships to their hosts and do not accurately distinguish types of red algal parasites. Based upon this research, we propose to redefine red algal parasites by their plastid origins as either Archaeplastid parasites (parasites that retain a native plastid) or Neoplastid parasites (those which incorporate a host plastid).

ACKNOWLEDGMENTS

I would like to thank my advisor, Chris Lane, for being a patient and supportive mentor, providing a wealth of knowledge and insight, as well as funding to enable me to complete this research. I also thank the members of my thesis committee, Dr. Jason Kolbe, and Dr. Tatiana Ryneerson for their guidance. I am grateful for the specimens, support, and numerous great conversations with my collaborator Gary Saunders. I would like to thank all the undergrads that have helped with this research, especially Katie Nickles, Danny Wolf, and Taylor Clement. I am grateful for the support of all the members of the Lane Lab including Nicolas Blouin, Thea Popolizio, Chris Paight, Jillian Freese, Daniela Lopes Paim Pinto, and Joselynn Wallace, you are hereby an honorary member too. Kristina Terpis, I cannot thank you enough for all your assistance, support and friendship. Funding is gratefully acknowledged from the following entities: NSF DEB #1257472, and the Phycological Society of America. Finally, thanks to Mary and David Salomaki, and Amy Moore for all your love, support, and encouragement.

PREFACE

This thesis is presented in manuscript format. Chapters 1 through 3 have been published in peer-reviewed journals and chapters 4 and 5 are in preparation for submission to peer-reviewed journals. Chapter one, “Are all red algal parasites cut from the same cloth?”, was published in *Acta Societatis Botanicorum Poloniae*. Chapter two, “The ghost plastid of *Choreocolax polysiphoniae*”, was published as a letter in *Journal of Phycology*. Chapter three, “Red algal mitochondrion genomes are more complete than previously reported”, was published in *Genome Biology and Evolution*. Chapter four, “Molecular analysis of parasites in the Choreocolacaceae (Rhodophyta) reveals a reduced *Harveyella mirabilis* (Reinsch) F. Schmitz et Reinke plastid genome and supports the transfer of genera to the Rhodomelaceae (Rhodophyta)”, is in preparation for submission to *The Journal of Eukaryotic Microbiology*. Chapter five, “Molecular data provides support to update terminology for distinguishing red algal parasites”, is in preparation for submission to *Journal of Phycology* as an opinion paper.

TABLE OF CONTENTS

ABSTRACT	ii
ACKNOWLEDGMENTS	iv
PREFACE	v
TABLE OF CONTENTS	vi
LIST OF TABLES	vii
LIST OF FIGURES	ix
CHAPTER 1	1
ARE ALL RED ALGAL PARASITES CUT FROM THE SAME CLOTH	1
CHAPTER 2	31
THE GHOST PLASTID OF <i>CHOREOCOLAX POLYSIPHONIAE</i>	31
CHAPTER 3	55
RED ALGAL MITOCHONDRIAL GENOMES ARE MORE COMPLETE THAN PREVIOUSLY REPORTED	55
CHAPTER 4	102
MOLECULAR ANALYSIS OF PARASITES IN THE CHOREOCOLACACEAE (RHODOPHYTA) REVEALS A REDUCED <i>HARVEYELLA MIRABILIS</i> (REINSCH) F. SCHMITZ ET REINKE PLASTID GENOME AND SUPPORTS THE TRANSFER OF GENERA TO THE RHODOMELACEAE (RHODOPHYTA)	102
CHAPTER 5	134
MOLECULAR DATA PROVIDES SUPPORT TO UPDATE TERMINOLOGY FOR DISTINGUISHING RED ALGAL PARASITES.....	134

LIST OF TABLES

TABLE	PAGE
Chapter 1	
Table 1. Taxonomic summary of red algal parasites including number of described species and type of parasite by Family. The shaded boxes are adelphoparasite numbers, highlighted to indicate their abundance, relative to alloparasites	19
Chapter 3	
Table 1. Table 1. Table of all currently available Florideophyceae mitochondrion genomes that were examined in this study with GenBank Accession, genome length, and A/T%. Genes previously reported as missing are listed along with notes regarding their status as a result of this study	82
Table 2. Current status of Florideophyceae mitochondrial genes previously reported as missing in Hancock <i>et al.</i> (2010) and Yang <i>et al.</i> (2015) or otherwise unannotated.	87
Table 3. A/T%, non-synonymous to synonymous mutation (dN/dS) ratio and individual dN and dS values for genes encoded on the Florideophyceae mitochondrion genomes. Species with a pseudogene, rather than a functional copy of the gene, are listed in the far right column and were left out of calculations of A/T% and dN/dS ratio	89
Table 4. The use of alternative start codons by gene based on published literature, and the proximity to the closest ATG start codon. Alternative initiation codons that are supported by the lack of a nearby ATG initiation codon and maintained gene start location as observed by alignment are indicated in bold	91

TABLE	PAGE
Chapter 4	
Table 1. GenBank Sequence accession numbers for taxa used in phylogenetic analyses (Figure 2)	128
Table 2. <i>Harveyella mirabilis</i> plastid genes and their function, which have been lost from the <i>Choreocolax polysiphoniae</i> plastid genome.....	129

LIST OF FIGURES

FIGURE	PAGE
Chapter 1	
Figure 1. Red algal parasites often evolve from a recent common ancestor with their host. These “adelphoparasites” (purple branch) are usually host-specific. However, parasites that have diversified into more than one species and/or infect distantly-related hosts, are known as “alloparasites” (blue branches)..	20
Figure 2. Infection strategies of red algal alloparasites and adelphoparasites. A) Alloparasites penetrate the host thallus and grow a network of filamentous cells into the host. Each cell is capable of fusing to a host cell via secondary pit connection and depositing its contents. Once inside the cell, the alloparasite nucleus does not divide, but causes the host nucleus to enlarge, or in the case of cortical cells, multiply. It is currently unclear whether the alloparasite plastid is derived from the host or parasite. B) Adelphoparasite spores fuse with a cortical cell and inject their contents. These parasite nuclei can multiply within the host, however no nuclear DNA synthesis has been observed by the host after infection. Parasite nuclei and organelles spread via the host primary pit connections. The host organelles also multiply in response to infection	21
Chapter 2	
Figure 1. The plastid genome of the parasitic red alga, <i>Choreocolax polysiphoniae</i> is 90,243 basepairs and contains 71 coding genes (colored by protein function), the 5S, 16S, and 23S rRNAs (Red), and 24 tRNAs (Pink). All genes involved with	

photosynthetic functions, except *petF*, have been lost. The *ftsH* gene is truncated but may still be transcribed, however *gltB* is a non-functional pseudogene (Yellow). If the plastid is a linear molecule, the breakpoint is indicated by the orange zig-zag line. The Venn diagram in the middle shows the plastid genome content of the free living red alga, *Vertebrata lanosa* (red box), and its obligate parasite *C. polysiphoniae* (black box with grey shading). Additional boxes represent the green algal facultative parasite, *Prototheca wickerhamii* (dark green box), the green algal obligate parasite, *Helicosporidium* sp. (light green box), and a composite Apicoplast (blue box). Genes are color coded by protein function. 43

Figure 2. Mauve alignment of *Choreocolax polysiphoniae* (top) and *Vertebrata lanosa* (second from top) and other published Florideophyceae plastid genomes. This alignment identifies 13 locally collinear blocks (LCBs) among all Florideophycean plastid genomes while highlighting the reduction and fragmentation evident in *C. polysiphoniae*. 45

Chapter 3

Figure 1. Translated alignment of *sdhD* genes from florideophycean mitochondria showing *Asparagopsis taxiformis*, *Ceramium japonicum* and *Ceramium sungminbooi* (top three sequences) share critical conserved residues with all other Florideophyceae *sdhD* genes. 93

Figure 2. Alignment of the original *Ceramium japonicum nad3* gene with the modified *Ceramium japonicum nad3* ('T' deleted from base 36; red box) and copies of the *nad3* gene from *Chondrus crispus*, *Gracilaria vermiculophylla*, *Gracilariopsis*

FIGURE	PAGE
<i>andersonii</i> , <i>Sporolithon durum</i> , <i>Choreocolax polysiphoniae</i> , and <i>Vertebrata lanosa</i> . Manual deletion of one ‘T’ from the string of 26 ‘T’s and 3 ‘C’s between 32-60 bp from the start codon restores conservation of the length and sequence of the <i>Ceramium japonicum nad3</i> gene. Genes are shown with the amino acid translation below.....	94

Chapter 4

Figure 1. Maximum likelihood phylogeny based on concatenated 18S, 28S, COI-5P, and *rbcL* data as shown in Table 1, of the parasites *Choreocolax polysiphoniae*, *Harveyella mirabilis*, *Leachiella pacifica*, an undescribed species of *Leachiella* found on *Polysiphonia paniculata*, and thirteen close relatives. This phylogenetic analysis supports a monophyletic clade of parasites arising within the Rhodomelaceae. Support values shown as Bayesian posterior probability/maximum likelihood bootstrap.....

Figure 2. The plastid genome of the parasitic red alga, *Harveyella mirabilis* is 90,654 basepairs and contains 84 protein coding genes (Green), the 5S, 16S, and 23S rRNAs (Red), and 23 tRNAs (Pink). All genes involved with photosynthetic functions, except *petF*, have been lost. The *ftsH* gene is truncated but may still be transcribed, however *gltB* is a non-functional pseudogene (Yellow)

Figure 3. Mauve alignment of the parasites *Choreocolax polysiphoniae* (top) and *Harveyella mirabilis* (second from top) with all published Rhodomelaceae plastid genomes as well as those from select representatives of other Florideophyceae families (*Dasya binghamiae*, Dasyaceae; *Chondrus crispus*, Gigartinaceae; *Calliarthron tuberculosum*, Corallinaceae; *Grateloupia taiwanensis*, Halymeniaceae; *Gracilaria tenuistipitata*, Gracilariaceae). This alignment identifies 16 locally collinear blocks

FIGURE	PAGE
<p>(LCBs) among the selected plastid genomes and demonstrates that even with the loss of photosynthesis genes overall synteny is shared between the parasite <i>H. mirabilis</i> and other plastid genomes, while <i>C. polysiphoniae</i> has undergone several genome rearrangements</p>	132
<p>Figure 4. Mauve alignment of the plastid genomes from the parasites <i>Harveyella mirabilis</i> (top) and <i>Choreocolax polysiphoniae</i> (bottom). This alignment identifies 11 locally collinear blocks (LCBs) among these parasite plastid genomes, highlighting the high level of genome fragmentation and rearrangements evident in the two parasite plastids</p>	133
<p>Chapter 5</p>	
<p>Figure 1. Mauve alignment of the parasites <i>Harveyella mirabilis</i> (top) and <i>Choreocolax polysiphoniae</i> (bottom) with all published Rhodomelaceae plastid genomes. This alignment identifies 12 locally collinear blocks (LCBs) among the selected plastid genomes and demonstrates that even with the loss of photosynthesis genes overall synteny is shared between the parasite <i>H. mirabilis</i> and other plastid genomes, while <i>C. polysiphoniae</i> has undergone several genome rearrangements..</p>	150
<p>Figure 2. Comparison of carotenoid biosynthesis related genes between the red algal parasite <i>Choreocolax polysiphoniae</i> and <i>Gracilariophila oryzoides</i>.....</p>	151

CHAPTER 1

Are all red algal parasites cut from the same cloth?

by

Eric D. Salomaki¹ and Christopher E. Lane¹

is published in *Acta Societatis Botanicorum Poloniae*

¹Department of Biological Sciences, University of Rhode Island, Kingston, RI 02881.

Abstract

Parasitism is a common life strategy throughout the eukaryotic tree of life. Many devastating human pathogens, including the causative agents of malaria and toxoplasmosis, have evolved from a photosynthetic ancestor. However, how an organism transitions from a photosynthetic to a parasitic life history strategy remains mostly unknown. This is largely because few systems present the opportunity to make meaningful comparisons between a parasite and a close free-living relative. Parasites have independently evolved dozens of times throughout the Florideophyceae (Rhodophyta), and often infect close relatives. The accepted evolutionary paradigm proposes that red algal parasites arise by first infecting a close relative and over time diversify and infect more distantly related species. This provides a natural evolutionary gradient of relationships between hosts and parasites that share a photosynthetic common ancestor. Elegant microscopic work in the late 20th century provided detailed insight into the infection cycle of red algal parasite and the cellular interactions between parasites and their hosts. Those studies led to the use of molecular work to further investigate the origins of the parasite organelles and reveal the evolutionary relationships between hosts and their parasites. Here we synthesize the research detailing the infection methods and cellular interactions between red algal parasites and their hosts. We offer an alternative hypothesis to the current dogma of red algal parasite evolution and propose that multiple possible evolutionary pathways for which red algae can adopt a parasitic life history strategy exist, including direct infection of distant relatives. Furthermore, we highlight potential directions for future research to further evaluate parasite evolution in red algae.

Introduction

Parasitism has evolved innumerable times throughout the eukaryotic tree of life (Lafferty et al. 2008). Some of the more virulent parasites have transitioned from a once photosynthetic ancestor, including the causative agents of malaria and related mammalian diseases (Wilson et al. 1996, Kohler et al. 1997). Therefore, understanding the evolutionary trajectory between photosynthesis and abandoning autotrophy for a parasitic strategy, is of particular importance. Red algal parasites are uniquely valuable to study this path because they have independently evolved many times, providing literally dozens of discrete events to compare (Goff 1982, Goff et al. 1996, 1997, Zuccarello et al. 2004). This system may provide novel insights into the evolution of parasitism, especially with regard to the early stages of transitioning from a photosynthetic past.

Red algal parasites exclusively infect other red algae, typically ones with which they share a recent common ancestor (Goff 1982, Goff and Coleman 1985, Goff et al. 1997, Zuccarello et al. 2004). The relationship between host and parasite was first recognized using morphological similarities in the life-cycles of parasites and their hosts (Setchell 1918). More recently, molecular data have confirmed this hypothesis (Goff et al. 1996, 1997, Zuccarello et al. 2004, Kurihara et al. 2010). Traditionally red algal parasites have been placed into two different groups, based on their phylogenetic relationships with their hosts (Goff et al. 1997). Adelphoparasites (adelpho is Greek for “kin”) are closely related to their host and often infect only one host, whereas alloparasites are more divergent from their host(s) (Goff et al. 1996,

1997). Currently, adelphoparasites are believed to make up roughly 90% of all red algal parasites (Goff et al. 1997).

Among the Florideophyceae, parasites belonging to at least 66 different red algal genera have evolved independently over 100 times (Table 1) (Verbruggen et al. 2010). The accepted evolutionary paradigm proposes that adelphoparasitism is the initial state, followed by parasite diversification, which leads to the development of alloparasites (Fig. 1). These “older” parasites can infect more distantly related taxa, and make up roughly 10% of red algal parasites (Goff 1982, Zuccarello et al. 2004). Due to their rarity, alloparasites are relatively unstudied, with the exception of *Choreocolax polysiphoniae*.

The Importance Of Red Algal Pit Connections

One of the defining characteristics of the florideophycean red algae is the ability of cells to form connections with their adjacent cells (Pueschel and Cole 1982). There are two distinct forms of these “pit-connections” formed by red algae. Primary pit-connections arise between a mother and daughter cell during apical growth (Ramus 1969). These connections result from a seemingly incomplete cell division where the septum begins to develop from the cell walls growing inward to separate the daughter nuclei (Ramus 1969). However, cytokinesis is incomplete and the septum does not fuse, leaving an opening that connects the two cells (Ramus 1969). A pit-plug composed of a polysaccharide-protein complex then forms sealing the pit connection and separating the two cells (Ramus 1969, 1971). Though pit connections result in an

aperture that is not entirely sealed by a septum, the pit plug prevents the transfer of cellular contents and photosynthate between adjacent cells (Turner and Evans 1978).

In addition to the primary pit connections, florideophytes also form secondary pit connections between adjacent non-daughter cells. These secondary pit connections are known to occur in a wide range of Florideophyceae and form between two genetically similar red algal cells (Hawkins 1972, Goff and Coleman 1985). Red algal parasite spores utilize secondary pit connections as a way to enter the cells of the host (Wetherbee and Quirk 1982a, Wetherbee et al. 1984, Goff and Coleman 1985). As evidence for the importance of secondary pit connections in parasitic infections, parasites are not known from red algal orders where secondary pit connections do not occur (Blouin and Lane 2012). An advantage of this strategy is that the similarity between host and parasite at the genetic level allows parasite spores to simply dump their organelles into the host cell and take over, spreading through primary or secondary connections (Wetherbee and Quirk 1982b, Goff and Coleman 1985). The widespread existence of secondary pit connections among red algae is undoubtedly a primary factor in the promiscuous nature of parasitism as a life history strategy in the lineage.

Differences in Adelphoparasite and Alloparasite Infection Cycles

Spore Germination and Host Infection

Rhodophytes lack flagella in all stages of their life cycle, making the initial stages of locating a host a passive process. Once a parasite spore lands upon a susceptible host, the parasite carospore (2N) or tetraspore (1N) will germinate and

undergo an initial cell division (Goff and Zuccarello 1994). Adelphoparasite cells will divide between 1 and 3 more times before one of the cells forms a rhizoid that penetrates the surface of the host, growing into the wall of a host epidermal cell (Goff and Zuccarello 1994). The tip of the rhizoid swells isolating a single parasite nucleus along with its organelles into a conjunctor cell which divides from the infection rhizoid (Goff and Zuccarello 1994, Goff and Coleman 1995). This conjunctor cell then fuses, via a secondary pit connection, with the adjacent host cell (Fig. 2a). The contents of the conjunctor cell, which include the parasite nucleus and organelles, are deposited into the host cell, thus forming a heterokaryotic cell (containing both parasite and host nuclei) (Wetherbee and Quirk 1982a, Goff and Coleman 1985, Goff and Zuccarello 1994). This connection between the parasite infection rhizoid and the transformed host cell is sealed by a pit-plug that was formed previously during the initial fusion between the conjunctor cell and the parasite infection rhizoid (Goff and Coleman 1985, Goff and Zuccarello 1994).

The differences between adelpho- and alloparasites become evident at the initial infection of the host cell. In alloparasites, rather than going through a few cell divisions before penetrating the surface, the parasite spore attaches to a suitable host, penetrating and forming a hyphae-like network of multicellular filaments between the host cells (Goff and Coleman 1984a, 1985, 1987). These filaments enable the parasite to spread numerous cells deep into the host away from the initial site of infection (Fig. 2b). Each alloparasite cell in these filaments contains a single nuclei and can form a conjunctor cell and secondary pit connection through which the parasite deposits its

cellular contents into the host, creating a heterokaryotic cell (Goff and Coleman 1984a, 1985, 1987).

Inside the Heterokaryon

After the host cell becomes heterokaryotic, both adelpho- and alloparasites take control of the host cellular machinery (Goff and Coleman 1985, Goff and Zuccarello 1994). Almost immediately upon infection, the hosts' central vacuole tonoplast is lysed, allowing cytoplasm to spread throughout the space previously occupied by the vacuole (Goff 1976, Goff et al. 1996). Subsequently, the number of organelles including plastids, mitochondria, and ribosomes increases throughout the cytoplasm causing it to appear more dense (Goff 1976, Goff and Coleman 1984a, 1985, Goff and Zuccarello 1994). Along with the increase in cytoplasmic organelles comes an increase in cell size (hypertrophy), a process in which the cell can grow to 40 times its original size (Kugrens and West 1973, Goff and Coleman 1985, 1995). In addition to the increased organelles, the host nuclei also increase in size and/or number (Goff and Coleman 1984a, 1985).

Goff and Coleman used 4',6-diamidino-2-phenylindole (DAPI) staining and microspectrofluorometry to examine the interactions between the alloparasite, *Choreocolax polysiphoniae* and its host *Vertebrata lanosa* (Goff and Coleman 1985). Their study showed that infected *V. lanosa* central cells will become enlarged and the nuclei will undergo DNA synthesis but not nuclear division resulting in polyploid host nuclei (Goff and Coleman 1985). Alternatively, in infected *V. lanosa* pericentral cells, the host nuclei will either increase in size, increase in number, or some combination of

both (Goff and Coleman 1985). An increase in the number of host nuclei is the most common response (Goff and Coleman 1985). Host cells that are adjacent to infected cells and connected by pit connections will not show any cytological transformation (Goff and Coleman 1984a, 1985).

Conversely, no DNA synthesis or nuclear division has been observed from the host nucleus after infection by the adelphoparasite *Janczewskia gardneri* (Goff and Coleman 1987). Instead, the adelphoparasite nucleus rapidly undergoes DNA synthesis, generating numerous parasite nuclei inside a single host cell (Fig. 2) (Goff and Coleman 1987). The adelphoparasite subsequently spreads to additional host cells through the formation of conjuctor cells that can infect adjacent host cells (Goff and Zuccarello 1994). Adelphoparasites can also form rhizoidal infection cells, which are multinucleate and contain large numbers of mitochondria, ribosomes, and dedifferentiated host-derived proplastids and can fuse with more distant host cells (Goff and Zuccarello 1994).

Alloparasites are capable of mitotic divisions to create the multicellular filaments that spread between host cells (Goff and Coleman 1984a, 1987). However, the alloparasite nucleus does not undergo DNA synthesis inside the host cell (Goff and Coleman 1984a, 1985, 1987). Therefore, parasite nuclei remain at a 1:1 ratio with the number of secondary pit connections between parasite and host cells (Goff and Coleman 1984a, 1985, 1987). Once inside a heterokaryon the adelphoparasite utilizes the host to progress through its lifecycle and reproduce rather than continue to spread to additional host cells.

Formation of Reproductive Structures

As the adelphoparasite spreads throughout the host, a gall or “erumpent pustule” begins to form as host cells continually expand upon infection by the parasite. Eventually the adelphoparasite will start to form reproductive structures (Goff and Zuccarello 1994). If the original infecting spore was haploid the parasite will form carposporangia that can be fertilized by a spermatia from another parasite forming a diploid carposporophyte that will eventually release carpospores (Goff and Coleman 1984b, Goff and Zuccarello 1994). If the original infection was from a diploid carpospores, the parasite will undergo meiosis forming haploid tetraspores that will be released from the erumpent pustule (Goff and Coleman 1984b, Goff and Zuccarello 1994).

An allopasite does not spread through the host like the adelphoparasite. Rather than forming a gall from invaded cells, a host pericentral cell containing a parasite nuclei and many host nuclei will form a protuberance (Goff and Coleman 1985). This protuberance will become isolated from the original host cell and undergo mitotic divisions, which produces the mature parasite pustule containing reproductive cells similar to that of adelphoparasites (Goff and Coleman 1985). After the reproductive cells are released, the former host tissue that made the erumpent pustule becomes necrotic (Goff and Coleman 1985). While the fate of future adelphoparasites lies in the released spores, allopasites are able to continue their infection of the same host as parasite filaments continue to grow into uninfected areas (Goff and Coleman 1985).

Organelles

Studies of cellular organelles have yielded particularly interesting findings during investigations of red algal parasite biology. Early studies established the role of secondary pit connections between the parasite and host cells and demonstrated their role in transferring the parasite nucleus to the host cell (Wetherbee and Quirk 1982a, Goff and Coleman 1985, Goff and Zuccarello 1994). However, it was unclear whether the parasite maintained its own mitochondrion and plastid or if it utilized the host organelles once the parasite nucleus was transferred into the host cell. Cytoplasmic organelles support many major metabolic pathways as well as play major roles in cellular energy and carbohydrate production. It has been well established that purifying selection is relaxed on parasite organellar genes that become unnecessary, leading to genome reduction in parasites as they increasingly rely on a host for energy and carbohydrates (Wolfe et al. 1992, Cai et al. 2003, Borza et al. 2005, de Koning and Keeling 2006, Vaidya and Mather 2009). Therefore, it seems likely that some red algal parasite mitochondrion and plastid genes would be truncated or even lost over time. The origin and roles of mitochondria and plastids in the parasite-host interaction may reveal key information regarding red algal parasite biology and their ability to infect and control the host cellular machinery.

First using the alloparasite *Choreocolax polysiphoniae*, and later the adelphoparasites *Gracilariophila oryzoides* and *Gardneriella tuberifera*, researchers noticed that in addition to the parasite nuclei, organelles are also transferred to the host cell via the conjuctor cell upon infection (Goff and Coleman 1987, Goff and Zuccarello 1994). Once molecular tools became more widely available, Goff and

Coleman utilized restriction fragment length polymorphisms (RFLPs) to investigate the origin of mitochondria in the adelphoparasites *Plocamicolax pulvinata*, *Gracilariophila oryzoides*, and *Gardneriella tuberifera* and their respective hosts, *Plocamium cartilagineum*, *Gracilariopsis lemaneiformis*, and *Sarcodiotheca gaudichaudii* (Goff and Coleman 1995). This work revealed that *P. pulvinata* and *G. oryzoides* maintain a genetically unique mitochondrion and that both the parasite and host mitochondria are present within the heterokaryotic cells (Goff and Coleman 1995). The study was unable to conclusively demonstrate that the mitochondrion from *Gardneriella tuberifera* is unique from that of *Sarcodiotheca gaudichaudii*, due to the extremely close relationship between the two species (Goff and Coleman 1995).

The mitochondrion genomes of the adelphoparasites *Plocamicolax pulvinata* and *Gracilariophila oryzoides*, as well as its host *Gracilariopsis andersonii*, were recently sequenced. These data paved the way for a new level of fine-scaled investigations of red algal parasites, elucidating details of the organellar genome architecture that was previously unattainable (Hancock et al. 2010). When comparing the mitochondrion genome sequences of the parasites with the free living host, the *atp8* and *sdhC* genes from *G. oryzoides* were determined to be pseudogenes (Hancock et al. 2010). Furthermore, *atp8* was determined to be absent from *P. pulvinata*. However, the authors noted that according to sequenced cDNA libraries, the genes were still transcribed (Hancock et al. 2010). Recent second generation sequencing of additional samples of these species has revealed that the “missing” genes are all present and lack frameshift mutations (Salomaki & Lane, unpublished). These data suggest that purifying selection is maintained on red algal parasite mitochondria and

that even though much of the parasite life cycle exists inside a host cell, red algal parasites still require their own mitochondrion for their survival.

While evident that parasites maintain their native mitochondrion, microscopy and molecular studies have demonstrated that red algal parasites do not maintain their own plastid (Goff and Zuccarello 1994, Goff and Coleman 1995). Microscopy shows that the spores of adelphoparasites *Gracilariophila oryzoides*, and *Gardneriella tuberifera* contain proplastids lacking photosynthetic pigments, phycobilisomes, and thylakoids (Goff and Zuccarello 1994). Once the parasite injects its nuclei and organelles into a host cell, the host plastids transform and their light harvesting phycobilisomes disappear from the thylakoids (Goff and Zuccarello 1994). After rapid dedifferentiation, simple proplastids that are similar to the infecting parasite plastid bud off the host plastid (Goff and Zuccarello 1994). As the parasite nuclei, mitochondrion, and host derived proplastids spread to adjacent cells through pit connections, plastids from the newly infected host cells also rapidly dedifferentiate into proplastids (Goff and Zuccarello 1994). Eventually cells emerge from the heterokaryotic host cell containing only parasite nuclei, parasite mitochondrion and a host-derived proplastid (Goff and Zuccarello 1994). RFLP analysis was utilized to investigate whether this plastid was a genetically unique parasite plastid, or instead, the parasite was incorporating a host-derived proplastid (Goff and Coleman 1995). This study revealed that adelphoparasites *Plocamiocolax pulvinata*, *Gracilariophila oryzoides*, and *Gardneriella tuberifera* and their hosts had identical banding patterns (Goff and Coleman 1995). Subsequent DNA sequencing of the variable plastid *rbcL-rbcS* spacer region revealed that the plastid from both hosts and parasites were

genetically identical, confirming that the parasite plastid is a dedifferentiated host plastid (Goff and Coleman 1995).

Interestingly, two studies suggest that the proplastid may be capable of differentiating again into a photosynthetic plastid. While investigating the photosynthetic rates and C^{14} transfer from *Polysiphonia* (= *Vertebrata*) *lanosa* to the alloparasite *Choreocolax polysiphoniae*, Callow *et al.* examined the alloparasite pustules after dissection from their host (Callow *et al.* 1979). The authors note that many of the parasite pustules had a pinkish hue and incorporated radioactively labeled C^{14} into their thallus (Callow *et al.* 1979). Furthermore, the carbon fixation rate increased over time (up to 66 hours) leading the authors to conclude that *C. polysiphoniae* is capable of photosynthesis on its own. However, the source of photosynthetic activity in dissected *C. polysiphoniae* pustules may be from host cells that have been incorporated into the pustule as observed in that study, and independently, by Kugrens and West, and Goff examining *Janczewskia gardneri* (Kugrens and West 1973, Callow *et al.* 1979, Goff and Zuccarello 1994). Additionally, the status of *J. gardneri* as a parasite or obligate epiphyte has been debated due to its pigmentation (Setchell 1914, Court 1980, Goff and Coleman 1987). Most recently, it was noted that during the early stages of the interaction between *J. gardneri* and its host, *Laurencia spectabilis*, *J. gardneri* exists as colorless cells and ‘infects’ host cells in the same manner as other adelphoparasites. As *J. gardneri* cells erupt from the host they remain colorless but the cells become pigmented once the adelphoparasite becomes reproductively mature (Goff and Coleman 1987). Whether this pigmentation

originates from host cells in the pustule matrix, or if the proplastid differentiates back into a photosynthetic plastid remains unknown.

Nutrient Transfer

With the exception of a few adelphoparasites that gain pigmentation upon reproductive maturity, red algal parasites are not capable of photosynthesis on their own and must obtain carbohydrates and other nutrients from a host. After parasite infection, the host (now heterokaryotic) cell loses the ability to photosynthesize as a result of plastid dedifferentiation (Goff and Zuccarello 1994). This leads to a differential gradient of carbon between the heterokaryotic cell and the adjacent normally functioning host cells (Goff 1976, Goff and Zuccarello 1994, Goff and Coleman 1995). To account for the loss of carbon fixation, uninfected host cells direct photosynthate to heterokaryon and parasite cells that they are connected to via pit connections (Goff 1979, Goff and Coleman 1987).

The first studies investigating carbon transfer between a red algal host and its parasite found three products of photosynthesis (floridoside, isofloridoside, and manitol) were transferred from the host to its parasite via a concentration gradient (Evans et al. 1973, Callow et al. 1979). Later five different sugar species were identified to be assimilated by the host *Rhodomela confervoides* and translocated to its parasite *Harveyella mirabilis* (Kremer 1983). Investigations into carbon translocation in *H. mirabilis* demonstrated the localization of carbon, from being fixed by the photosynthetic host through its movement into the parasite cells and revealed that heterokaryon cells incorporated more C^{14} than neighboring uninfected host cells (Goff

1979). Furthermore, it was determined that starch was not distributed evenly throughout the parasite cells as might be expected, but instead was being directed preferentially to parasite reproductive cells (Goff 1979, Goff and Coleman 1985). Given the capabilities of parasites for obtaining carbon from the host, the role of the maintained proplastid in parasite cells remains in question.

Host Specificity and Parasite Resistance

Red algal parasites are known to be extremely host specific, usually infecting one to a few, closely related host species (Goff 1982, Zuccarello and West 1994a, 1994b). A study using the adelphoparasite *Janczewskia morimotoi* tested its ability to infect 15 other species including close relatives of its natural host, *Laurencia nipponica*, as well as members of different genera (Nonomura and West 1981). While *J. morimotoi* was capable of infecting two close relatives of its natural host, the more distantly related potential hosts prevented parasite infections (Nonomura and West 1981). Additionally, the host specificity of *Leachiella pacifica* was assessed through culture studies attempting to use parasites isolated from *Polysiphonia paniculata* to infect *Pterocladia bipinnata* and vice-versa (Zuccarello and West 1994b). Parasites isolated from *P. paniculata* could infect other populations of the same species as well as some other *Polysiphonia* species, however they could not infect *Pt. bipinnata* populations that were susceptible to parasites isolated from other *Pt. bipinnata* specimens (Zuccarello and West 1994b). These *L. pacifica* isolates showed strong genus-level host specificity. However, due to the greatly reduced morphology of red algal parasites, it cannot be ruled out that parasites isolated from different genera are,

in fact, different host-specific species. Revisiting this study with molecular data would strengthen our understanding of host specificity and potentially reveal cryptic parasite species.

Dawsoniocolax bostrychiae and *Bostrychiocolax australis* are parasites that infect *Bostrychia radicans* (Zuccarello and West 1994a). A study on host range and specificity of these parasites on a variety of potential hosts, yielded similar results to the *J. morimotoi* study: the genetic distance between parasite and host has a strong negative correlation with susceptibility to parasite penetration and infection (Zuccarello and West 1994a). The authors note that they encountered hosts that are resistant to parasite infection, including some host populations that contained resistant and susceptible specimens (Zuccarello and West 1994a). In several cases the parasite was capable of forming an initial infection in a resistant host, however the host cell or cells adjacent to the infected cell died off, preventing the parasite from spreading further into the host (Zuccarello and West 1994a). However, subsequent molecular studies revealed phenotypic plasticity and cryptic diversity in *B. radicans* (Zuccarello and West 2003). Therefore, the possibility remains that resistant and susceptible hosts from the host resistance study were actually different species. These findings emphasize the need for ongoing taxonomic evaluation of red algal parasites and their hosts. Without the taxonomic framework questions about whether or not the host is actually resisting parasite infection cannot be answered conclusively.

Many Questions Remain

Why does the parasite maintain a copy of the host plastid as it is forming its own reproductive cells and spores? Other parasites that have evolved from a plastid bearing ancestor, including the apicomplexans *Eimeria tenella* and *Plasmodium falciparum*, the parasitic plant *Epifagus virginiana* and many others, maintain a reduced plastid for cellular functions other than photosynthesis, such as fatty acid biosynthesis (Wolfe et al. 1992, Wilson et al. 1996, Cai et al. 2003). However, none of these plastid-bearing parasites steal a plastid from their host like the red algal adelphoparasites. Are adelphoparasites genetically similar enough to their hosts that they can target nuclear-encoded proteins to the host-derived proplastid and utilize those products for fatty acid biosynthesis? Genomic analyses of signaling and targeting peptides for plastid targeted nuclear genes in red algal parasites, combined with transcriptomic and proteomic approaches, will provide valuable insight into the role of the plastids in the infection mechanism and parasite life cycle. The use of additional molecular tools including *in-situ* hybridization would enable researchers to localize parasite nuclear-encoded proteins in the heterokaryotic cell.

Furthermore, the taxonomic range and multiple independent origins of red algal parasites makes it difficult to make generalizations based on a few observations. Thus far, the origin of red algal parasite plastids has only been investigated in adelphoparasites and the origin of the alloparasite plastid remains unknown. The assumption is that alloparasites first progress through an adelphoparasite stage and also maintain a co-opted host plastid. However, there are distinct developmental differences between adelpho- and alloparasites, including the initial steps of infection and alloparasites inability to synthesize DNA in heterokaryon cells (Fig. 2). Therefore,

it seems plausible to propose an alternative hypothesis that, rather than passing through an adelphoparasite stage, alloparasites are capable of directly evolving infection mechanisms to parasitize distantly related hosts. In the proposed scenario, alloparasites would presumably maintain their own plastid, as they are likely incapable of utilizing such a genetically distant host plastid. Preliminary genomic data from *C. polysiphoniae* indicates this may, in fact, be the case (Salomaki & Lane, unpublished). Future research investigating red algal parasite evolution will provide unique insight into the effects of transitioning from a free-living to a parasitic life strategy. Molecular data has supported morphological observations that red algal parasites share a recent common ancestor with their hosts (Goff et al. 1996, 1997, Zuccarello et al. 2004, Kurihara et al. 2010). However, further use of molecular tools is essential to provide a robust taxonomic framework of red algal parasites and their hosts. Only then can meaningful observations be made about host specificity and parasite resistance. With the technological advances of the past few decades and continually decreasing costs of DNA sequencing, information about the relationships between parasites and their hosts, unraveling the roles of parasite and host interactions, and the origins and function of organelles is within our grasp.

Table 1. Taxonomic summary of red algal parasites including number of described species and type of parasite by Family. The shaded boxes are adelphoparasite numbers, highlighted to indicate their abundance, relative to alloparasites.

Order	Family	Type and Number of Parasites
Ceramiales	Ceramiaceae	Adelphoparasites = 1
		Alloparasites = 2
	Delesseriaceae	Adelphoparasites = 14
	Rhodomelaceae	Adelphoparasites = 37
		Alloparasites = 8
Spyridiaceae	Adelphoparasites = 1	
Wrangeliaceae	Adelphoparasites = 1	
Corallinales	Corallinaceae	Adelphoparasites = 2
		Alloparasites = 1
Gigartinales	Hapalidiaceae	Adelphoparasites = 13
		Alloparasites = 1
	Cystocloniaceae	Adelphoparasites = 2
	Kallymeniaceae	Adelphoparasites = 3
	Phyllophoraceae	Adelphoparasites = 1
Solieriaceae	Adelphoparasites = 1	
Gracilariales	Gracilariaceae	Adelphoparasites = 2
	Pterocladophilaceae	Adelphoparasites = 1
		Alloparasites = 12
Halymeniales	Halymeniaceae	Adelphoparasites = 1
Palmariales	Palmariaceae	Adelphoparasites = 1
	Rhodophysemataceae	Adelphoparasites = 1
Plocamiales	Plocamiaceae	Adelphoparasites = 2
Rhodymeniales	Faucheaceae	Adelphoparasites = 2
	Rhodymeniaceae	Adelphoparasites = 2
Incertae sedis	Incertae sedis	Adelphoparasites = 1
		Alloparasites = 6

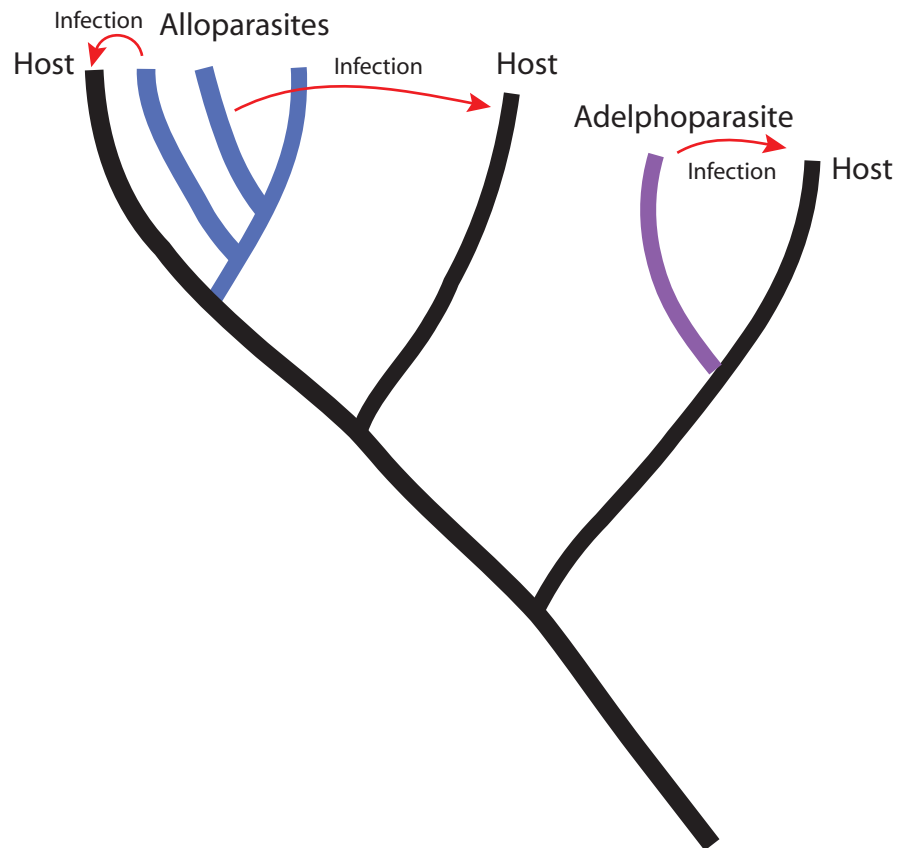


Figure 1: Red algal parasites often evolve from a recent common ancestor with their host. These “adelphoparasites” (purple branch) are usually host-specific. However, parasites that have diversified into more than one species and/or infect distantly-related hosts, are known as “alloparasites” (blue branches).

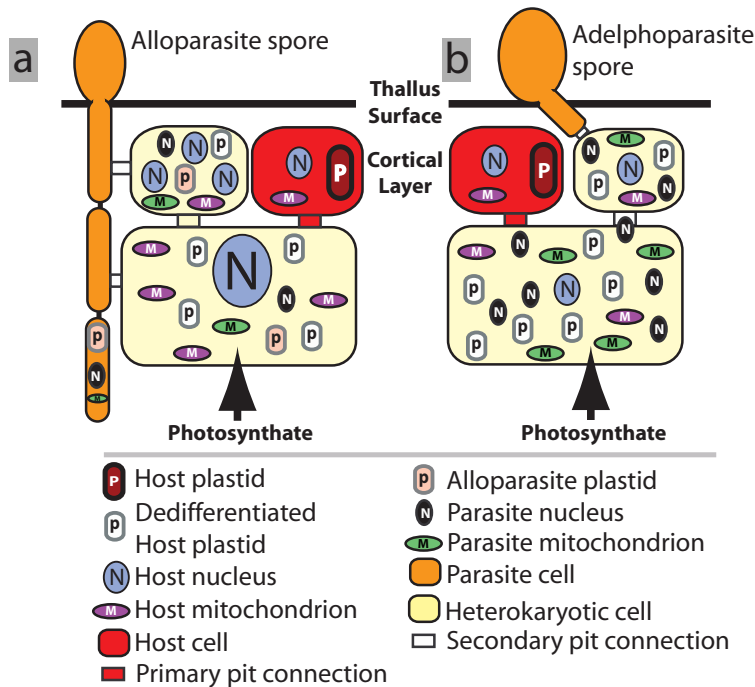


Figure 2: Infection strategies of red algal alloparasites and adelphoparasites. A)

Alloparasites penetrate the host thallus and grow a network of filamentous cells into the host. Each cell is capable of fusing to a host cell via secondary pit connection and depositing its contents. Once inside the cell, the alloparasite nucleus does not divide, but causes the host nucleus to enlarge, or in the case of cortical cells, multiply. It is currently unclear whether the alloparasite plastid is derived from the host or parasite.

B) Adelphoparasite spores fuse with a cortical cell and inject their contents. These parasite nuclei can multiply within the host, however no nuclear DNA synthesis has been observed by the host after infection. Parasite nuclei and organelles spread via the host primary pit connections. The host organelles also multiply in response to infection.

References

- Bendich, A.J. 2004. Circular Chloroplast Chromosomes : The Grand Illusion. *Plant Cell*. 16:1661–6.
- Bendich, A.J. & Smith, S.B. 1990. Moving pictures and pulsed-field gel electrophoresis show linear DNA molecules from chloroplasts and mitochondria. *Curr. Genet.* 17:421–5.
- Blouin, N.A. & Lane, C.E. 2012. Red algal parasites: models for a life history evolution that leaves photosynthesis behind again and again. *Bioessays*. 34:226–35.
- Borza, T., Popescu, C.E. & Lee, R.W. 2005. Multiple metabolic roles for the nonphotosynthetic plastid of the green alga *Prototheca wickerhamii*. *Eukaryot. Cell*. 4:253–61.
- Cai, X.M., Fuller, A.L., McDougald, L.R. & Zhu, G. 2003. Apicoplast genome of the coccidian *Eimeria tenella*. *Gene*. 321:39–46.
- Callow, J.A., Callow, M.E. & Evans, L. V 1979. Nutritional studies on the parasitic red alga *Choreocolax polysiphoniae*. *New Phytol.* 83:451–62.
- Court, G.J. 1980. Photosynthesis and Translocation Studies of *Laurencia spectabilis* and its Symbiont *Janczewskia gardneri* (Rhodophyceae). *J. Phycol.* 16:270–9.
- 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.

- de Koning, A.P. & Keeling, P.J. 2006. The complete plastid genome sequence of the parasitic green alga *Helicosporidium* sp. is highly reduced and structured. *BMC Biol.* 4.
- DePriest, M.S., Bhattacharya, D. & López-Bautista, J.M. 2013. The plastid genome of the red macroalga *Grateloupia taiwanensis* (Halymeniaceae). *PLoS One.* 8:e68246.
- Douzery, E.J.P., Snell, E. a, Baptiste, E., Delsuc, F. & Philippe, H. 2004. The timing of eukaryotic evolution: does a relaxed molecular clock reconcile proteins and fossils? *Proc. Natl. Acad. Sci. U. S. A.* 101:15386–91.
- Evans, L. V, Callow, J.A. & Callow, M.E. 1973. Structural and physiological studies on the parasitic red alga *Holmsella*. *New Phytol.* 72:393–402.
- Fast, N.M., Kissinger, J.C., Roos, D.S. & Keeling, P.J. 2001. Nuclear-encoded, plastid-targeted genes suggest a single common origin for apicomplexan and dinoflagellate plastids. *Mol. Biol. Evol.* 18:418–26.
- Feldmann, G. & Feldmann, J. 1958. Recherches sur quelques Floridees parasites. *Rev. Gen. Bot.* 65:49–128.
- Freshwater, D.W., Fredericq, S., Butler, B.S., Hommersand, M.H. & Chase, M.W. 1994. A gene phylogeny of the red algae (Rhodophyta) based on plastid *rbcL*. *Proc. Natl. Acad. Sci. U. S. A.* 91:7281–5.
- Gardner, M.J., Bishop, R., Shah, T., de Villiers, E.P., Carlton, J.M., Hall, N., Ren, Q. et al. 2005. Genome sequence of *Theileria parva*, a bovine pathogen that transforms lymphocytes. *Science.* 309:134–7.

- Goff, L.J. 1976. The Biology of *Harveyella mirabilis* (Cryptonemiales, Rhodophyceae). V. Host Responses to Parasite Infection. *J. Phycol.* 12:313–28.
- Goff, L.J. 1979. The Biology of *Harveyella mirabilis* (Cryptonemiales, Rhodophyceae). VI. Translocation of Photoassimilated ¹⁴C. *J. Phycol.* 15:82–7.
- Goff, L.J. 1982. The biology of parasitic red algae. *Prog. Phycol. Res.* 1:289–369.
- Goff, L.J., Ashen, J. & Moon, D. 1997. The evolution of parasites from their hosts: A case study in the parasitic red alga. *Evolution (N. Y.)* 51:1068–78.
- Goff, L.J. & Coleman, A.W. 1984a. Transfer of nuclei from a parasite to host. *Proc. Natl. Acad. Sci.* 81:5420–4.
- Goff, L.J. & Coleman, A.W. 1984b. Elucidation of fertilization and development in a red alga by quantitative DNA microspectrofluorometry. *Dev. Biol.* 102:173–94.
- 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. 1987. Nuclear transfer from parasite to host: A new regulatory mechanism of parasitism. *Ann. N. Y. Acad. Sci.* 503:402–23.
- 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.

- Hancock, L., Goff, L.J. & Lane, C.E. 2010. Red algae lose key mitochondrial genes in response to becoming parasitic. *Genome Biol. Evol.* 2:897–910.
- Happe, T. & Naber, J.D. 1993. Isolation, characterization and N-terminal amino acid sequence of hydrogenase from the green alga *Chlamydomonas reinhardtii*. *Eur. J. Biochem.* 214:475–81.
- Hawkins, E.K. 1972. Observations on the Developmental Morphology and Fine Structure of Pit Connections in Red Algae. *Cytologia (Tokyo)*. 37:759–68.
- Jacobs, J., Pudollek, S., Hemschemeier, A. & Happe, T. 2009. A novel, anaerobically induced ferredoxin in *Chlamydomonas reinhardtii*. *FEBS Lett.* 583:325–9.
- Janouškovec, J., Liu, S.-L., Martone, P.T., Carré, W., Leblanc, C., Collén, J. & Keeling, P.J. 2013. Evolution of red algal plastid genomes: ancient architectures, introns, horizontal gene transfer, and taxonomic utility of plastid markers. *PLoS One*. 8:e59001.
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S. et al. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 28:1647–9.
- Knauf, U. & Hachtel, W. 2002. The genes encoding subunits of ATP synthase are conserved in the reduced plastid genome of the heterotrophic alga *Prototheca wickerhamii*. *Mol. Genet. Genomics*. 267:492–7.
- Kohler, S., Delwiche, C.F., Denny, P.W., Tilney, L.G., Webster, P., Wilson, R.J.M., Palmer, J.D. et al. 1997. A Plastid of Probable Green Algal Origin in

- Apicomplexan Parasites. *Science* (80-). 275:1485–9.
- Kremer, B.P. 1983. Carbon economy and nutrition of the alloparasitic red alga *Harveyella mirabilis*. *Mar. Biol.* 239:231–9.
- Kugrens, P. & West, J.A. 1973. The ultrastructure of an alloparasitic red alga *Choreocolax polysiphoniae*. *Phycologia*. 12:175–86.
- Kuo, C.-H., Wares, J.P. & Kissinger, J.C. 2008. The Apicomplexan whole-genome phylogeny: an analysis of incongruence among gene trees. *Mol. Biol. Evol.* 25:2689–98.
- 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 *Ululania* (Ceramiales). *J. Phycol.* 46:580–90.
- Lafferty, K.D., Allesina, S., Arim, M., Briggs, C.J., De Leo, G., Dobson, A.P., Dunne, J.A. et al. 2008. Parasites in food webs: the ultimate missing links. *Ecol. Lett.* 11:533–46.
- Lim, L., Linka, M., Mullin, K. a, Weber, A.P.M. & McFadden, G.I. 2010. The carbon and energy sources of the non-photosynthetic plastid in the malaria parasite. *FEBS Lett.* 584:549–54.
- Lim, L. & McFadden, G.I. 2010. The evolution, metabolism and functions of the apicoplast. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 365:749–63.
- Nonomura, A.M. & West, J.A. 1981. Host specificity of *Janczewskia* (Ceramiales, Rhodophyta). *Phycologia*. 20:251–8.
- Pueschel, C.M. & Cole, K.M. 1982. Rhodophycean Pit Plugs: An Ultrastructural

- Survey with Taxonomic Implications. *Am. J. Bot.* 69:703–20.
- Ramus, J. 1969. Pit Connection Formation in the Red Alga *Pseudogloiophloea*. *J. Phycol.* 5:57–63.
- Ramus, J. 1971. Properties of septal plugs from the red alga *Griffithsia pacifica*. *Phycologia*. 10:99–103.
- Reinsch, P.F. 1875. *Contributiones ad Algologiam et Fungologiam*. Lipsiae, Nuremberg. 61 pp.
- Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B. et al. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61:539–42.
- Salomaki, E.D. & Lane, C.E. 2014. Are all red algal parasites cut from the same cloth? *Acta Soc. Bot. Pol.* 83:369–75.
- Saunders, G.W. 1993. Gel purification of red algal genomic DNA: An inexpensive and rapid method for the isolation of polymerase chain reaction-friendly DNA. *J. Phycol.* 29:251–4.
- Saunders, G.W. & Moore, T.E. 2013. Refinements for the amplification and sequencing of red algal DNA barcode and RedToL phylogenetic markers: a summary of current primers, profiles and strategies. *Algae*. 28:31–43.
- Schattner, P., Brooks, A.N. & Lowe, T.M. 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res.* 33:686–9.
- Setchell, W.A. 1914. Parasitic Florideae I. *Univ. Calif. Publ. Bot.* 6:1–34.

- Setchell, W.A. 1918. Parasitism among red algae. *Proc. Am. Philosophical Soc.* 57:155–72.
- Sherwood, A.R., Kurihara, A., Conklin, K.Y., Sauvage, T. & Presting, G.G. 2010. The Hawaiian Rhodophyta Biodiversity Survey (2006-2010): a summary of principal findings. *BMC Plant Biol.* 10:258.
- Sherwood, A.R. & Presting, G.G. 2007. Universal primers amplify a 23S rDNA plastid marker in eukaryotic algae and cyanobacteria. *J. Phycol.* 43:605–8.
- Stamatakis, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics.* 22:2688–90.
- Tamayo, D.A. & Hughey, J.R. 2016. Organellar genome analysis of the marine red alga *Dasya binghamiae* (Dasyaceae, Rhodophyta) reveals an uncharacteristic florideophyte mitogenome structure. *Mitochondrial DNA Part B.* 1:510–1.
- Turner, C.H.C. & Evans, L.V. 1978. Translocation of photoassimilated ¹⁴C in the red alga *Polysiphonia lanosa*. *Br. Phycol. J.* 13:51–5.
- Vaidya, A.B. & Mather, M.W. 2009. Mitochondrial evolution and functions in malaria parasites. *Annu. Rev. Microbiol.* 63:249–67.
- 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.
- Viola, R., Nyvall, P. & Pedersén, M. 2001. The unique features of starch metabolism in red algae. *Proc. Biol. Sci.* 268:1417–22.
- Wetherbee, R. & Quirk, H. 1982a. The Fine Structure of Secondary Pit Connection

- Formation Between the Red Algal Alloparasite *Holmsella australis* and Its Red Algal Host *Gracilaria furcellata*. *Protoplasma*. 176:166–76.
- Wetherbee, R. & Quirk, H. 1982b. The Fine Structure and Cytology of the Association Between the Parasitic Red Alga *Holmsella australis* and Its Red Algal Host *Gracilaria furcellata*. *Protoplasma*. 165:153–65.
- Wetherbee, R., Quirk, H.M., Mallett, J.E. & Ricker, R.W. 1984. The Structure and Formation of Host-Parasite Pit Connections Between the Red Algal Alloparasite *Harveyella mirabilis* and Its Red Algal Host *Odonthalia floccosa*. *Protoplasma*. 73:62–73.
- Wilson, R.J.M.I., Denny, P.W., Preiser, D.J., Rangachari, K., Roberts, K., Roy, A., Whyte, A. et al. 1996. Complete gene map of the plastid-like DNA of the malaria parasite *Plasmodium falciparum*. *J. Mol. Biol.* 261:155–72.
- Wolfe, K.H., Morden, C.W. & Palmer, J.D. 1992. Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. *Proc. Natl. Acad. Sci. U. S. A.* 89:10648–52.
- Yoon, H.S., Muller, K.M., Sheath, R.G., Ott, F.D. & Bhattacharya, D. 2006. Defining the major lineages of red algae (Rhodophyta). *J. Phycol.* 42:482–92.
- Zhu, G., Marchewka, M.J. & Keithly, J.S. 2002. *Cryptosporidium parvum* appears to lack a plastid genome. *Microbiology*. 146:315–21.
- 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. 1994a. Host Specificity In The Red Algal Parasites *Bostrichiocolax australis* And *Dawsoniocolax bostrychia* (Choreocolacaceae, Rhodophyta). *J. Phycol.* 30:462–73.
- Zuccarello, G.C. & West, J.A. 1994b. Genus and race specificity in the red algal parasite *Leachiella pacifica* (Choreocolacaceae, Rhodophyta). *Phycologia.* 33:213–8.
- Zuccarello, G.C. & West, J. a. 2003. Multiple Cryptic Species: Molecular Diversity and Reproductive Isolation in the *Bostrychia Radicans*/*B. Moritziana* Complex (Rhodomelaceae, Rhodophyta) With Focus on North American Isolates. *J. Phycol.* 39:948–59.

CHAPTER 2

The ghost plastid of *Choreocolax polysiphoniae*

by

Eric D. Salomaki¹, Katie R. Nickles¹, and Christopher E. Lane¹

is published in *Journal of Phycology*

¹Department of Biological Sciences, University of Rhode Island, Kingston, RI 02881.

Abstract

Parasitism has evolved innumerable times among eukaryotes. Red algal parasites alone have independently evolved over 100 times. The accepted evolutionary paradigm proposes that red algal parasites arise by first infecting a close relative and over time diversifying and infecting more distantly related species. This provides a natural evolutionary gradient of relationships between hosts and parasites that share a photosynthetic common ancestor. Upon infection, the parasite deposits its organelles into the host cell and takes over, spreading through cell-cell connections. Microscopy and molecular studies have demonstrated that the parasites do not maintain their own plastid, but rather abscond with a dedifferentiated host plastid as they pack up spores for dispersal. However, we sequenced a ~90 kb plastid genome from the parasite *Choreocolax polysiphoniae* and the 167 kb plastid genome of its host, *Vertebrata lanosa*. Sequencing a native red algal parasite plastid indicates additional evolutionary pathways to parasitism exist among red algae.

Red Algal Parasite Evolution

With over 100 extant species of parasitic red algae (Blouin and Lane 2012, Salomaki and Lane 2014), most of which evolved independently, red algae are a spectacular group to investigate the evolutionary mechanism by which a species transitions from free-living to parasitic. Early after the discovery of red algal parasites, phycologists postulated that parasites arise sympatrically and infect the species with which they share their most recent common ancestry (Setchell 1918). Subsequent molecular studies have provided additional support for this hypothesis (Goff and Coleman 1987, Goff et al. 1996, Zuccarello et al. 2004, Kurihara et al. 2010). These red algal parasites take advantage of the close relationship with their host in their unique infection mechanism.

Elegant studies by Goff and colleagues provided the fundamental understanding of red algal parasite biology and how they interact with their hosts. Upon infection, red algal parasites fuse with a host cell and deposit their cellular contents (Goff and Coleman 1984a, 1985, 1987, 1995, Goff and Zuccarello 1994). The resulting heterokaryotic host cell rapidly increases carbohydrate production and starch formation, becoming enlarged (Goff and Coleman 1987). The parasite eventually directs the host to form spores, which will be released to start the cycle again (Goff and Coleman 1984b, Goff and Zuccarello 1994). Interestingly, all red algal parasites studied, including specimens from three different orders, have lost their own plastid. Instead of maintaining their native plastid, they incorporate a dedifferentiated host plastid into their spores (Goff and Coleman 1995).

Until now, only the plastids of red algal adelphoparasites (adelpho is Greek for “kin”) that share a recent common ancestor with their host have been examined. Here we describe the plastid of *Choreocolax polysiphoniae* Reinsch, a parasite that is evolutionarily distant from its host (termed an alloparasite), and that of its host *Vertebrata lanosa* (Linnaeus) T.A. Christensen. These data reveal the parasite-host plastid interactions may be quite different in alloparasites. The presence of a native non-photosynthetic plastid in *C. polysiphoniae*, indicates that multiple evolutionary pathways to parasitism exist among red algae.

Descriptions of the Plastids

Vertebrata lanosa is a multicellular polysiphonious red alga that belongs to the Rhodomelaceae (Supplemental Figure 1), and like other red algae with published plastid sequences, is a canonical photosynthetic red alga. At 167 kilobases (kb) in length, the *V. lanosa* plastid genome (GenBank Accession KP308097) is slightly smaller than those previously sequenced from florideophytes, which range from 178 kb (*Calliarthron tuberculosum*) to 191 kb (*Grateloupia taiwanensis*) (DePriest et al. 2013, Janouškovec et al. 2013). The *V. lanosa* plastid encodes 193 protein coding genes, the 16S, 23S, and 5S rRNAs, 27 tRNAs (Supplemental Figure 1), and is similar to other florideophyte plastid genomes in gene content and arrangement (Supplemental Figure 2).

Choreocolax polysiphoniae is also a member of the Rhodomelaceae, however it is an obligate parasite of *V. lanosa*, which appears as a multicellular unpigmented erumpent pustule growing from the *V. lanosa* thallus (Supplementary Figure 1). In the

course of gathering genome-scale data from *C. polysiphoniae*, a 90 kb contig with high coverage (815x) emerged from the data. Closer examination revealed that the contig represents a highly reduced plastid genome sequence (GenBank Accession KP308096) in *C. polysiphoniae* (Figure 1). The plastid encodes 71 protein coding genes, 3 rRNAs, and 24 tRNAs. All 71 protein-coding genes are shared with the *V. lanosa* plastid. However *petF* is the only photosynthesis-related protein encoded by the *C. polysiphoniae* plastid. In *C. polysiphoniae*, as in *Chlamydomonas reinhardtii*, it is likely to serve as an electron carrier in additional metabolic pathways (Happe and Naber 1993, Jacobs et al. 2009). Although the *C. polysiphoniae* plastid is no longer capable of light harvesting and photosynthesis, genes involved in amino acid, fatty acid, isoprene, and protein biosynthesis, transcription and translation as well as other cellular maintenance, are conserved.

The *C. polysiphoniae* *ftsH* gene, which is involved in photosystem II repair, is missing the first ~150 residues, however the conservation of the remaining 452 amino acid residues does not indicate a loss of selection pressure on the gene. Conversely, only the remnants of *gltB* are detectable by BLAST. Unlike the conservation observed in *ftsH*, the *gltB* region of the plastid genome contains no substantial ORFs and is laden with stop codons, indicative of a vanishing pseudogene. BLAST searches against the preliminary assembly of *C. polysiphoniae* genomic data was used to locate plastid genes found in the *V. lanosa* plastid genome, but absent from *C. polysiphoniae*. These searches were performed to account for the possibility of transfer from the plastid to the nucleus. None of the ‘missing’ plastid genes could be found in the

assembly of nuclear data, however low coverage of these data (<10x) does not rule out transfer as a possibility.

Comparisons to Other Florideophyceae Plastids

Typical plastid genes used for Rhodophyte phylogenies, including the large subunit of RuBisCo (*rbcL*) and the photosystem I & II genes *psaA* and *psbA* (Freshwater et al. 1994, Yoon et al. 2006, Verbruggen et al. 2010, Saunders and Moore 2013), have been lost in *C. polysiphoniae*. However, the UPA marker provided meaningful comparative data for phylogenetics (Sherwood and Presting 2007, Sherwood et al. 2010). Despite the expected low levels of statistical support for deep branches in the tree when using such a variable gene region, the phylogeny supports the placement of the *C. polysiphoniae* and *V. lanosa* plastids among the Rhodomelaceae (Supplemental Figure 3).

Aligning the plastid genomes of *V. lanosa* and *C. polysiphoniae* with other published Florideophyceae plastids demonstrates that Florideophyte plastids are generally conserved (Figure 2). Even though the *V. lanosa* plastid is ~13 kb smaller than other Florideophyte plastids there are very few genes that are missing (Supplementary Table 1). The *C. polysiphoniae* plastid maintains blocks of synteny with other Florideophytes, but it is clear that the *C. polysiphoniae* plastid has undergone many losses and genome rearrangements (Figure 2). This is likely a result of relaxed selection due to the parasite relying on the host for light harvesting and carbohydrate production.

Despite the annular representation of the *C. polysiphoniae* plastid in Figure 1, we have thus far been unable to unequivocally demonstrate that the genome is a circular molecule like most other plastid genomes. Reports of linear plastids have been made from some flowering plants (Bendich and Smith 1990, Bendich 2004). While the read mapping data indicates there are reads spanning from one end of the assembly to the other, multiple attempts to PCR across the ‘gap’ have been unsuccessful. Additionally, two of the three independent *C. polysiphoniae* data sets (and combined datasets) assembled the plastid with the same beginning and end. The third dataset assembled the plastid in two separate contigs though this was a result of lack of overlapping coverage between the *tufA* and *rps12* genes and the ends of those contigs corresponded to the same start and end of the previous assemblies (see Supplemental Materials for additional information). If the molecule were circular, a random “breakpoint” in the assembly would be expected. This similarity between the beginning and end to the sequence suggests either there is a distinct region resistant to DNA amplification and sequencing, or that it may actually be a linear molecule. Future work using restriction digests and pulse-gel electrophoresis will help to confirm the annular or linear nature of the *C. polysiphoniae* plastid.

Comparisons to Other Parasite Plastids

Most “crown” Apicomplexa are parasites that contain a relict plastid (apicoplast) obtained through an ancient secondary endosymbiosis (Fast et al. 2001, Zhu et al. 2002), illustrating plastid genome reduction in a lineage that lost its photosynthetic abilities long ago (Douzery et al. 2004, Kuo et al. 2008, Lim et al.

2010). The apicoplast genomes from *Plasmodium falciparum*, *Theileria parva*, and *Eimeria* are more reduced than *C. polysiphoniae* (all three are ~35 kb), but maintain function in fatty acid biosynthesis, heme biosynthesis, iron-sulfur cluster synthesis, and isoprenoid biosynthesis (Figure 1) (Wilson et al. 1996, Cai et al. 2003, Gardner et al. 2005, Lim and McFadden 2010).

Non-photosynthetic plastids from parasites in the Viridiplantae lineage have also been sequenced (Wolfe et al. 1992, Knauf and Hachtel 2002, Borza et al. 2005, de Koning and Keeling 2006). The *Helicosporidium* sp. plastid genome has similar coding capacity to *C. polysiphoniae*, as both maintain genes involved with amino acid and fatty acid synthesis and transcriptional processes (Figure 1) (de Koning and Keeling 2006). The *Epifagus virginiana* plastid-encoded functions are more limited, lacking the ability to perform amino acid synthesis (Wolfe et al. 1992). Unlike all other parasites discussed, *Prototheca wickerhamii* is a facultative parasite and its plastid (55.6 kb, GenBank accession KJ001761) has maintained genes for additional processes including starch synthesis (which is not a role of red algal plastids (Viola et al. 2001)), carbohydrate metabolism, and *de novo* purine synthesis (Figure 1) (Knauf and Hachtel 2002, Borza et al. 2005). The *C. polysiphoniae* plastid demonstrates that similar patterns of degradation occur in plastids as species transition from a photosynthetic life strategy to parasitism.

Implications for Red Algal Parasite Evolution

The description of a plastid in *C. polysiphoniae* suggests there are multiple mechanisms for parasite evolution in Rhodophytes. Current dogma suggests that red

algal parasites first evolve sympatrically with their host (Goff and Coleman 1987, Goff et al. 1996, Zuccarello et al. 2004, Kurihara et al. 2010). However, there are several differences in the cellular interactions between adelpho- and alloparasites and their respective hosts (Salomaki and Lane 2014). Furthermore, the non-photosynthetic plastid of the alloparasite, *C. polysiphoniae*, indicates this is not the sole evolutionary pathway. If *C. polysiphoniae* had passed through an adelphoparasite stage, previous research suggests it would no longer maintain a plastid (Goff and Coleman 1987, Goff et al. 1996, Zuccarello et al. 2004, Kurihara et al. 2010). Conversely, if parasites can directly infect distantly related hosts, the remaining non-photosynthetic parasite plastid implicates an alternative to recognized organelle evolution in these parasites. Adelphoparasites and alloparasites rely on different infection mechanisms, based on the relationship with their host (Goff and Coleman 1987, 1995, Goff and Zuccarello 1994). The lack of a native plastid in all adelphoparasites examined to date suggests that the genetic similarity between adelphoparasites and their hosts allows the parasite to utilize the host plastid rather than maintaining its own. Clearly it is necessary for some red algal parasites to maintain their own plastid for certain functions while relying on the host for carbohydrate production. Perhaps future studies targeted at alloparasites will determine a link between phylogenetic relatedness and ability to utilize the host plastid.

Methods and Materials

Specimens of both *Vertebrata lanosa* and *Choreocolax polysiphoniae* were collected at Beavertail State Park, Jamestown, RI. The *V. lanosa* was inspected for parasites and epiphytes under a dissecting microscope and clean material was ground under liquid nitrogen. Individual *C. polysiphoniae* were excised from the thallus of *V. lanosa* and collected in a 1.5 mL microcentrifuge tube. Specimens were hand-ground using a Corning Axygen® PES-15-B-SI disposable tissue grinder pestle in a 1.5 mL microcentrifuge tube, submerged in 100µL of DNA extraction buffer (Saunders 1993). DNA was extracted from two batches containing ~50 individual *C. polysiphoniae* using a standard phenol/chloroform extraction (Saunders 1993). Additionally, an individual *C. polysiphoniae* was extracted using only 100µL of buffer and the same volume for the phenol/chloroform steps. Whole-genome amplification of the individual parasite was completed using the illustra GenomiPhi Hy DNA Amplification Kit (GE Healthcare Life Sciences, Pittsburgh, Pa) according to manufacturer protocols. The *cox1* and 18S genes were PCR amplified from the whole-genome amplified *C. polysiphoniae* DNA to screen for contamination. The resulting DNA from five separate amplifications of the same *C. polysiphoniae* extraction was combined to reduce amplification biases that may have resulted during the whole-genome amplification. Libraries for Illumina sequencing of the amplified sample and the two batch extraction samples were prepared on the Apollo 324 robot using the PrepX ILM DNA Library Kit (Wafergen Biosystems, Fremont, Ca). Two libraries, one from the *C. polysiphoniae* genome amplification and one from one of the batch *C. polysiphoniae* extractions were each sequenced on separate full-cell Illumina MiSeq paired-end 300 x 300 basepair runs. The libraries for the *V. lanosa* DNA and the

second batch *C. polysiphoniae* extraction were multiplexed on a third Illumina MiSeq paired-end 300 x 300 basepair run. Each library was assembled independently using the CLC Genomics Workbench (CLC Bio, Boston, Ma).

All *C. polysiphoniae* assemblies were aligned and found to be similar with few indels and single nucleotide polymorphisms between them. The assembly from the amplified library was chosen as the representative plastid genome sequence for *C. polysiphoniae* due to the quality and depth of coverage. Open reading frame (ORF) prediction on the *V. lanosa* and *C. polysiphoniae* plastid sequences was done using Geneious Pro v6.1 (Kearse et al. 2012) and the resulting ORFs were manually annotated using GenBank and UniProt databases. Genes from other red algal plastid genomes that were absent from the *V. lanosa* and *C. polysiphoniae* plastids were BLAST searched against the plastid sequences and the genomic assemblies to verify their absence and check for evidence of transfer to another genetic compartment. The assembled plastid genome sequences were submitted to the tRNAscan-SE online server v1.21 for identification of tRNA sequences (Schattner et al. 2005). Attempts to ‘close the gap’ by using PCR to sequence from one end of the *C. polysiphoniae* plastid genome was attempted unsuccessfully using 12 different primer pairs.

The universal plastid amplicon (UPA) (Sherwood and Presting 2007) sequences were extracted from the *V. lanosa* and *C. polysiphoniae* plastid genome assemblies and aligned with Rhodomelaceae and Ceramiaceae UPA downloaded from GenBank (accessed Aug 1, 2014). The aligned UPA dataset was subjected to Bayesian inference (BI) analysis in MrBayes v3.2 (Ronquist et al. 2012) using two Metropolis-coupled Markov chain Monte Carlo (MCMCMC) runs consisting of three

hot chains and one cold chain. Each run was sampled every 100 generations for 2,020,000 generations. After the average standard deviation of split frequencies was below 0.01 to confirm that the runs converged, the first 200 trees were removed as a burn-in. The resulting tree and posterior probabilities were generated from the remaining trees generated for the dataset. Maximum likelihood (ML) analysis was conducted using RAxML (Stamatakis 2006) with the same model parameters that were used for the BI analysis. ML bootstrap support values were calculated using 500 bootstrap replicates.

To compare the arrangement and structure of the *V. lanosa* and *C. polysiphoniae* genomes with other Florideophyceae plastids, the published plastid genomes of *Calliarthron tuberculosum* (KC153978), *Chondrus crispus* (HF562234), *Gracilaria salicornia* (KF861575), *Gracilaria tenuistipitata* (NC006137), and *Grateloupia taiwanensis* (KC894740), were downloaded from GenBank. All circular plastid genomes were arranged to be linear, starting at the same point to aid in visualization, since this was the beginning of the *C. polysiphoniae* plastid sequence. The plastid genomes were aligned using the default settings for the progressiveMauve algorithm in the Mauve v2.3.1 (Darling et al. 2004) Geneious plugin.

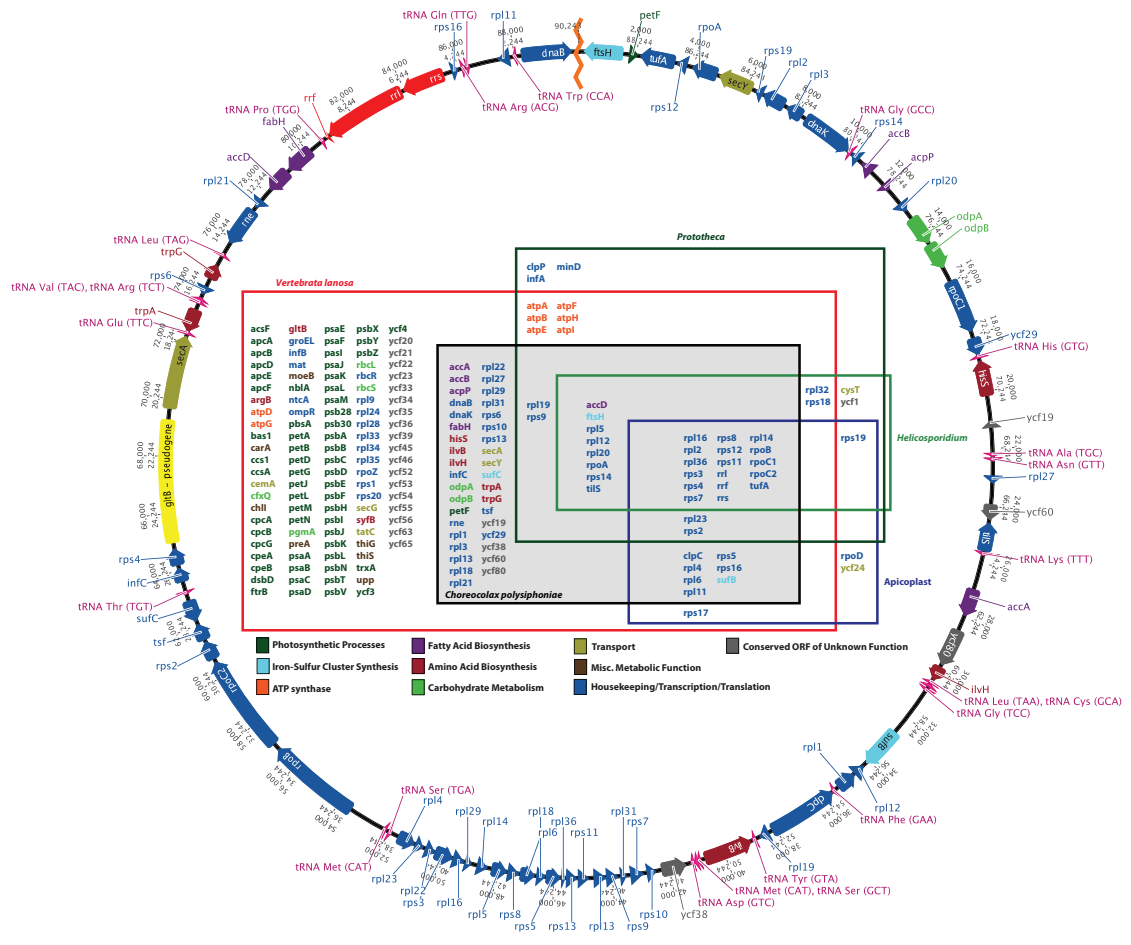


Figure 1. The plastid genome of the parasitic red alga, *Choreocolax polysiphoniae* is 90,243 basepairs and contains 71 coding genes (colored by protein function), the 5S, 16S, and 23S rRNAs (Red), and 24 tRNAs (Pink). All genes involved with photosynthetic functions, except *petF*, have been lost. The *ftsH* gene is truncated but may still be transcribed, however *gltB* is a non-functional pseudogene (Yellow). If the plastid is a linear molecule, the breakpoint is indicated by the orange zig-zag line. The Venn diagram in the middle shows the plastid genome content of the free living red alga, *Vertebrata lanosa* (red box), and its obligate parasite *C. polysiphoniae* (black box with grey shading). Additional boxes represent the green algal facultative parasite,

Prototheca wickerhamii (dark green box), the green algal obligate parasite, *Helicosporidium* sp. (light green box), and a composite Apicoplast (blue box). Genes are color coded by protein function.

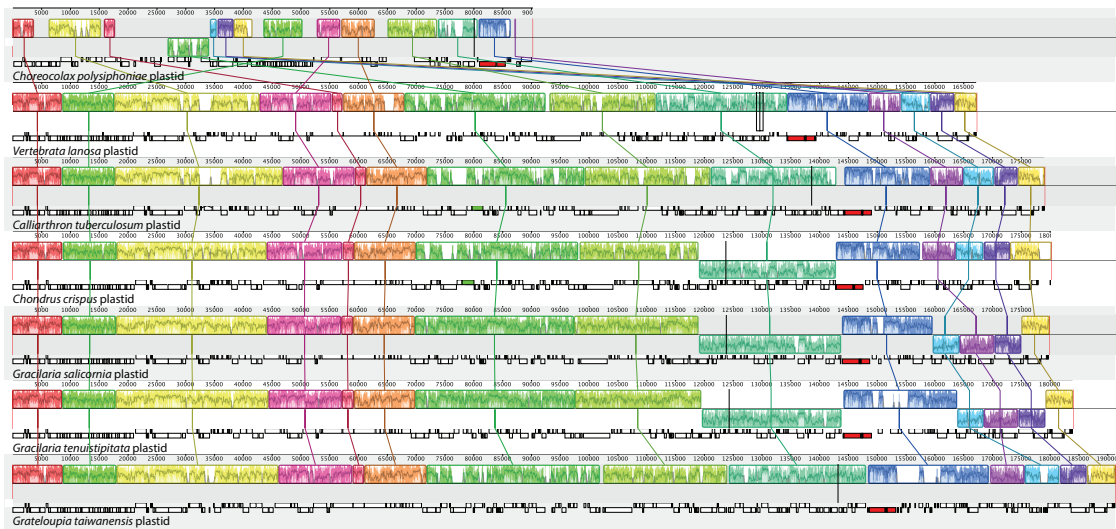


Figure 2. Mauve alignment of *Choreocolax polysiphoniae* (top) and *Vertebrata lanosa* (second from top) and other published Florideophyceae plastid genomes. This alignment identifies 13 locally collinear blocks (LCBs) among all Florideophyceae plastid genomes while highlighting the reduction and fragmentation evident in *C. polysiphoniae*.

References

- Bendich, A.J. 2004. Circular Chloroplast Chromosomes : The Grand Illusion. *Plant Cell*. 16:1661–6.
- Bendich, A.J. & Smith, S.B. 1990. Moving pictures and pulsed-field gel electrophoresis show linear DNA molecules from chloroplasts and mitochondria. *Curr. Genet.* 17:421–5.
- Blouin, N.A. & Lane, C.E. 2012. Red algal parasites: models for a life history evolution that leaves photosynthesis behind again and again. *Bioessays*. 34:226–35.
- Borza, T., Popescu, C.E. & Lee, R.W. 2005. Multiple metabolic roles for the nonphotosynthetic plastid of the green alga *Prototheca wickerhamii*. *Eukaryot. Cell*. 4:253–61.
- Cai, X.M., Fuller, A.L., McDougald, L.R. & Zhu, G. 2003. Apicoplast genome of the coccidian *Eimeria tenella*. *Gene*. 321:39–46.
- Callow, J.A., Callow, M.E. & Evans, L. V 1979. Nutritional studies on the parasitic red alga *Choreocolax polysiphoniae*. *New Phytol.* 83:451–62.
- Court, G.J. 1980. Photosynthesis and Translocation Studies of *Laurencia spectabilis* and its Symbiont *Janczewskia gardneri* (Rhodophyceae). *J. Phycol.* 16:270–9.
- 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.
- de Koning, A.P. & Keeling, P.J. 2006. The complete plastid genome sequence of the

- parasitic green alga *Helicosporidium* sp. is highly reduced and structured. *BMC Biol.* 4.
- DePriest, M.S., Bhattacharya, D. & López-Bautista, J.M. 2013. The plastid genome of the red macroalga *Grateloupia taiwanensis* (Halymeniaceae). *PLoS One.* 8:e68246.
- Douzery, E.J.P., Snell, E. a, Baptiste, E., Delsuc, F. & Philippe, H. 2004. The timing of eukaryotic evolution: does a relaxed molecular clock reconcile proteins and fossils? *Proc. Natl. Acad. Sci. U. S. A.* 101:15386–91.
- Evans, L. V, Callow, J.A. & Callow, M.E. 1973. Structural and physiological studies on the parasitic red alga *Holmsella*. *New Phytol.* 72:393–402.
- Fast, N.M., Kissinger, J.C., Roos, D.S. & Keeling, P.J. 2001. Nuclear-encoded, plastid-targeted genes suggest a single common origin for apicomplexan and dinoflagellate plastids. *Mol. Biol. Evol.* 18:418–26.
- Feldmann, G. & Feldmann, J. 1958. Recherches sur quelques Floridees parasites. *Rev. Gen. Bot.* 65:49–128.
- Freshwater, D.W., Fredericq, S., Butler, B.S., Hommersand, M.H. & Chase, M.W. 1994. A gene phylogeny of the red algae (Rhodophyta) based on plastid *rbcL*. *Proc. Natl. Acad. Sci. U. S. A.* 91:7281–5.
- Gardner, M.J., Bishop, R., Shah, T., de Villiers, E.P., Carlton, J.M., Hall, N., Ren, Q. et al. 2005. Genome sequence of *Theileria parva*, a bovine pathogen that transforms lymphocytes. *Science.* 309:134–7.
- Goff, L.J. 1976. The Biology of *Harveyella mirabilis* (Cryptonemiales,

- Rhodophyceae). V. Host Responses to Parasite Infection. *J. Phycol.* 12:313–28.
- Goff, L.J. 1979. The Biology of *Harveyella mirabilis* (Cryptonemiales, Rhodophyceae). VI. Translocation of Photoassimilated ¹⁴C. *J. Phycol.* 15:82–7.
- Goff, L.J. 1982. The biology of parasitic red algae. *Prog. Phycol. Res.* 1:289–369.
- Goff, L.J., Ashen, J. & Moon, D. 1997. The evolution of parasites from their hosts: A case study in the parasitic red alga. *Evolution (N. Y.)*. 51:1068–78.
- Goff, L.J. & Coleman, A.W. 1984a. Transfer of nuclei from a parasite to host. *Proc. Natl. Acad. Sci.* 81:5420–4.
- Goff, L.J. & Coleman, A.W. 1984b. Elucidation of fertilization and development in a red alga by quantitative DNA microspectrofluorometry. *Dev. Biol.* 102:173–94.
- 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. 1987. Nuclear transfer from parasite to host: A new regulatory mechanism of parasitism. *Ann. N. Y. Acad. Sci.* 503:402–23.
- 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.

- Hancock, L., Goff, L.J. & Lane, C.E. 2010. Red algae lose key mitochondrial genes in response to becoming parasitic. *Genome Biol. Evol.* 2:897–910.
- Happe, T. & Naber, J.D. 1993. Isolation, characterization and N-terminal amino acid sequence of hydrogenase from the green alga *Chlamydomonas reinhardtii*. *Eur. J. Biochem.* 214:475–81.
- Hawkins, E.K. 1972. Observations on the Developmental Morphology and Fine Structure of Pit Connections in Red Algae. *Cytologia (Tokyo)*. 37:759–68.
- Jacobs, J., Pudollek, S., Hemschemeier, A. & Happe, T. 2009. A novel, anaerobically induced ferredoxin in *Chlamydomonas reinhardtii*. *FEBS Lett.* 583:325–9.
- Janouškovec, J., Liu, S.-L., Martone, P.T., Carré, W., Leblanc, C., Collén, J. & Keeling, P.J. 2013. Evolution of red algal plastid genomes: ancient architectures, introns, horizontal gene transfer, and taxonomic utility of plastid markers. *PLoS One*. 8:e59001.
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S. et al. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 28:1647–9.
- Knauf, U. & Hachtel, W. 2002. The genes encoding subunits of ATP synthase are conserved in the reduced plastid genome of the heterotrophic alga *Prototheca wickerhamii*. *Mol. Genet. Genomics*. 267:492–7.
- Kohler, S., Delwiche, C.F., Denny, P.W., Tilney, L.G., Webster, P., Wilson, R.J.M., Palmer, J.D. et al. 1997. A Plastid of Probable Green Algal Origin in

- Apicomplexan Parasites. *Science* (80-). 275:1485–9.
- Kremer, B.P. 1983. Carbon economy and nutrition of the alloparasitic red alga *Harveyella mirabilis*. *Mar. Biol.* 239:231–9.
- Kugrens, P. & West, J.A. 1973. The ultrastructure of an alloparasitic red alga *Choreocolax polysiphoniae*. *Phycologia*. 12:175–86.
- Kuo, C.-H., Wares, J.P. & Kissinger, J.C. 2008. The Apicomplexan whole-genome phylogeny: an analysis of incongruence among gene trees. *Mol. Biol. Evol.* 25:2689–98.
- 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 *Ululania* (Ceramiales). *J. Phycol.* 46:580–90.
- Lafferty, K.D., Allesina, S., Arim, M., Briggs, C.J., De Leo, G., Dobson, A.P., Dunne, J.A. et al. 2008. Parasites in food webs: the ultimate missing links. *Ecol. Lett.* 11:533–46.
- Lim, L., Linka, M., Mullin, K. a, Weber, A.P.M. & McFadden, G.I. 2010. The carbon and energy sources of the non-photosynthetic plastid in the malaria parasite. *FEBS Lett.* 584:549–54.
- Lim, L. & McFadden, G.I. 2010. The evolution, metabolism and functions of the apicoplast. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 365:749–63.
- Nonomura, A.M. & West, J.A. 1981. Host specificity of *Janczewskia* (Ceramiales, Rhodophyta). *Phycologia*. 20:251–8.
- Pueschel, C.M. & Cole, K.M. 1982. Rhodophycean Pit Plugs: An Ultrastructural

- Survey with Taxonomic Implications. *Am. J. Bot.* 69:703–20.
- Ramus, J. 1969. Pit Connection Formation in the Red Alga *Pseudogloiophloea*. *J. Phycol.* 5:57–63.
- Ramus, J. 1971. Properties of septal plugs from the red alga *Griffithsia pacifica*. *Phycologia*. 10:99–103.
- Reinsch, P.F. 1875. *Contributiones ad Algologiam et Fungologiam*. Lipsiae, Nuremberg. 61 pp.
- Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B. et al. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61:539–42.
- Salomaki, E.D. & Lane, C.E. 2014. Are all red algal parasites cut from the same cloth? *Acta Soc. Bot. Pol.* 83:369–75.
- Saunders, G.W. 1993. Gel purification of red algal genomic DNA: An inexpensive and rapid method for the isolation of polymerase chain reaction-friendly DNA. *J. Phycol.* 29:251–4.
- Saunders, G.W. & Moore, T.E. 2013. Refinements for the amplification and sequencing of red algal DNA barcode and RedToL phylogenetic markers: a summary of current primers, profiles and strategies. *Algae*. 28:31–43.
- Schattner, P., Brooks, A.N. & Lowe, T.M. 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res.* 33:686–9.
- Setchell, W.A. 1914. Parasitic Florideae I. *Univ. Calif. Publ. Bot.* 6:1–34.

- Setchell, W.A. 1918. Parasitism among red algae. *Proc. Am. Philosophical Soc.* 57:155–72.
- Sherwood, A.R., Kurihara, A., Conklin, K.Y., Sauvage, T. & Presting, G.G. 2010. The Hawaiian Rhodophyta Biodiversity Survey (2006-2010): a summary of principal findings. *BMC Plant Biol.* 10:258.
- Sherwood, A.R. & Presting, G.G. 2007. Universal primers amplify a 23S rDNA plastid marker in eukaryotic algae and cyanobacteria. *J. Phycol.* 43:605–8.
- Stamatakis, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics.* 22:2688–90.
- Tamayo, D.A. & Hughey, J.R. 2016. Organellar genome analysis of the marine red alga *Dasya binghamiae* (Dasyaceae, Rhodophyta) reveals an uncharacteristic florideophyte mitogenome structure. *Mitochondrial DNA Part B.* 1:510–1.
- Turner, C.H.C. & Evans, L.V. 1978. Translocation of photoassimilated ¹⁴C in the red alga *Polysiphonia lanosa*. *Br. Phycol. J.* 13:51–5.
- Vaidya, A.B. & Mather, M.W. 2009. Mitochondrial evolution and functions in malaria parasites. *Annu. Rev. Microbiol.* 63:249–67.
- 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.
- Viola, R., Nyvall, P. & Pedersén, M. 2001. The unique features of starch metabolism in red algae. *Proc. Biol. Sci.* 268:1417–22.
- Wetherbee, R. & Quirk, H. 1982a. The Fine Structure of Secondary Pit Connection

- Formation Between the Red Algal Alloparasite *Holmsella australis* and Its Red Algal Host *Gracilaria furcellata*. *Protoplasma*. 176:166–76.
- Wetherbee, R. & Quirk, H. 1982b. The Fine Structure and Cytology of the Association Between the Parasitic Red Alga *Holmsella australis* and Its Red Algal Host *Gracilaria furcellata*. *Protoplasma*. 165:153–65.
- Wetherbee, R., Quirk, H.M., Mallett, J.E. & Ricker, R.W. 1984. The Structure and Formation of Host-Parasite Pit Connections Between the Red Algal Alloparasite *Harveyella mirabilis* and Its Red Algal Host *Odonthalia floccosa*. *Protoplasma*. 73:62–73.
- Wilson, R.J.M.I., Denny, P.W., Preiser, D.J., Rangachari, K., Roberts, K., Roy, A., Whyte, A. et al. 1996. Complete gene map of the plastid-like DNA of the malaria parasite *Plasmodium falciparum*. *J. Mol. Biol.* 261:155–72.
- Wolfe, K.H., Morden, C.W. & Palmer, J.D. 1992. Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. *Proc. Natl. Acad. Sci. U. S. A.* 89:10648–52.
- Yoon, H.S., Muller, K.M., Sheath, R.G., Ott, F.D. & Bhattacharya, D. 2006. Defining the major lineages of red algae (Rhodophyta). *J. Phycol.* 42:482–92.
- Zhu, G., Marchewka, M.J. & Keithly, J.S. 2002. *Cryptosporidium parvum* appears to lack a plastid genome. *Microbiology*. 146:315–21.
- 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. 1994a. Host Specificity In The Red Algal Parasites *Bostrichiocolax australis* And *Dawsoniocolax bostrychia* (Choreocolacaceae, Rhodophyta). *J. Phycol.* 30:462–73.
- Zuccarello, G.C. & West, J.A. 1994b. Genus and race specificity in the red algal parasite *Leachiella pacifica* (Choreocolacaceae, Rhodophyta). *Phycologia.* 33:213–8.
- Zuccarello, G.C. & West, J. a. 2003. Multiple Cryptic Species: Molecular Diversity and Reproductive Isolation in the *Bostrychia Radicans*/*B. Moritziana* Complex (Rhodomelaceae, Rhodophyta) With Focus on North American Isolates. *J. Phycol.* 39:948–59.

CHAPTER 3

Red Algal Mitochondrial Genomes are More Complete than Previously Reported

by

Eric D. Salomaki¹ and Christopher E. Lane¹

is published in *Genome Biology and Evolution*

¹Department of Biological Sciences, University of Rhode Island, Kingston, RI 02881.

Abstract

The enslavement of an alpha-proteobacterial endosymbiont by the last common eukaryotic ancestor resulted in large-scale gene transfer of endosymbiont genes to the host nucleus as the endosymbiont transitioned into the mitochondrion. Mitochondrial genomes have experienced widespread gene loss and genome reduction within eukaryotes and DNA sequencing has revealed that most of these gene losses occurred early in eukaryotic lineage diversification. On a broad scale, more recent modifications to organelle genomes appear to be conserved and phylogenetically informative. The first red algal mitochondrial genome was sequenced more than 20 years ago, and an additional 29 Florideophyceae mitochondria have been added over the past decade. A total of 32 genes have been described to have been missing or considered non-functional pseudogenes from these Florideophyceae mitochondria. These losses have been attributed to endosymbiotic gene transfer or the evolution of a parasitic life strategy. Here we sequenced the mitochondrial genomes from the red algal parasite *Choreocolax polysiphoniae* and its host *Vertebrata lanosa* and found them to be complete and conserved in structure with other Florideophyceae mitochondria. This result led us to resequence the previously published parasite *Gracilariophila oryzoides* and its host *Gracilariopsis andersonii*, as well as reevaluate reported gene losses from published Florideophyceae mitochondria. Multiple independent losses of *rpl20* and a single loss of *rps11* can be verified. However by reannotating published data and resequencing specimens when possible, we were able to identify the majority of genes that have been reported as lost or pseudogenes from Florideophyceae mitochondria.

Introduction

Endosymbiotic events have had a profound impact on eukaryotic evolution (Lane & Archibald 2008; Keeling 2010; Koonin 2010; Zimorski et al. 2014; Martin et al. 2015). All eukaryotes (with one recent exception (Karnkowska et al. 2016)) possess a mitochondrion or mitochondrion-related organelle (MRO) that was initially acquired from an alpha-proteobacteria endosymbiont (Gray et al. 1999; Lang et al. 1999; Koonin 2010; Gray 2012). Additionally, photosynthetic lineages maintain a plastid that originated as a cyanobacterial endosymbiont in the shared ancestor of Glaucophytes, Rhodophytes, and Viridiplantae (Chlorophytes and Streptophytes), and was subsequently spread through the eukaryotic tree of life via secondary and tertiary endosymbiotic events (Bhattacharya et al. 2004; Keeling 2004; Stiller 2007; Gould et al. 2008; Lane & Archibald 2008; Keeling 2010; Stiller et al. 2014). There is evidence of massive gene transfer from the endosymbiont to the host nucleus upon the initial acquisition of these organelles, resulting in host control and regulation of the organelle's function (Martin et al. 1998; Timmis et al. 2004; Qiu et al. 2013; Ku, Nelson-Sathi, Roettger, Garg, et al. 2015). Further organellar genome modifications appear to be mostly lineage specific, with gene losses and transfers being restricted within lineages (Tucker 2013; Janouškovec et al. 2013; Ku, Nelson-Sathi, Roettger, Sousa, et al. 2015; Qiu et al. 2015; Tanifuji et al. 2016). The conservation among organellar genomes, in addition to their being inherited predominately uniparentally, has made organelles prime targets for understanding evolutionary relationships across and within the eukaryotic tree of life.

Red algae (phylum Rhodophyta) diversified from their last common ancestor, shared with green algae, more than 1 billion years ago (Yoon et al. 2004). There are ~7,100 currently described species of rhodophytes that are divided into 7 classes; Bangiophyceae, Compsopogonophyceae, Cyanidiophyceae, Florideophyceae, Porpyridiophyceae, Rhodellophyceae, and Stylonematophyceae (Guiry & Guiry 2016). The Florideophyceae exhibit a wide range of morphological complexity and are by far the most species rich class, containing ~6,750 species spread across 30 orders (Guiry & Guiry 2016). Understanding the evolutionary relationships within the Florideophyceae has traditionally been complicated by phenotypic plasticity (Cianciola et al. 2010). More recently, molecular data have been analyzed and great progress has been made in describing new genera and species (Cianciola et al. 2010; Saunders & McDonald 2010; Le Gall & Saunders 2010). However, teasing apart the evolutionary histories of red algal orders has proven quite difficult even with the abundance of sequence data currently available (Verbruggen et al. 2010; Lam et al. 2016). Resolving the evolutionary relationships among florideophytes will provide a robust framework for asking a wide range of evolutionary questions including, but not limited to, transitions from marine to freshwater habitats, the evolution of the complex triphasic life-cycle found in many Florideophyceae orders, and the evolution of parasitism, a life strategy that has arisen many times among the Florideophyceae (Blouin & Lane 2012; Salomaki & Lane 2014; Blouin & Lane 2015; Lam et al. 2016). The use of the maternally inherited mitochondrial genome to resolve evolutionary relationships among the Florideophyceae shows promise (Yang et al. 2015).

The number of sequenced red algal organellar genomes has been increasing exponentially over the past decade. In part, this is a result decreasing sequencing costs allow for increasing use of next-generation sequencing technologies. Currently there are 30 published Florideophyceae mitochondrial genomes species available on GenBank (Table 1). However, only 16 of the 30 florideophycean orders are represented in these data, and 10 of those orders are represented by a single mitochondrion genome sequence.

Analyses of mitochondrion and MRO genomes across the tree of life have shown they are highly variable in gene content, arrangement, and structure (Smith & Keeling 2015). More recently, the oxymonad *Monocercomonoides* was shown to have entirely lost its MRO and all genes of mitochondrial origin that had been transferred to the nucleus (Karnkowska et al. 2016). Wide variability of mitochondrial genome content and structure has been implicated in the Florideophyceae as well. A study investigating the impact of adopting a parasitic life strategy on mitochondrial genomes of red algae described the *atp8* and *sdhC* genes of red algal parasite *Gracilariophila oryzoides* as pseudogenes, and that the *atp8* gene in the parasite *Plocamionocolax pulvinata* has been lost entirely (Hancock et al. 2010). The authors concluded that the products of these genes may be provided either from the parasite nucleus as a result of endosymbiotic gene transfer, or perhaps the proteins are being obtained from their hosts. More recently, Yang et al. (2015) sequenced 11 Florideophyceae mitochondrial genomes. Analysis of their data, in combination with all previously sequenced red algal mitochondria led them to describe widespread variation in gene arrangement and

multiple independent losses of *atp8*, *nad4L*, *rpl20*, *rps11*, *sdhC*, *sdhD*, *secY*, and *yfm39* across the Florideophyceae (Yang et al. 2015).

Prior to this study, 30 Florideophyceae mitochondrial genomes have been sequenced. Of those, 19 are reported to be missing a functional copy of at least one gene. A total of 8 different genes have been reported as pseudogenes or missing entirely from a Florideophyceae mitochondrial genome. Previous speculation on what is driving gene loss from Florideophyceae mitochondria include endosymbiotic gene transfer (EGT) from the mitochondrion to the nucleus (Hancock et al. 2010; Yang et al. 2015), and decreasing selective pressures in parasite mitochondria as a parasite may be obtaining products of those genes from the host (Hancock et al. 2010). Both explanations seem plausible, with the later hypothesis being directly responsible for the sequencing of the mitochondrial genome from the parasitic red alga, *Choreocolax polysiphoniae* and its host *Vertebrata lanosa* (this study).

The mitochondrial genomes of the parasitic red alga, *Choreocolax polysiphoniae* and its host *Vertebrata lanosa* represent the first mitochondrial genomes available from the family Rhodomelaceae, which comprises $\sim 1/7^{\text{th}}$ of species diversity within the phylum Rhodophyta (Guiry & Guiry 2016). In 2010 our lab reported that two mitochondrial respiratory protein-coding genes were degraded in the red algal parasites, *Gracilariophila oryzoides* and *Plocamiocolax pulvinata* (Hancock et al. 2010). Unexpectedly, the *C. polysiphoniae* mitochondrion has no degradation of respiratory mitochondrial genes. To reconcile these datasets we resequenced the mitochondrial genomes of the parasite *Gracilariophila oryzoides* and its host *Gracilariopsis andersonii*. Furthermore, we systematically reevaluate the described

gene losses from the other 30 previously published Florideophyceae mitochondrial genomes, revealing that more than two-thirds of the described losses are the result of errors in sequencing or downstream analyses. We find Florideophyceae mitochondrial genomes to be highly conserved and that gene losses are rare and predominately, if not entirely, observed in genes encoding ribosomal proteins.

Materials and Methods

Mitochondrial Genome Sequencing

Specimens of *Vertebrata lanosa* and *Choreocolax polysiphoniae* were collected from Beavertail State Park, Jamestown, RI, USA (voucher RI 0423). *Gracilariopsis andersonii* and *Gracilariophila oryzoides* were collected at Pigeon Point, Pescadero, CA, USA (voucher CL031613). Representatives of these parasite and host pair populations are retained as vouchers in the Lane Lab herbarium at the University of Rhode Island. Vegetative tissue from *V. lanosa* and *G. andersonii* was inspected for parasites and epiphytes under a dissecting microscope and clean tissue was ground under liquid nitrogen. Erumpent pustules of *Choreocolax polysiphoniae* (n=50) and *Gracilariophila oryzoides* (n=10) were excised from the thallus of their hosts *V. lanosa*, and *G. andersonii*, respectively, and collected in a 1.5 mL microcentrifuge tube. The parasite tissue was hand-ground using a Corning Axygen® PES-15-B-SI disposable tissue grinder pestle in a 1.5 mL microcentrifuge tube while submerged in 100 μ L of DNA extraction buffer (Saunders 1993). DNA was extracted from all specimens using a standard phenol/chloroform extraction (Saunders 1993).

All DNA libraries were prepared for Illumina sequencing on the Apollo 324 robot using the PrepX ILM DNA Library Kit (Wafergen Biosystems, Fremont, CA). The *G. andersonii* library was sequenced on a full-cell of an Illumina MiSeq paired-end 250 x 250 basepair run yielding 30,330,114 sequences in pairs. The *Gr. oryzoides* and *C. polysiphoniae* libraries were each sequenced on full-cells of an Illumina MiSeq paired-end 300 x 300 basepair run yielding 26,097,992 and 29,355,470 sequences in pairs, respectively. The *V. lanosa* library was sequenced on a partial-cell of an Illumina MiSeq paired-end 300 x 300 basepair run yielding 12,888,082 sequences in pairs. For all datasets, sequences with PHRED scores <30 were removed and the remaining reads were trimmed of adapter sequences. Additionally, fifteen 5' and five 3' nucleotides were trimmed from the remaining reads and all reads under 100 nucleotides were removed from the dataset. All trimming was completed using CLC Genomics Workbench v. 8.5.1 (CLC Bio-Qiagen, Aarhus, Denmark) and the remaining reads were assembled using default parameters in CLC Genomics Workbench v. 8.5.1. Trimmed sequencing reads for the *G. andersonii* and *Gr. oryzoides* mitochondria were mapped back to the previously published mitochondrion to compare the two assemblies and confirm support for differences.

Open reading frame (ORF) prediction on the *V. lanosa* and *C. polysiphoniae* mitochondrion sequences was done using translation table 4 (Protazoa Mitochondrion) using ATG as a start-codon in Geneious Pro v6.1 (Kearse et al. 2012). The resulting ORFs were manually annotated using blastN against GenBank. If blastN was insufficient for annotating an ORF, the ORF was translated to an amino acid sequence and then searched against the non-redundant protein sequence database (nr) in

GenBank using blastP and the Pfam database (Finn et al. 2010, 2015). If no conserved domain or sequence similarity could be found after searches using blastP or Pfam, the ORF remained without further annotation. Mitochondrion genome sequences were submitted to the tRNAscan-SE online server v1.21 for identification of tRNA sequences (Schattner et al. 2005). Ribosomal RNA predictions were based on annotations produced by MFannot (<http://megasun.bch.umontreal.ca/cgi-bin/mfannot/mfannotInterface.pl>).

Red Algal Mitochondrial Genome Conservation

All 30 currently available Florideophyceae mitochondrial genomes (Table 1) were downloaded from GenBank and imported into GeneiousPro v9.1. These mitochondrial genomes were combined with those from *V. lanosa* and *C. polysiphoniae* to create a database of Florideophyceae mitochondrial genomes. Sequences that have previously been found to have missing genes or pseudogenes were reanalyzed for ORFs in GeneiousPro v9.1. In cases where an ORF was found in a conserved location that had not previously been annotated as a gene, the ORF was translated and searched against the non-redundant protein sequence database (nr) in GenBank using blastP and Pfam. If this was insufficient to annotate an ORF in a conserved location, representatives of the missing genes were mapped back to the genome of interest for further evaluation and the region was manually curated. Translations of ORFs from locations of missing genes were aligned with annotated copies of those genes to manually assess annotation. When an apparent premature stop codon was found in a conserved location for a missing gene or pseudogene, the region

was resequenced using PCR amplification for confirmation when material or DNA from that species could be obtained.

To determine the AT content (%) and nonsynonymous to synonymous substitution ratio (dN/dS ratio), all protein-coding genes widely shared throughout the 29 Florideophyceae mitochondrial genomes were aligned using GeneiousPro v9.1. The average AT content (%) was calculated for each gene across the Florideophyceae in GeneiousPro v9.1. The CODEML program in PAML v. 4.8 (Yang 2007) was utilized to estimate the pairwise dN/dS ratio across all published Florideophyceae mitochondrion genes. For each gene analyzed, a nucleotide alignment was created using the translation align function in GeneiousPro v9.1 utilizing the Blosum62 cost matrix. Additionally, a Neighbor-Joining tree was constructed from the alignment in GeneiousPro v9.1 using the Tamura-Nei substitution model and the gene from *Hildenbrandia rubra* was used to root the tree. For the *rpl20* gene, which has been lost in *H. rubra*, *Palmaria palmata* was used to root the tree. The alignment and Neighbor-Joining tree were used as input files and the following parameters were specified in CODEML: runmode = 0; seqtype = 1; codonfreq = 0; model = 0; icode=4; and omega (measures dN/dS ratio) and kappa (measures transitions/transversions) were estimated.

Results and Discussion

Description of a Red Algal Alloparasite and Host Mitochondrion

The mitochondrial genomes of the red algal alloparasite, *Choreocolax polysiphoniae* (KX687877) and its host, *Vertebrata lanosa* (KX687880) represent the

first mitochondrial genomes sequenced from the Rhodomelaceae (Florideophyceae, Rhodophyta), further expanding the diversity of available red algal mitochondrial genomes. The *C. polysiphoniae* mitochondrial genome is a 25,357 basepair (bp) long circular molecule with an AT content of 79.4%. The mitochondrial genome of *C. polysiphoniae* encodes 23 protein coding genes, 2 rRNAs and 20 tRNAs and contains only 9.9% intergenic, non-coding DNA. The *V. lanosa* mitochondrial genome is also a circular molecule that is 25,337 bp long, has an AT content of 76.4%, and encodes 23 protein coding genes, 2 rRNAs and 19 tRNAs with 10.4% of the mitochondrial DNA being non-coding. Both the *C. polysiphoniae* and *V. lanosa* mitochondrial genomes maintain a similar genome architecture to other published red algal mitochondria.

Parasitic Red Algal Mitochondrial Genomes are Conserved

The *atp8* gene has previously been reported missing from five Florideophyceae mitochondria (Hancock et al. 2010; Yang et al. 2015) including the parasites *Gracilariophila oryzoides* and *Plocamiocolax pulvinata*. A re-annotation of the mitochondrion of *P. pulvinata*, identified the *atp8* as an ORF that was annotated only as a hypothetical protein CDS in the sequence downloaded from GenBank (Tables 1 & 2). Subsequently, resequencing the *Gracilariophila oryzoides* mitochondrial genome (KX687879) revealed a complete copy of the *atp8* gene, rather than the pseudogene that was previously reported using both Illumina and Sanger sequencing (Hancock et al. 2010). The *sdhC* gene was also previously reported to be a pseudogene in the adelphoparasite *Gracilariophila oryzoides* (Hancock et al. 2010). As with *atp8*, resequencing of the *G. oryzoides* mitochondrion (KX687879) demonstrated that there

was no frameshift mutation, as originally published, and that *sdhC* remains complete in red algal parasite mitochondria. These findings indicate that red algal parasites have not found alternative mechanisms for acquiring mitochondrion proteins and rely on their own mitochondrion for generating cellular energy as was previously hypothesized (Hancock et al. 2010).

Gene Loss in Other Red Algal Mitochondria

As a result of identifying conserved copies of genes originally reported to have been lost, we reevaluated all reported gene losses from red algal mitochondria. Investigation of the other reported *atp8* losses revealed that the gene was an ORF annotated as a hypothetical protein CDS in *Plocamium cartilagineum*, and that an ORF corresponding to *atp8* could be found in the published *Hildenbrandia rubra* sequence data that was not previously annotated (Tables 1 & 2). The *Ahnfeltia plicata* mitochondrion had a premature stop codon resulting in a pseudogene where *atp8* is normally found, however targeted PCR and sequencing showed the gene (KX687876) is complete. Analysis of the ratio of non-synonymous to synonymous substitutions (dN/dS ratio) in Florideophyceae copies of the *atp8* gene show a higher rate on non-synonymous mutations in *atp8* than *atp6* and *atp9*, which combine with *atp8* to make up the F₀ domain of the F₁F₀-ATP synthase complex involved in ATP synthesis (Table 3). However the dN/dS ratio of all three proteins remains <1 indicating that purifying selection is acting on the mitochondrial F₁F₀-ATP synthase complex in red algae, as is expected from genes essential for mitochondrion function.

Although the biological implications of losing the *nad4L* gene was not discussed in previous literature, the gene was noted as being absent in the mitochondrial genomes of both *Plocamiocolax pulvinata* and *Plocamium cartilagineum* (Hancock et al. 2010; Yang et al. 2015). In *Plocamiocolax pulvinata* an ORF was identified in the same location as other red algal copies of *nad4L*, between the 16S ribosomal RNA and the 26S ribosomal RNA, and both Pfam sequence search and blastP search of the translation strongly supports it coding for a functional *nad4L*. The published sequence for the mitochondrial genome of *Plocamium cartilagineum* splits an ORF here identified as the *nad4L* gene in two pieces, with the 5' portion of the sequence found from bases 26,172 - 26,431 and the 3' portion of the sequence is located from bases 1-43. With a *dN/dS* ratio of 0.14619, the *nad4L* gene remains under strong purifying selection in red algal mitochondria. Therefore, the loss of *nad4L* in any red algal mitochondria would represent a strong departure from this heavy selective pressure.

The *sdhD* gene encodes an essential protein that serves to anchor the succinate dehydrogenase complex II to the inner-membrane of the mitochondrion (Elorza et al. 2004; Bayley et al. 2005, 2006). The *sdhD* gene was reported missing from the mitochondria of *Ceramium japonicum* and *Asparagopsis taxiformis* (Yang et al. 2015) and the gene is also not annotated in the more recently published mitochondrial genome of *Ceramium sungminbooi* (Hughey & Boo 2016). Upon reanalysis, an ORF was identified between *nad4* and *nad2*, in the conserved Florideophyceae location of *sdhD* in the published mitochondrial genomes for all three species (see Tables 1 & 2). Furthermore, a translated alignment of these ORFs with other Florideophyceae copies

of *sdhD* show they are conserved in frame, retaining several critical conserved residues (Figure 1), and therefore should be annotated as *sdhD*.

Four mitochondria are reportedly missing copies of *sdhC*. Similarly to our findings with the *sdhD* genes, unannotated ORFs that are conserved with other Florideophyceae *sdhC* genes were identified from the mitochondrial genomes of *Ceramium sungminbooi* and *Dasya bingamiae* (Table 2). Based on the published *Sebdenia flabellata* mitochondrial genome, using an ATG as the only start-codon, there is no ORF that can be attributed to *sdhC*. However, using all start-codons in translation table 4 (Protozoa Mitochondrion) an ORF that is highly conserved in comparison with other Florideophyceae copies of *sdhC* is found with a TTA start-codon (Tables 1 & 2). Alternative start codons have previously been invoked for annotating red algal mitochondrion genes with variable support, which is discussed in more detail below (and see Table 4). The *Sebdenia flabellata sdhC* appears to be a well-justified case for using an alternative initiation codon.

The *Ceramium japonicum* mitochondrion is the other reported case of an *sdhC* gene loss (Yang et al. 2015). Although it appears to be highly conserved throughout the 5' region in comparison to other species, the *C. japonicum sdhC* is truncated by ~81 nucleotides (27 amino acids) at the 3' end when aligned with copies of the *sdhC* gene from other Florideophyceae. The *Coeloseira compressa sdhC* is similarly conserved at the 5' region and truncated at the 3' end. A Pfam search of the *C. japonicum* and *Co. compressa* sequences, translated to amino acids, confirms their identity as Succinate dehydrogenase/Fumarate reductase transmembrane subunit proteins though suggests they may be truncated as well. Although material was not

available for experimental validation, we speculate that this observed truncation has little effect on the functionality of *sdhC* as an anchor protein in succinate dehydrogenase complex II. The length of Florideophyceae *sdhC* genes (excluding *C. japonicum* and *Co. compressa*) is quite variable, ranging from 339 bp in *Plocamiocolax pulvinata* to 411 bp in *Asparagopsis taxiformis*. Furthermore, the dN/dS ratio remains at 0.47084 indicating that purifying selection is acting fairly strongly on deleterious mutations in *sdhC*. The alternative would seem that the *sdhC* gene in *C. japonicum* and *Co. compressa* is losing its functional capacity, which would hinder the ability of these free-living species to generate cellular energy.

Although not reported as a loss, the published *Gracilariopsis andersonii rps11* gene is inverted in comparison with all other Florideophyceae copies of the gene (Hancock et al. 2010; Yang et al. 2015). Resequencing this genome revealed an ORF in the conserved location between *nad3* and *atp9* that was not inverted and maintained strong homology with red algal *rps11* genes. Analysis of this ORF in comparison to the previously published *G. andersonii* mitochondrion identified a string of seven 'A's stretching from bases 9,158-9,164 correspond to only six 'A's in the newly sequenced mitochondrion. This apparent frameshift mutation resulted in a premature stop codon in the conserved direction that led to identifying an ORF in the same location but inverted as *rps11* in the earlier publication. The *rps11* gene in the resequenced *G. andersonii* mitochondrion, extending from bases 14,568 to 14,209, maintains strong homology with, and is encoded in the same direction as other Florideophyceae copies of *rps11*.

Although no genes were explicitly described as being lost in the recently sequenced *Dasya binghamiae* mitochondrial genome (Tamayo and Hughey 2016), annotations for *rpl20* and *sdhC* are absent from the published sequence. Additionally, alignments demonstrate that the *cox3*, *rps3* and *TatC* genes are truncated in comparison with other Florideophyceae. Perhaps even more interesting is the report of two inverted multi-gene rearrangements that are unprecedented in light of the highly conserved synteny in florideophycean mitochondria. Unfortunately a thorough evaluation of the losses, truncations and rearrangements in this mitochondrial genome is difficult as the publication is extremely brief (<500 words) and lacks essential details such as the sequencing platform from the materials and methods.

Frameshift Mutations are Overstated

In addition to the annotation of an inverted *rps11* in *Gracilariopsis andersonii*, frameshift mutations have been described as the cause for genes being lost or becoming pseudogenes in Florideophyceae mitochondria including *atp8* in *Ahnfeltia plicata* and *Gracilariophila oryzoides* and *sdhC* in *Gr. oryzoides*. The *Gracilaropsis andersonii rps12* gene is another case of an apparent frameshift mutation causing a gene to be truncated. In *G. andersonii*, the *rps12* gene is annotated at 240 nucleotides in length while other red algal copies of the gene range from 366-390 basepairs long. As a part of this study we resequenced the *G. andersonii* mitochondrion (KX687878) and identified that the ‘CT’ found at bases 25,864-25,865 in the previously published *G. andersonii* mitochondrion appears to be the result of sequencing or assembly error.

Without these additional bases, the *rps12* gene remains conserved and is 369 basepairs long.

The *Ceramium japonicum nad3* gene appears to be an instance of an unreported frameshift mutation causing a gene to be truncated. Although the *C. japonicum nad3* gene is annotated in its mitochondrial genome, it is only 234 bp long, whereas all other Florideophyceae *nad3* genes are either 363 or 366 bp long. An alignment of other Florideophyceae *nad3* genes to the *Ceramium japonicum nad3* region indicates that this truncation is the result of a frameshift mutation in a string of 26 ‘T’ and 3 ‘C’ between 32-60 bp from the start codon. A translated alignment of the annotated *Ceramium japonicum nad3* with all other Florideophyceae *nad3* genes shows little conservation in the annotated *Ceramium japonicum nad3*. Additionally, a blastP search of the NCBI nr database, and a Pfam sequence search of the translated original annotation shows the region is not homologous with any gene sequenced to date. However, manually deleting a ‘T’ from the previously mentioned string yields a 366 bp *nad3* gene that is highly conserved with copies of the *nad3* gene sequenced from other Ceramiales mitochondria (Figure 2) and is homologous with *nad3* genes in the NCBI nr database and Pfam database. Long homopolymer runs are notoriously challenging for both sequencing and assembly (Kieleczawa 2006; Gilles et al. 2011; Loman et al. 2012; Laehnemann et al. 2016) but is a more likely explanation than a frameshift resulting in two conserved sections of the gene.

The *Ceramium japonicum TatC (secY)* initially appears to be another case of a Florideophyceae mitochondrion gene losing function and becoming a pseudogene due to a frameshift mutation, and again, pinpointing the exact location of the mutation is

difficult. By manually manipulating the sequence and deleting a nucleotide from a string of 43 T's and 7 C's between 23,501 and 23,550 basepairs into the published sequence, an ORF that is highly conserved with other Florideophyceae *TatC* genes containing an ATT initiation codon is observed. Due to high levels of variation in length and sequence of Florideophyceae *TatC* genes, we continue to recognize the *Ceramium japonicum* *TatC* gene as a pseudogene until firm evidence contradicts this. However, based on our findings that all frameshift mutations previously discussed in this manuscript were the result of sequencing error or downstream analysis, it seems likely that is again the case here. Resequencing of this region is essential before considering *TatC* (*SecY*) as a true loss in *Ceramium japonicum* and the addition of RNA sequence data would help to confirm or reject this hypothesis.

Some Genes Have Degraded Into Pseudogenes

Even though secondary analysis of published sequences combined with subsequent PCR and resequencing efforts have found many of the genes that have been reported missing, this is not the case for all losses. The *rpl20* gene seems to blur the lines of deciphering when a gene is lost, and it appears to be the least conserved gene in Florideophyceae mitochondria. Interestingly, aside from its presence in red algal mitochondria, the only other lineage of eukaryotes reported to maintain *rpl20* are the jakobids (Burger & Nedelcu 2012). Retaining up to 67 genes, the most of any known mitochondria, Jakobid mitochondria are considered to most closely resemble the alpha-proteobacteria endosymbiont that became the contemporary mitochondrion (Gray et al. 2004; Burger & Nedelcu 2012; Burger et al. 2013). In the

Florideophyceae, *rpl20* has been reported missing or a pseudogene in 11 species including the two new additions from this study.

Annotation of *rpl20* has been complicated because, in addition to ATG, which is the most commonly used initiation-codon for Florideophyceae mitochondrion genes, it appears that ATA may serve as an initiation-codon for *rpl20* in *Ahnfeltia plicata*, and ATT in *Rhodymenia pseudopalmata* and *Sporolithon durum* (Table 4). Without these alternative initiation-codons, *rpl20* is likely a pseudogene in *A. plicata*, *R. pseudopalmata*, and *S. durum*. In addition to the aforementioned species, a conserved copy of *rpl20* using the ATG start codon was located in *Schimmelmannia schousboei* (previously not annotated) and *Schizymenia dubyi* (previously annotated as a hypothetical protein CDS).

In *Ceramium japonicum*, *Ceramium sungminbooi*, *Choreocolax polysiphoniae*, *Gelidium elegans*, *Gelidium vagum*, *Hildenbrandia rubra*, and *Vertebrata lanosa*, the 3' region of *rpl20* gene remains somewhat conserved, however the 5' end of the sequence is laden with stop codons, or appears to be missing entirely. Therefore *rpl20* is considered a pseudogene in these species. No substantial region in the *Dasya binghamiae* mitochondrial genome appears to be homologous to the *rpl20* gene. Furthermore, *rpl20* is annotated as a gene/coding sequence (CDS) in the *Coeloseira compressa* mitochondrial genome, however the 3' region is slightly truncated and not highly conserved with other *rpl20* copies, suggesting that perhaps this also is a pseudogene. This wide variability in *rpl20* initiation codons and conservation cause annotation to be extremely difficult. Confirming the presence or absence of a

functional *rpl20* localized in the mitochondrion is difficult and will likely require RNA sequencing and nuclear genome sequencing to identify possible cases of EGT.

The only unique Florideophyceae mitochondrion gene loss that appears to stand up to further scrutiny also encodes a ribosomal protein. Based on the published sequence of the *Sporolithon durum* mitochondrion, *rps11* has degraded to a pseudogene. In all other Florideophyceae, *rps11* is found adjacent to the 3' end of *nad3*, however in *S. durum*, this region contains no ORFs that can be attributed to a full-length copy of *rps11*. As in other genes, a frameshift mutation appears to be initially responsible for *rps11* becoming a pseudogene. However, in all previously discussed frameshift derived pseudogenes, it was apparent that the insertion or deletion of a nucleotide or two would 'repair' the gene and result in a conserved copy, that could then subsequently be confirmed by PCR. In the case of the *Sporolithon durum rps11*, artificially 'fixing' the gene could restore a conserved 3' end of the gene, however a 6 residue gap upstream of this 'fix' remained in translated alignments adding further support that *rps11* is no longer functional in *S. durum*. RNA and nuclear genome sequencing work remains necessary to identify whether this is a complete loss or a case of EGT from mitochondrion to the nucleus.

The Importance of Nomenclature

Identifying gene losses in Florideophyceae mitochondrial genomes has been further complicated by the use of two different names for a homologous gene. In Yang et al. (2015), the *ymf39* gene was reported as the most widely lost gene in Florideophyceae mitochondria, and was noted as being absent in six species:

Ceramium japonicum, *Gracilariopsis andersonii*, *Hildenbrandia rubra*, *Kappaphycus striatus*, *Schizymenia dubyi*, and *Sebdenia flabellata*. Furthermore, this gene is annotated only as a hypothetical protein CDS in *Gracilariopsis chorda*. Resequencing of the *Gracilariopsis andersonii* (KX687878) and reanalysis of the published *Hildenbrandia rubra* data reveals that the *ymf39* gene is present in both mitochondria. Interestingly, the other four species lacking *ymf39* are also the only Florideophyceae mitochondria with an annotated *atp4* gene, which is found between the *cox3* and *cob* genes, the same location as *ymf39* in other Florideophyceae mitochondria (Yang et al. 2015). According to Burger et al. (2003), *ymf39* encodes subunit b of mitochondrial F₀F₁-ATP synthase and should formally be designated as *atp4*.

Although it has not led to reports of gene loss, it is of note that three names have been applied to the *TatC* gene in Florideophyceae mitochondria. In *Chondrus crispus* the gene currently annotated as *ymf16* was initially described as a gene of unknown function called ORF 262 (Leblanc et al. 1995). In the publication of the *Porphyra purpurea* mitochondrial genome it was noted that *ymf16* is recognized as a homolog of *E. coli* *TatC* encoding a protein in the Sec-independent protein translocation pathway (Burger et al. 1999). *Coeloseira compressa*, *Dasya binghamiae*, and *Kappaphycus striatus* use the name *TatC*, which is a sec-independent protein translocase protein. In all other published Florideophyceae mitochondria this gene is called *SecY*, a sec dependent protein translocase protein. When these sequences are searched against the Pfam database, all similarity hits match sec-independent protein translocase protein (*TatC*). This gene was initially incorrectly annotated as *SecY* with the publication of the second, third, and fourth florideophycean mitochondrion

genomes (Hancock et al. 2010) with subsequent sequencing efforts transferring that nomenclature throughout the Florideophyceae. Furthermore, in their review of algal mitochondrial genomes, Burger and Nedelcu (2012) note that *SecY* is not found in the mitochondrial DNA of algae.

It seems reasonable for annotation efforts to rely largely upon previous publications as a reference, however in the case of *atp4*, our understanding of gene function surpassed the nomenclatural usage. The annotation of the *TatC* gene as *SecY* may have been the result of available comparative data or knowledge of mitochondrial translocase proteins at the time of the initial publication. In this effort to correct the course of mitochondrial genome annotations we support following the recommendations of Burger et al. (2003) that all *ymf39* annotations in Florideophyceae mitochondria be updated to *atp4* to reduce further confusion. Additionally, it is recommended that *SecY* and *ymf16* annotations be changed to *TatC*.

The Use of Alternative Initiation-Codons

The most widespread initiation-codon in Florideophyceae mitochondrion genes is ATG, though some exceptions have been previously proposed (Table 4). For example, in the *Grateloupia angusta* mitochondria the use of ATT, TTA, or TTG as initiation-codons was reported for 9 genes (Kim et al. 2014). Further examination of the published *G. angusta* mitochondrial genome revealed that 7 of the genes reported with an alternative initiation-codon (*atp4* (as *ymf39*), *atp6*, *cox2*, *orf-Gang5*, *rps11*, *sdhB*, and *sdhC*), an ORF starting with ATG could be found within a few basepairs of the previously annotated gene, and the current *G. angusta atp4* (as *ymf39*) gene

annotation on GenBank does use an ATG start-codon. The reasoning behind the decision to opt for an alternative codon rather than ATG at the beginning of the gene was not described in the genome announcement.

Alternative start-codons have been suggested in a few other Florideophyceae mitochondrion genes besides *Grateloupia angusta*. For example, in *Asparagopsis taxiformis* the *atp4* gene is annotated with the initiation-codon ATT, yet 6 basepairs (2 amino acid residues) away in the same reading frame is an ATG, which could also serve as the initiation-codon (Table 4). The *Grateloupia angusta* and *Sporolithon durum* copies of *atp6* are both annotated to start with ATT codons that are 9 and 3 basepairs (3 and 1 amino acid residues), respectively, upstream of an ATG (Table 4). A complete assessment of Florideophyceae mitochondrion genes that have been annotated using non-ATG protist mitochondrion initiation-codons and their proximity to a potential ATG start-codon is shown in Table 4.

Even though some of the alternative start-codon usage is questionable, though not necessarily incorrect, there appear to be several Florideophyceae mitochondrion genes that likely are using alternative start-codons (Table 4). In several of these cases the alternative hypothesis is that the genes are severely truncated and have been rendered non-functional. For example, it seems much more reasonable to believe that *Sebdenia flabellata* utilizes ATT as a start-codon for *cob* as opposed to losing the need to transcribe the first 42 amino acids of the protein. A similar situation occurs with the *sdhB* gene in *Grateloupia angusta*, where reliance on an ATG initiation-codon would result in the first 49 amino acids not being transcribed. Furthermore, maintaining a functional *TatC* (annotated as *SecY*) and *rpl16* in the *G. angusta* mitochondrion is

dependent on the use of alternative initiation-codons (ATT and TTA respectively) (Table 4).

The *Chondrus crispus* *TatC* (annotated as *ymf16*) gene is also likely reliant on an alternative start codon. Currently the *Chondrus crispus* *TatC* gene is annotated with a GTT initiation codon. However, GTT has not been used as an initiation codon in any other Florideophyceae mitochondrion gene, nor is it one of the start codon options in translation table 4 (Protozoa Mitochondrion). Four other ORFs in the same reading frame that use either ATA or TTA as a start codon for *TatC* gene are found from 12 - 39 nucleotides downstream of the GTT codon. All *Chondrus crispus* ORFs that can be reasonably attributed to *TatC* use a start-codon other than ATG, suggesting that this is another reliable instance for invoking an alternative, however the use of GTT is questionable.

Why Were Genes Reported Missing?

There are several technical and biological reasons that could explain the previous results of missing genes in Florideophyceae mitochondrial genomes. Each cell maintains numerous mitochondria and some of these may in fact maintain the frameshift mutations that have led to gene losses being reported in published literature (Hancock et al. 2010; Yang et al. 2015). Preferential amplification of these mitochondrial genomes, or segments of the genome when using targeted PCR, would lead to the aforementioned findings even in cases where other mitochondria in the cell remain fully functional. The first four Florideophyceae mitochondrial genomes to be sequenced were completed primarily using nuclease digestions or PCR amplification

and Sanger sequencing methods to assemble the genome at ~2x depth (Leblanc et al. 1995; Hancock et al. 2010). The advances in sequencing technologies and reduction in costs since the *Chondrus crispus* mitochondrion was first sequenced over 20 years ago have enabled much greater sequencing depths. For example the *Vertebrata lanosa* mitochondrial genome published here has average read coverage of 391x. This increased depth allows for the correction of ‘errors’ either in the biology or technical aspects of sequencing by utilizing the deeper coverage of sequencing reads when forming a consensus sequence. It is noteworthy that all frameshift mutations that have been reported and led to missing genes, pseudogenes or the inversion of the *Gracilariopsis andersonii rps11* were found in long homopolymer regions. Assembling sequences containing long regions of low-complexity, often dominated by a single nucleotide, has been recognized as a major complication for sequencing (Kieleczawa 2006; Laehnemann et al. 2016) and was especially difficult for the 454 FLX technology used in Hancock et al. (2010) (Gilles et al. 2011; Loman et al. 2012).

RNA data can sometimes help identify problematic annotation or assembly. The apparent *atp8* and *sdhC* pseudogenes observed in their *Gracilariophila oryzoides* DNA data were confusing to the authors as they noted that both genes were still being transcribed based on RNA sequencing efforts (Hancock et al. 2010). However, transfer to the nucleus was invoked at the time as a possible explanation. In retrospect, the RNA data was a strong indication to reexamine the data assembly. The sequencing errors reported in the first few published red algal mitochondrial genomes formed the foundation that was used as a reference for the annotation of subsequently sequenced Florideophyceae mitochondrial genomes. The apparent flexibility of mitochondrial

genomes based on early sequencing efforts set a precedent for gene loss in Florideophyceae mitochondria. Building on a flawed foundation has allowed for the gene loss to be overstated without a deeper reanalysis of results. This is in no way meant as a criticism of the researchers themselves and it is plausible that results shown in data being published with current technology will be revised with future advances.

Conclusions

A detailed investigation of previously reported gene losses in Florideophyceae mitochondria reveals that losses are much less common and widespread than the published literature indicates. Prior to this study there genes had been either described as lost, or annotations were overlooked from 18 of the 30 published mitochondrial genomes (Tables 1 & 2). Thoroughly examining each loss using the available published sequence data, in combination with resequencing those specimens that we could obtain material from, has positively identified 23 of the ‘missing’ genes. Overwhelmingly, the ‘missing’ genes or pseudogenes were the result of overlooked ORFs in the available sequence data and artificial frameshift mutations that resulted from sequencing and/or downstream assembly and analysis. In light of these findings, it is essential to thoroughly investigate results that indicate genes are degrading into pseudogenes or being lost entirely.

The *C. japonicum* mitochondrion was described as having lost five genes (*atp4*, *rpl20*, *sdhC*, *sdhD*, & *TatC*) of which three were identified here. Additionally, the translation of the existing annotation of the *nad3* gene shares no homology with any other gene, though that homology is restored through the manual deletion of a ‘T’

in a low complexity region of that gene. Furthermore a gap in the sequence is annotated between *cob* and *nad6*. Resequencing this mitochondrion to confirm the presence or absence of the *rpl20* and *TatC*, and close the gap is essential prior to inferring biological relevance resulting from these losses. However it does seem likely that at least *rpl20* remains absent from the *C. japonicum* mitochondrion considering it has also been truncated in all other Ceramiales mitochondrial genomes.

It is logical that gene losses would be rare in red algal mitochondria since the core genes encoded are essential for cellular respiration and oxidative phosphorylation. The loss of genes involved in these processes would interfere with the organisms' ability to produce cellular energy and would likely be a lethal mutation. The ribosomal proteins *rps11* and *rpl20* have been lost in the mitochondria of other lineages (Burger & Nedelcu 2012) and may be examples of gene transfer from the mitochondrion to the nucleus. Additional red algal genome data will allow for the identification of nuclear-encoded, mitochondrial target proteins.

Table 1. Table of all currently available Florideophyceae mitochondrion genomes that were examined in this study with GenBank Accession, genome length, and A/T%. Genes previously reported as missing are listed along with notes regarding their status as a result of this study.

Species	GenBank Accession	Length	AT Content (%)	Reported Missing Genes	Notes
<i>Ahnfeltia plicata</i>	KF649303	32,878	66.6	<i>atp8</i> , <i>rpl20</i>	atp8 : PCR of <i>atp8</i> region shows gene is present (KX687876). Previously published assembly was missing a single nucleotide, which resulted in a pseudogene. rpl20 : With ATA start codon, gene is fully present (243 bp), but ATG start codon gene is short (147 bp). Translations of both align with other red algal <i>rpl20</i> .
<i>Asparagopsis taxiformis</i>	KJ398158	26,097	73.3	<i>sdhD</i>	sdhD : Annotated in GenBank as ‘gene/hypothetical protein CDS’, BlastP of translated sequence hits other red algal <i>sdhD</i> sequences and aligns well with other translated red algal <i>sdhD</i> genes.
<i>Ceramium japonicum</i>	KJ398159	26,200	71.5	<i>atp4</i> (as <i>ymf39</i>), <i>sdhC</i> , <i>sdhD</i> , <i>TatC</i> (as <i>secY</i>), <i>rpl20</i>	atp4 : Annotated in GenBank as ‘gene/hypothetical protein CDS’, BlastP of translated sequence hits other red algal <i>atp4</i> (ATP synthase B chain precursor) sequences and aligns well with other translated red algal <i>atp4</i> (<i>ymf39</i>) genes. In Yang et al. (2015), considered <i>atp4</i> , this is a case of nomenclature causing confusion. sdhC : In the correct location there is a region annotated as ‘gene/hypothetical protein CDS’, but it is shorter than other Florideophyceae <i>sdhC</i> genes. Translation shows conservation of residues, particularly at 5’ region of the sequence. Possibly truncated at 3’ end, resequencing would help clarify. sdhD : Found in published data, but was unannotated 16,636>16,403. TatC : Seems to be the result of homopolymer sequence error, though possibility of pseudogene remains. Homology based on translation alignment of region from 23,168>23,590.

					rpl20 : Likely an actual loss, all sequenced members of Ceramiales also missing <i>rpl20</i> . Truncated from 5' end and about half the gene remains as a pseudogene.
<i>Ceramium sungminbooi</i>	KU145004 and KU145005	24,508 24,494	71.2 71.2	noted as partial in manuscript (Hughey & Boo 2016)	rpl20 : Likely a loss, all sequenced members of Ceramiales also missing <i>rpl20</i> . Truncated from 5' end and about half the gene remains as a pseudogene. sdhC : Found, as unannotated 10,977 > 10,678 in conserved location between <i>atp9</i> and <i>sdhB</i> in KU145004. sdhD : Found, as unannotated ORF 15,284 > 15,036 in conserved location between <i>nad4</i> and <i>nad2</i> in KU145004. TatC : Found, as unannotated ORF 15,277 > 15,029 in conserved location between <i>nad4</i> and <i>nad2</i> in KU145004.
<i>Chondrus crispus</i>	NC 001677	25,836	72.1	-	Complete
<i>Choreocolax polysiphoniae</i>	KX687877	25,357	79.4	<i>rpl20</i>	rpl20 : Likely an actual loss, all sequenced members of Ceramiales also missing <i>rpl20</i> . Truncated from 5' end and about half the gene remains as a pseudogene.
<i>Coeloseira compressa</i>	KU053956	25,391	74.3	-	rpl20 : Truncated and here considered a pseudogene.
<i>Corallina officinalis</i>	KU641510	26,504	69.9		Although the manuscript is published, not yet available on GenBank Reported use of GTG (<i>nad2</i>) and ATT (<i>sdhC</i>) as start codons, otherwise reported as complete
<i>Dasya binghamiae</i>	KX247283	26,052	77.4		sdhC : Found, as unannotated 21,403 > 21,789 in conserved location between <i>atp9</i> and <i>sdhB</i> . rpl20 : No good evidence of <i>rpl20</i> homology rps3 : Truncated version published seems to be the result of homopolymer sequence error, though possibility of pseudogene remains. Insertion of a nucleotide around 23,590 restores conservation of start and end of gene compared to other red algal <i>rps3</i> genes rps11 : Moving the initiation codon from 12,194 to 12,044 removes overlap with 3' of the <i>nad3</i> gene and results in conserved start with other Florideophyceae <i>rps11</i> genes based on alignment.
<i>Gelidium elegans</i>	KF290995	24,922	70.5	<i>rpl20</i>	rpl20 : With ATA start codon, gene is present but short (177 bp) with some conserved residues. Absent with ATG start.

					Area somewhat conserved but appears to be degraded too much to actually encode <i>rpl20</i> .
<i>Gelidium vagum</i>	KC875854	24,901	69.5	<i>rpl20</i>	rpl20 : With ATA start codon, gene is mostly present (222 bp) with some conserved residues. With ATG start it could be either 297 bp with overlapping tRNAs or short at 186 bp. Still considered a pseudogene here.
<i>Gracilaria chilensis</i>	KP728466	26,898	72.4	-	Complete
<i>Gracilaria salicornia</i>	KF852534	25,272	71.6	-	Complete
<i>Gracilaria vermiculophylla</i>	KJ526627	25,973	71.9	-	Complete
<i>Gracilariophila oryzoides</i>	NC_014771 and KX687879	25,161	71.9	<i>atp8, sdhC</i>	atp8 : Present and complete, early stop codon assumed to be sequencing error. sdhC : Present and complete, early stop codon assumed to be sequencing error. Newly sequenced <i>Gracilariophila oryzoides</i> mitochondrion available at GenBank (KX687879).
<i>Gracilariopsis andersonii</i>	NC_014772 and KX687878	27,036	72	<i>atp4 (as ymf39)</i>	atp4 : Present in resequenced <i>G. andersonii</i> mitochondrion. rps11 : Originally reported as inversion which appears to result from frameshift mutation in original sequence. This is corrected in resequenced mitochondrion. Newly sequenced <i>Gracilariopsis andersonii</i> mitochondrion available at GenBank (KX687878).
<i>Gracilariopsis chorda</i>	NC_023251	26,534	72.4	<i>atp4 (as ymf39)</i>	atp4 : Annotated as gene/hypothetical protein CDS in GenBank but considered as 'present' in Yang et al. (2015). Definitely present.
<i>Gracilariopsis lemaneiformis</i>	JQ071938	25,883	72.5	-	Complete
<i>Grateloupia angusta</i>	NC_023094	27,943	69.8	-	Complete, but uses multiple start codons, seemingly where unnecessary (see Table 4). Contains hypothetical protein CDS in <i>coxI</i> intron.
<i>Grateloupia taiwanensis</i>	KM999231	28,906	68.6	-	Complete. Contains hypothetical protein CDS in <i>coxI</i> intron that is also annotated as <i>coxI</i> . atp4 : Initiation codon questionable, see Table 4.
<i>Hildenbrandia rubra</i>	KF649304	33,066	67.8	<i>atp4, atp8, rpl20</i>	atp4 : Found, as unannotated 13,225 > 13785. ORF was at conserved location immediately after <i>cox3</i> . Translation conserved with other Florideophyceae <i>atp4</i> genes. atp8 : Found, as unannotated 24,587 > 24,213. ORF present

					in conserved location between <i>atp8</i> and <i>nad5</i> , 375 bp and conserved residues at n-terminus. rpl20 : Truncated 5' region – 135 bp, here considered a pseudogene.
<i>Kappaphycus striatus</i>	KF833365	25,242	69.9	<i>atp4</i> (as <i>ymf39</i>)	atp4 : Annotated as <i>atp4</i> gene. This is a case of nomenclature causing confusion.
<i>Mastocarpus papillatus</i>	KX525587	25,906	65.0		Complete
<i>Palmaria palmata</i>	KF649305	29,735	67.8	-	Complete
<i>Plocamiocholax pulvinata</i>	HQ586061	25,894	76.1	<i>atp8</i> , <i>nad4L</i>	atp8 : Found, as annotated gene/hypothetical protein CDS in GenBank. 20,389 > 19,985. nad4L : Found, as unannotated 25,592 > 25,894. atp4 : Moving the initiation codon from 7,772 to 7,790 will remove overlap with 3' of the <i>cox3</i> gene and result in conserved start with other Florideophyceae <i>atp4</i> genes based on alignment.
<i>Plocamium cartilagineum</i>	KJ398160	26,431	76.4	<i>atp8</i> , <i>nad4L</i>	atp8 : Present, as annotated gene/ATP synthase F0 subunit 8 CDS in GenBank. 20,528 > 20,127. nad4L : Found, as unannotated 26,172 > 43 (linear sequence of circular molecule ends and starts over from the beginning of sequence). atp4 : Moving the initiation codon from 7,766 to 7,784 will remove overlap with 3' of the <i>cox3</i> gene and result in conserved start with other Florideophyceae <i>atp4</i> genes based on alignment.
<i>Rhodymenia pseudopalmata</i>	KC875852	26,351	70.5	<i>rpl20</i>	rpl20 : ATG start codon leaves it 40 residues shorter than other, while ATT start codon leaves it 4 residues shorter than others. Here considered present with ATT start. RNA for this gene would be very useful to confirm whether this is transcribed and not a pseudogene.
<i>Riquetophycus</i> sp.	KJ398161	26,351	74.3	-	Complete
<i>Schimmelmannia schousboei</i>	KJ398162	25,906	73.3	<i>rpl20</i>	rpl20 : Found, as unannotated 23,912 > 24,148.
<i>Schizymenia dubyi</i>	KJ398163	26,438	74.1	<i>atp4</i> (as <i>ymf39</i>), <i>rpl20</i>	atp4 : Annotated in GenBank as 'gene/hypothetical protein CDS', BlastP of translated sequence hits other red algal <i>atp4</i> & ATP synthase B chain precursor sequences and aligns well with other translated red algal <i>atp4</i> genes. In Yang et al.

					(2015), considered <i>atp4</i> . rpl20 : Annotated in GenBank as ‘gene/hypothetical protein CDS’, BlastP of translated sequence shows similarity to other red algal <i>rpl20</i> genes.
<i>Sebdenia flabellata</i>	KJ398164	26,767	71.5	<i>atp4</i> (as <i>yfm39</i>), <i>sdhC</i>	atp4 : Annotated in GenBank as ‘gene/hypothetical protein CDS’, BlastP of translated sequence hits other red algal <i>atp4</i> & ATP synthase B chain precursor sequences and aligns well with other translated red algal <i>atp4</i> genes. In Yang et al. (2015), considered <i>atp4</i> . sdhC : Found a conserved ORF that relies on a TTA start codon (11,358 > 10,966).
<i>Sporolithon durum</i>	KF186230	26,202	71.6	<i>rps11</i> , <i>rpl20</i>	rpl20 : Many options exist for start codons other than ATG, clearly conserved residues present. Considered present in this study using ATT start codon (Table 4). rps11 : Pseudogene or sequencing error, regional conservation remains.
<i>Vertebrata lanosa</i>	KX687880	25,119	71.7	<i>rpl20</i>	rpl20 : Likely an actual loss, all sequenced members of Ceramiales also missing <i>rpl20</i> . Truncated from 5’ end and about half the gene remains as a pseudogene.

Table 2. Current status of Florideophyceae mitochondrial genes previously reported as missing in Hancock *et al.* (2010) and Yang *et al.* (2015) or otherwise unannotated.

Gene	Current Status		
	Gene Present	Location in Published Sequence or GenBank Accession Number	Pseudogene
<i>atp8</i>	<i>Ahnfeltia plicata</i> <i>Gracilariophila oryzoides</i> <i>Hildenbrandia rubra</i> <i>Plocamiocolax pulvinata</i> <i>Plocamium cartilagineum</i>	KX687876 20,431 > 20,024 ^d 24,587 > 24,213 20,389 > 19,985 20,528 > 20,127	
<i>nad4L</i>	<i>Plocamiocolax pulvinata</i> <i>Plocamium cartilagineum</i>	25,592 > 25,894 26,129 > 26,474 (43)	
<i>rpl20</i>	<i>Ahnfeltia plicata</i> ^a <i>Rhodymenia pseudopalmata</i> ^a <i>Schimmelmanna schousboei</i> <i>Schizymenia dubyi</i> <i>Sporolithon durum</i> ^a	30,851 > 31,093 24,127 > 24,351 23,912 > 24,148 24,440 > 24,709 24,248 > 24,487	<i>Ceramium japonicum</i> <i>Ceramium sungminbooi</i> <i>Choreocolax polysiphoniae</i> <i>Coeloseira compressa</i> <i>Dasya binghamiae</i> ^f <i>Gelidium elegans</i> <i>Gelidium vagum</i> <i>Hildenbrandia rubra</i> <i>Vertebrata lanosa</i>
<i>rps11</i>			<i>Sporolithon durum</i>
<i>sdhC</i>	<i>Ceramium japonicum</i> <i>Ceramium sungminbooi</i> <i>Dasya binghamiae</i> <i>Gracilariophila oryzoides</i> <i>Sebdenia flabellata</i> ^a	12,331 > 12,044 10,977 > 10,678 21,403 > 21,789 11,593 > 11,234 ^d 11,358 > 10,966	
<i>sdhD</i>	<i>Asparagopsis taxiformis</i>	15,759 > 15,514	

	<i>Ceramium japonicum</i>	16,636 > 16,403	
	<i>Ceramium sungminbooi</i>	15,284 > 15,036	
TatC	<i>Ceramium sungminbooi</i>	15,277 > 15,029	<i>Ceramium japonicum</i> ^g
atp4	<i>Ceramium japonicum</i> ^b	6,935 > 7,489	
	<i>Gracilariopsis andersonii</i>	7,560 > 8,102 ^e	
	<i>Gracilariopsis chorda</i> ^c	7,238 > 7,780	
	<i>Hildenbrandia rubra</i>	13,225 > 13,785	
	<i>Schizymenia dubyi</i> ^b	7,261 > 7,803	
	<i>Sebdenia flabellata</i> ^b	7,234 > 7,776	

^a Indicates presence of functional gene is dependent on non-ATG start codon, alternatively these could be pseudogenes. RNA sequence data would be required to confirm gene function.

^b The *atp4* gene was annotated as hypothetical protein CDS in GenBank but considered as *atp4* in Yang *et al.* (2015), Fig. 2.

∞ ^c The *atp4* gene was annotated as hypothetical protein CDS in GenBank but considered as *ymf39* in Yang *et al.* (2015), Fig. 2.

^d Location in newly sequenced *Gracilariophila oryzoides* mitochondrion (GenBank KX687879)

^e Location in newly sequenced *Gracilariopsis andersonii* mitochondrion (GenBank KX687878)

^f No evidence for remnant pseudogene

^g The *Ceramium japonicum* *TatC* gene seems likely to be the result of homopolymer sequence error, though possibility of pseudogene remains. Due to high levels of variation in length and sequence of Florideophyceae *TatC* genes, we continue to recognize the *Ceramium japonicum* *TatC* gene as a pseudogene until firm evidence contradicts this.

Table 3. A/T%, non-synonymous to synonymous mutation (dN/dS) ratio and individual dN and dS values for genes encoded on the Florideophyceae mitochondrion genomes. Species with a pseudogene, rather than a functional copy of the gene, are listed in the far right column and were left out of calculations of A/T% and dN/dS ratio.

Gene	AT Content (%)	dN/dS ratio	dN	dS	Species with pseudogene only
<i>atp4</i>	79.8	0.51645	0.22882	0.44307	
<i>atp6</i>	72.5	0.12864	0.07486	0.58195	
<i>atp8</i>	78.8	0.51248	0.20110	0.39241	
<i>atp9</i>	66.1	0.01550	0.00212	0.13673	
<i>cob</i>	70.3	0.11691	0.05443	0.46555	
<i>cox1</i> ^a	67.2	0.06936	0.01711	0.24675	
<i>cox2</i>	69.5	0.14959	0.04549	0.30414	
<i>cox3</i>	68.3	0.12819	0.07186	0.56058	
<i>nad1</i>	69.0	0.09454	0.06667	0.70520	
<i>nad2</i>	74.6	0.31481	0.15877	0.50434	
<i>nad3</i> ^b	74.5	0.15718	0.07060	0.44919	
<i>nad4</i>	72.7	0.18548	0.07790	0.41999	
<i>nad4L</i>	75.3	0.14619	0.05654	0.38677	
<i>nad5</i>	71.9	0.21716	0.08547	0.39356	
<i>nad6</i>	75.1	0.26439	0.17651	0.66762	
<i>rpl16</i>	75.9	0.36193	0.13664	0.37752	
<i>rpl20</i>	79.1	0.62160	0.46012	0.74023	<i>Ceramium japonicum</i> <i>Ceramium sungminbooi</i> <i>Choreocolax polysiphoniae</i> <i>Coeloseira compressa</i> <i>Dasya binghamiae</i> ^d <i>Gelidium elegans</i> <i>Gelidium vagum</i> <i>Hildenbrandia rubra</i> <i>Vertebrata lanosa</i>
<i>rps3</i> ^c	76.9	0.51571	0.28532	0.55326	
<i>rps11</i>	77.3	0.45177	0.19375	0.42886	<i>Sporolithon durum</i>
<i>rps12</i>	67.5	0.20794	0.11986	0.57642	
<i>sdhB</i>	71.7	0.20603	0.10707	0.51969	
<i>sdhC</i>	78.6	0.47084	0.17487	0.37140	
<i>sdhD</i>	78.2	0.46556	0.26263	0.56412	
<i>TatC</i>	80.0	0.50046	0.34550	0.69037	<i>Ceramium japonicum</i>

^a Introns removed and coding sequences only were used for A/T% and dN/dS analysis

^b *Ceramium japonicum nad3* left out of dN/dS analysis

^c *Dasya binghamiae rps3* left out of dN/dS analysis

^d No evidence for remnant pseudogene, appears to be a complete loss

Table 4. The use of alternative start codons by gene based on published literature, and the proximity to the closest ATG start codon. Alternative initiation codons that are supported by the lack of a nearby ATG initiation codon and maintained gene start location as observed by alignment are indicated in bold.

Gene	Species	Location with Alternative Initiation Codon	Location with ATG Initiation Codon	Difference in Gene Length in nucleotides (Amino Acids)
<i>atp4</i>	<i>Asparagopsis taxiformis</i>	7,298 (ATT)	7,304	6 (2)
<i>atp6</i>	<i>Grateloupia angusta</i>	23,475 (ATT)	23,466	9 (3)
	<i>Sporolithon durum</i>	21,462 (ATT)	21,465	3 (1)
<i>atp8</i>	-	-	-	-
<i>atp9</i>	-	-	-	-
<i>cob</i>	<i>Sebdenia flabellata</i>^a	8,070 (ATT)	8,196	126 (42)
<i>cox1</i>	<i>Schizymenia dubyi</i>	3,889 (ATT)	3,904	15 (5)
<i>cox2</i>	<i>Grateloupia angusta</i>	7,845 (TTG)	7,818	27 (9)
	<i>Palmaria palmata</i>	6,556 (ATT)	6,535	21 (7)
<i>cox3</i>	-	-	-	-
<i>nad1</i>	-	-	-	-
<i>nad2</i>	-	-	-	-
<i>nad3</i>	-	-	-	-
<i>nad4</i>	<i>Hildenbrandia rubra</i>	22,165	22,150	15 (5)
<i>nad4L</i>	-	-	-	-
<i>nad5</i>	-	-	-	-
<i>nad6</i>	-	-	-	-
<i>rpl16</i>	<i>Ahnfeltia plicata</i>^a	5,899 (TTA)	6,169	270 (90)
	<i>Grateloupia angusta</i>^a	3,324 (TTA)	3,585	261 (87)
<i>rpl20</i>	<i>Ahnfeltia plicata</i> ^{a,c}	30,851 (ATA)	30,947	96 (32)
	<i>Rhodymenia pseudopalmata</i> ^a	24,127 (ATT)	24,235	108 (36)
	<i>Sporolithon durum</i> ^a	24,248 (ATT)	-	-
<i>rps3</i>	<i>Plocamium cartilagineum</i>	2,693 (ATC)	2,732	39 (13)
	<i>Sporolithon durum</i>	2,574 (ATA)	-	-
<i>rps11</i>	<i>Ahnfeltia plicata</i>^a	17,436 (ATA)	17,277	159 (53)
	<i>Gracilaria chilensis</i>	10,426 (ATT)	10,423	3 (1)
	<i>Grateloupia angusta</i>	14,946 (ATT)	14,938	9 (3)
	<i>Palmaria palmata</i>^a	14,083 (TTA)	-	-
<i>rps12</i>	-	-	-	-
<i>sdhB</i>	<i>Grateloupia angusta</i>^a	13,459 (TTA)	13,312	147 (49)
<i>sdhC</i>	<i>Grateloupia angusta</i>	13,852 (TTG)	13,849	3 (1)
	<i>Hildenbrandia rubra</i> ^{a,b}	17,061 (ATT)	17,022	39 (13)
	<i>Schizymenia dubyi</i>	11,319 (CTT)	11,310	9 (3)
<i>sdhD</i>	<i>Ahnfeltia plicata</i>	20,718 (TTA)	20,733	15 (5)
	<i>Hildenbrandia rubra</i> ^a	20,662 (ATA)	-	-
<i>TatC</i>	<i>Ahnfeltia plicata</i>	29,469 (ATT)	29,487	18 (6)
	<i>Chondrus crispus</i>^{a,d}	348 (GTT)	681	333 (111)
	<i>Grateloupia angusta</i>^a	24,671 (ATT)	25,031	360 (120)
	<i>Hildenbrandia rubra</i>^a	29,085 (ATA)	-	-

<i>Kappaphycus striatus</i> ^a	24,437 (TTG)	24,176	261 (87)
<i>Palmaria palmata</i> ^a	24,152 (ATC)	23,687	465 (155)
<i>Rhodomenia pseudopalmata</i> ^a	22,834 (TTA)	23,350	516 (172)
<i>Schimmelmania schousboei</i> ^a	22,636 (ATT)	23,188	552 (184)
<i>Sebdenia flabellata</i> ^a	23,492 (TTA)	23,546	54 (18)
<i>Sporolithon durum</i>	22,915 (ATC)	22,942	27 (9)

^a Indicates examples where other non-ATG initiation codons from translation table 4 are also possible locations for the gene to start although no ATG codon is found within 30 nucleotides (10 amino acid residues) upstream or downstream from the start of the currently annotated gene.

^b The *Hildenbrandia rubra sdhC* gene annotation is longer than other copies of *sdhC* and the beginning of the gene overlaps with a tRNA. Starting annotation at ATG makes the gene much more similar in length to other Florideophyceae copies of *sdhC*.

^c Gene not previously annotated in GenBank

^d The *Chondrus crispus TatC (ymf16)* gene is currently annotated with a GTT initiation codon, which is not found for any other Florideophyceae mitochondrion gene nor is it a start codon in translation table 4 (Protozoa Mitochondrion). Four other ORFs in the same reading frame that use either ATA or TTA as a start codon for *TatC* gene are found from 12 -39 nucleotides downstream of the GTT codon.



Figure 1. Translated alignment of *sdhD* genes from florideophycean mitochondria showing *Asparagopsis taxiformis*, *Ceramium japonicum* and *Ceramium sungminbooi* (top three sequences) share critical conserved residues with all other Florideophyceae *sdhD* genes.



Figure 2. Alignment of the original *Ceramium japonicum nad3* gene with the modified *Ceramium japonicum nad3* (“T” deleted from base 36; red box) and copies of the *nad3* gene from *Chondrus crispus*, *Gracilaria vermiculophylla*, *Gracilariopsis andersonii*, *Sporolithon durum*, *Choreocolax polysiphoniae*, and *Vertebrata lanosa*. Manual deletion of one ‘T’ from the string of 26 ‘T’s and 3 ‘C’s between 32-60 bp from the start codon restores conservation of the length and sequence of the *Ceramium japonicum nad3* gene. Genes are shown with the amino acid translation below.

References

- Bayley J-P, Devilee P, Taschner PEM. 2005. The SDH mutation database: an online resource for succinate dehydrogenase sequence variants involved in pheochromocytoma, paraganglioma and mitochondrial complex II deficiency. *BMC Med. Genet.* 6:39.
- Bayley J-P et al. 2006. Mutation analysis of SDHB and SDHC: novel germline mutations in sporadic head and neck paraganglioma and familial paraganglioma and/or pheochromocytoma. *BMC Med. Genet.* 7:1.
- Bhattacharya D, Yoon HS, Hackett JD. 2004. Photosynthetic eukaryotes unite: endosymbiosis connects the dots. *Bioessays.* 26:50–60.
- Blouin NA, Lane CE. 2012. Red algal parasites: models for a life history evolution that leaves photosynthesis behind again and again. *Bioessays.* 34:226–35.
- Blouin NA, Lane CE. 2015. Red algae provide fertile ground for exploring parasite evolution. *Perspect. Phycol.* 3:11–19.
- Burger G, Gray MW, Forget L, Lang BF. 2013. Strikingly bacteria-like and gene-rich mitochondrial genomes throughout jakobid protists. *Genome Biol. Evol.* 5:418–438.
- Burger G, Lang BF, Braun HP, Marx S. 2003. The enigmatic mitochondrial ORF ymf39 codes for ATP synthase chain b. *Nucleic Acids Res.* 31:2353–2360.
- Burger G, Nedelcu AM. 2012. Mitochondrial Genomes of Algae. In: *Genomics of Chloroplasts and Mitochondria*. Bock, R & Knoop, V, editors. Springer pp. 127–157.
- Burger G, Saint-Louis D, Gray MW, Lang BF. 1999. Complete sequence of the mitochondrial DNA of the red alga *Porphyra purpurea*. Cyanobacterial introns and shared ancestry of red and green algae. *Plant Cell.* 11:1675–1694.

- Cianciola EN, Popolizio TR, Schneider CW, Lane CE. 2010. Using molecular-assisted alpha taxonomy to better understand red algal biodiversity in Bermuda. *Diversity*. 2:946–958.
- Elorza A et al. 2004. Nuclear SDH2-1 and SDH2-2 genes, encoding the iron-sulfur subunit of mitochondrial complex II in Arabidopsis, have distinct cell-specific expression patterns and promoter activities. *Plant Physiol*. 136:4072–4087.
- Finn RD et al. 2010. The Pfam protein families database. *Nucleic Acids Res*. 38:D211–22.
- Finn RD et al. 2015. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res*. 44:D279–D285.
- Gilles A et al. 2011. Accuracy and quality assessment of 454 GS-FLX Titanium pyrosequencing. *BMC Genomics*. 12:245.
- Gould SB, Waller RR, McFadden GI. 2008. Plastid evolution. *Annu. Rev. Plant Biol*. 59:491–517.
- Gray MW. 2012. Mitochondrial Evolution. *Cold Spring Harb. Perspect. Biol*. 4:a011403.
- Gray MW, Burger G, Lang BF. 1999. Mitochondrial Evolution. *Science*. 283:1476–1482.
- Gray MW, Lang BF, Burger G. 2004. Mitochondria of protists. *Annu. Rev. Genet*. 38:477–524.
- Guiry MD, Guiry GM. 2016. Algaebase. World-wide Electron. Publ. Natl. Univ. Ireland, Galway. <http://www.algaebase.org>. searched on 21 July 2016.
- Hancock L, Goff LJ, Lane CE. 2010. Red algae lose key mitochondrial genes in

- response to becoming parasitic. *Genome Biol. Evol.* 2:897–910.
- Hughey JR, Boo GH. 2016. Genomic and phylogenetic analysis of *Ceramium cimbricum* (Ceramiales, Rhodophyta) from the Atlantic and Pacific Oceans supports the naming of a new invasive Pacific entity *Ceramium sungminbooi* sp. nov. *Bot. Mar.* 59:211–222.
- Janouškovec J et al. 2013. Evolution of red algal plastid genomes: ancient architectures, introns, horizontal gene transfer, and taxonomic utility of plastid markers. *PLoS One.* 8:e59001.
- Karnkowska A et al. 2016. A Eukaryote without a Mitochondrial Organelle. *Curr. Biol.* 1–11.
- Kearse M et al. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics.* 28:1647–1649.
- Keeling PJ. 2004. Diversity and evolutionary history of plastids and their hosts. *Am. J. Bot.* 91:1481–1493.
- Keeling PJ. 2010. The endosymbiotic origin, diversification and fate of plastids. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 365:729–748.
- Kieleczawa J. 2006. Fundamentals of sequencing of difficult templates-An overview. *J. Biomol. Tech.* 17:207–217.
- Kim SY, Yang EC, Boo SM, Yoon HS. 2014. Complete mitochondrial genome of the marine red alga *Grateloupia angusta* (Halymeniales). *Mitochondrial DNA.* 25:269–270.
- Koonin E V. 2010. The origin and early evolution of eukaryotes in the light of

phylogenomics. *Genome Biol.* 11:209.

Ku C, Nelson-Sathi S, Roettger M, Garg S, et al. 2015. Endosymbiotic gene transfer from prokaryotic pangenomes: Inherited chimerism in eukaryotes. *Proc. Natl. Acad. Sci.* 112:10139-10146.

Ku C, Nelson-Sathi S, Roettger M, Sousa FL, et al. 2015. Endosymbiotic origin and differential loss of eukaryotic genes. *Nature.* 524:427-432.

Laehnemann D, Borkhardt A, McHardy AC. 2016. Denoising DNA deep sequencing data-high-throughput sequencing errors and their correction. *Brief. Bioinform.* 17:154–179.

Lam DW, Verbruggen H, Saunders GW, Vis ML. 2016. Multigene phylogeny of the red algal subclass Nemaliophycidae. *Mol. Phylogenet. Evol.* 94:730–736.

Lane CE, Archibald JM. 2008. The eukaryotic tree of life: endosymbiosis takes its TOL. *Trends Ecol. Evol.* 23:268–275.

Lang BF, Gray MW, Burger G. 1999. Mitochondrial genome evolution and the origin of Eukaryotes. *Annu. Rev. Genet.* 33:351–397.

Leblanc C, Kloareg B, Loiseaux-de Goër S, Boyen C. 1995. DNA sequence, structure, and phylogenetic relationship of the mitochondrial small-subunit rRNA from the red alga *Chondrus crispus* (Gigartinales, Rhodophytes). *J. Mol. Evol.* 41:196–202.

Le Gall L, Saunders GW. 2010. DNA barcoding is a powerful tool to uncover algal diversity: a case study of the Phylloporaceae (Gigartinales, Rhodophyta) in the Canadian flora. *J. Phycol.* 46:374–389.

Loman NJ et al. 2012. Performance comparison of benchtop high-throughput sequencing platforms. *Nat. Biotechnol.* 30:434–439.

- Martin W et al. 1998. Gene transfer to the nucleus and the evolution of chloroplasts. *Nature*. 393:162–165.
- Martin WF, Garg S, Zimorski V. 2015. Endosymbiotic theories for eukaryote origin. *Phil. Trans. R. Soc. B*. 370:20140330.
- Qiu H et al. 2013. Assessing the bacterial contribution to the plastid proteome. *Trends Plant Sci*. 18:680–687.
- Qiu H, Price DC, Yang EC, Yoon HS, Bhattacharya D. 2015. Evidence of ancient genome reduction in red algae (Rhodophyta). *J. Phycol.* 51:624-636.
- Salomaki ED, Lane CE. 2014. Are all red algal parasites cut from the same cloth? *Acta Soc. Bot. Pol.* 83:369–375.
- Saunders GW. 1993. Gel purification of red algal genomic DNA: An inexpensive and rapid method for the isolation of polymerase chain reaction-friendly DNA. *J. Phycol.* 29:251–254.
- Saunders GW, McDonald B. 2010. DNA barcoding reveals multiple overlooked Australian species of the red algal order Rhodymeniales (Florideophyceae), with resurrection of *Halopeltis* J. Agardh and description of *Pseudohalopeltis* gen. nov. *Botany-Botanique*. 88:639–667.
- Schattner P, Brooks AN, Lowe TM. 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res*. 33:686–689.
- Smith DR, Keeling PJ. 2015. Mitochondrial and plastid genome architecture: Reoccurring themes, but significant differences at the extremes. *Proc. Natl. Acad. Sci.* 112:10177-10184.

- Stiller JW. 2007. Plastid endosymbiosis, genome evolution and the origin of green plants. *Trends Plant Sci.* 12:391–396.
- Stiller JW et al. 2014. The evolution of photosynthesis in chromist algae through serial endosymbioses. *Nat. Commun.* 5:1–7.
- Tamayo DA, Hughey JR. 2016. Organellar genome analysis of the marine red alga *Dasya binghamiae* (Dasyaceae, Rhodophyta) reveals an uncharacteristic florideophyte mitogenome structure. *Mitochondrial DNA Part B.* 1:510–511.
- Tanifuji G, Archibald JM, Hashimoto T. 2016. Comparative genomics of mitochondria in chlorarachniophyte algae: endosymbiotic gene transfer and organellar genome dynamics. *Sci. Rep.* 6:21016.
- Timmis JN, Ayliffe MA, Huang CY, Martin W. 2004. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat. Rev. Genet.* 5:123–35.
- Tucker RP. 2013. Horizontal gene transfer in choanoflagellates. *J. Exp. Zool. B. Mol. Dev. Evol.* 320:1–9.
- Verbruggen H et al. 2010. Data mining approach identifies research priorities and data requirements for resolving the red algal tree of life. *BMC Evol. Biol.* 10:16.
- Yang E et al. 2015. Highly conserved mitochondrial genomes among multicellular red algae of the Florideophyceae. *Genome Biol. Evol.* 7:2394-2406.
- Yang Z. 2007. PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24:1586–1591.
- Yoon HS, Hackett JD, Ciniglia C, Pinto G, Bhattacharya D. 2004. A molecular timeline for the origin of photosynthetic eukaryotes. *Mol. Biol. Evol.* 21:809–818.
- Zimorski V, Ku C, Martin WF, Gould SB. 2014. Endosymbiotic theory for organelle

origins. *Curr. Opin. Microbiol.* 22:38–48.

CHAPTER 4

Molecular Analysis of Parasites in the Choreocolacaceae (Rhodophyta) Reveals a Reduced *Harveyella mirabilis* (Reinsch) F. Schmitz et Reinke Plastid Genome and Supports the Transfer of Genera to the Rhodomelaceae (Rhodophyta)

Eric D. Salomaki¹, Gary W. Saunders², and Christopher E. Lane¹

is formatted for submission to The Journal of Eukaryotic Microbiology

¹ Department of Biological Sciences, University of Rhode Island, Kingston, RI, USA.

² Centre for Environmental and Molecular Algal Research, Department of Biology, University of New Brunswick, Fredericton, NB, E3B 5A3, Canada.

Abstract:

Parasitism is a life strategy that has repeatedly evolved within the Florideophyceae. Until recently, the accepted paradigm of red algal parasite evolution was that parasites arise by first infecting a close relative and, either through host jumping or diversification, adapt to infect more distant relatives. The terms adelphoparasite and alloparasite have been used to distinguish parasites that are closely related to their hosts from those more distantly related to their hosts, respectively. Phylogenetic studies have cast doubt on the utility of these terms as data show that even alloparasites predominately infect with the same family. All adelphoparasites that have been investigated have lost a native plastid and instead hijack and incorporate a copy of their hosts' plastid when packaging spores. In contrast, a highly reduced plastid that has lost all genes involved with photosynthesis was sequenced from the alloparasite *Choreocolax polysiphoniae*, which indicates that it did not pass through an adelphoparasite stage. In this study we investigate whether other species in the Choreocolacaceae, a family of alloparasites, also retains its native plastid, as well as test the hypothesis that alloparasites can arise and subsequently speciate to form monophyletic clades that infect a range of hosts. We present the plastid genome for *Harveyella mirabilis* which, similar to that of *C. polysiphoniae*, has lost genes involved in photosynthesis. The *H. mirabilis* plastid shares more synteny with free-living red algal plastids than that of *C. polysiphoniae*. Phylogenetic analysis identifies a well-supported monophyletic clade of parasites in the Choreocolacaceae, which retain their own plastid genomes, within the Rhodomelaceae. We therefore transfer genera in the Choreocolacaceae to the Rhodomelaceae.

INTRODUCTION:

Since the 1845, when *Microcolax botryocarpa* Hooker & Harvey, became the first formally described parasitic red alga, the known biodiversity of red algal parasites has been steadily increasing (Reinsch 1875; Setchell 1914, 1923; Goff 1982; Preuss, Nelson, and Zuccarello 2017). Red algal parasites exclusively infect other rhodophytes and are predominately unpigmented, appearing as galls or irregular growths on their free-living red algal hosts. Recent counts have identified 121 distinct species of red algal parasites distributed across eight different orders of the Florideophyceae (Salomaki and Lane 2014; Blouin and Lane 2015; Preuss, Nelson, and Zuccarello 2017). Florideophyceae seem to be prone to adopting parasitism, in large part due to their ability to form direct cell-to-cell connections called secondary pit connections between adjacent, non daughter cells (Wetherbee and Quirk 1982; Goff and Coleman 1985). With few exceptions [e.g. *Choreonema thuretii* (Broadwater and LaPointe 1997)], red algal parasites leverage their ability to form secondary pit connections as a means of infecting their host (Goff and Coleman 1985; Zuccarello, Moon, and Goff 2004).

Two main hypotheses have been proposed for the origin of red algal parasites. Setchell (1918) initially proposed that parasites arose as spores from their host, which had mutated to no longer be capable of photosynthesis. Sturch (1926) later proposed that parasites evolved from epiphytes that penetrated and, over time, became nutritionally reliant on the host. Feldmann and Feldmann (1958) later added to Sturch's hypothesis by suggesting that epiphytes that are closely related to their hosts

are more likely to succeed in forming secondary pit connections and therefore increase the likelihood of successfully establishing a parasitic relationship.

Despite their diminutive habit, parasitic red algae share morphological characteristics with other close relatives, and thus were assigned to tribes or families at the time of their initial discovery (Reinsch 1875, Feldmann and Feldmann 1958). Historically parasites that have infected close relatives have been considered ‘adelphoparasites’, while those more distantly related to their hosts are called ‘alloparasites’ (Feldmann and Feldmann 1958; Goff et al. 1996). Approximately 80% of the described red algal parasite diversity is considered to be adelphoparasitic, while the remaining 20% is alloparasitic (Goff 1982; Salomaki and Lane 2014; Blouin and Lane 2015). Based initially on morphological observations, and later coupled with molecular data, it was proposed that parasites evolve sympatrically with their hosts, as adelphoparasites, and over time diversify or adapt to infect new and more distantly related hosts, becoming alloparasites (Feldmann and Feldmann 1958; Goff, Ashen, and Moon 1997).

Sturch (1926) initially described the family Choreocolacaceae as a family in the Gigartinales, which consisted of morphologically reduced parasites lacking chlorophyll including members of the genera *Choreocolax*, *Harveyella*, and *Holmsella*. The Choreocolacaceae was the subject of a thorough phylogenetic analysis of alloparasites to confirm whether the family was truly a monophyletic clade of parasitic red algae (Zuccarello, Moon, and Goff 2004). This study supported previous morphological observations (Fredericq and Hommersand 1990) confirming that the

genus *Holmsella* was a member of the Gracilariaceae and questioned the legitimacy of recognizing a family of red algal parasites.

In the few other cases where molecular tools have been applied to assess evolutionary histories of red algal parasites, data suggests that red algal parasites arise through independent evolutionary events (Goff et al. 1996; Goff, Ashen, and Moon 1997; Kurihara et al. 2010). In addition to phylogenetic analyses, molecular tools have also been applied to investigate the parasite-host dynamics throughout parasite development. Analyses of the adelphoparasites *Gardneriella tuberifera* Kylin, *Gracilariophila oryzoides* Setchell & H. L. Wilson demonstrated that, although the parasites maintain a native mitochondrion, they have lost their native plastid and instead ‘hijack’ a host plastid when packaging their own spores (Goff and Coleman 1995). To date, all red algal parasites examined maintain a fully functional mitochondrion (Salomaki and Lane 2016). All adelphoparasites that have been investigated have lost their native plastid (Goff and Coleman 1995; Salomaki and Lane 2014; Salomaki and Lane unpublished). In contrast, a highly reduced native plastid was sequenced from the alloparasite *Choreocolax polysiphoniae* Reinsch, which has lost genes involved in photosynthesis, yet maintains functions including fatty acid and amino acid biosynthesis (Salomaki, Nickles, and Lane 2015). The lack of plastids in adelphoparasites, in combination with finding a native plastid in the alloparasite *C. polysiphoniae*, demonstrates that not all parasites pass through an adelphoparasite stage and that there are multiple paths to parasitism in red algae.

In their study examining relationships in the Choreocolacaceae, Zuccarello *et al.* (2004) found that *Holmsella pachyderma* and *Holmsella australis* form a

monophyletic clade within the Gracilariaceae. Additionally they identify that parasite genera *Choreocolax*, *Harveyella*, and *Leachiella* are members of the Rhodomelaceae, though have fairly low support for their relationships to each other. Using molecular data we investigate the relationships of *Choreocolax*, *Harveyella*, and *Leachiella*, and test the hypothesis that alloparasites can arise and subsequently speciate, forming monophyletic clades that infect a range of hosts. Furthermore, we set out to determine if another alloparasite besides *C. polysiphoniae* retains its own native plastid.

MATERIALS AND METHODS

Sample Collection and DNA Extraction

Choreocolax polysiphoniae, found on its host, *Vertebrata lanosa* (Linneaus) T.A. Christensen, was collected at Beavertail State Park, Jamestown, RI, USA. *Harveyella mirabilis* (Reinsch) F.Schmitz & Reinke, on its host, *Odonthalia washingtoniensis* Kylin, was collected from Cattle Point, Friday Harbor, WA, USA (48.452428, -122.962774), and *Leachiella pacifica* Kugrens, on *Polysiphonia hendryi* N.L. Gardner, was collected off the dock at Friday Harbor Laboratories, Friday Harbor WA, USA (48.545066, -123.012268). Individual parasite galls from *Harveyella mirabilis* and *Leachiella pacifica* were excised from their respective hosts and collected in a 1.5 mL microcentrifuge tube. The parasite tissue was hand-ground using a Corning Axygen® PES-15-B-SI disposable tissue grinder pestle in a 1.5 mL microcentrifuge tube while submerged in 100µL of DNA extraction buffer (Saunders 1993). DNA was extracted from specimens using a standard phenol/chloroform extraction with all ratios adjusted for an initial buffer volume of 100µL (Saunders

1993). Additionally, lyophilized DNA from specimen GWS021225 (Saunders 2014) was acquired from the Saunders lab and rehydrated in 20 μ L of 5mM Tris/HCl pH 8.5.

Molecular analyses

A 664 bp fragment at the 5-prime end of the mitochondrial cytochrome oxidase 1 gene (COI-5P), which has been used extensively for barcoding red algal species (Saunders 2005), was PCR amplified using the GWSFn (Le Gall and Saunders 2010) and GWSRx (Saunders and Moore 2013) primer pair according to protocols in Saunders and Moore (2013). All PCR reactions were completed with Ex TaqTM DNA Polymerase (Takara Bio USA, Mountainview, CA, USA) following the manufacturer's recommendations. The BigDye Terminator v3.1 Cycle Sequencing Kit (PE Applied Biosystems [ABI], Foster City, CA, USA) was used for the sequencing of PCR products, and reactions were analyzed using an ABI Prism 3130XL genetic analyzer. Sequences were trimmed and assembled using Geneious v. 9.1.5. The consensus sequences resulting from the assembly were BLAST searched against the NCBI GenBank database and the Barcode Of Life Database (boldsystems.org) to find similar publicly available sequences.

In addition to PCR amplification of the COI-5P, genomic DNA was amplified from all specimens using the illustra Single Cell GenomiPhi DNA Amplification kit (GE Healthcare Life Sciences, Pittsburgh, Pa) according to manufacturer protocols. Whole-genome amplification of an individual *C. polysiphoniae* parasite was completed using the illustra GenomiPhi Hy DNA Amplification Kit according to manufacturer protocols. The COI-5P was PCR amplified and sequenced, as previously

described, from the whole-genome amplified DNA to screen for contamination resulting from the amplification process. Libraries for Illumina sequencing were constructed from the amplified DNA on the Apollo 324 robot using the PrepX ILM DNA Library Kit (Wafergen Biosystems, Fremont, CA, USA). The *Harveyella mirabilis*, *Leachiella pacifica*, and GWS021225 libraries were multiplexed and sequenced on an Illumina MiSeq paired-end 250 x 250 basepair run. The *C. polysiphoniae* library was sequenced on separate full-cell Illumina MiSeq paired-end 300 x 300 basepair run. The sequencing effort resulted in 29,355,470 paired-end reads for *C. polysiphoniae*, 7,923,094 paired end reads for *Harveyella mirabilis*, 7,502,360 paired-end reads for *Leachiella pacifica*, and 8,869,740 paired end reads for GWS021225. For all datasets, sequences with PHRED scores <30 were removed and the remaining reads were trimmed of adapter sequences. Additionally, fifteen 5' and five 3' nucleotides were trimmed from the remaining reads and all reads under 100 nucleotides were removed from the dataset. All trimming was completed using CLC Genomics Workbench v. 9.5.2 (CLC Bio-Qiagen, Aarhus, Denmark) and the remaining reads were assembled using default parameters in CLC Genomics Workbench v. 9.5.2.

Phylogenetic Analysis

A total of 823 Rhodomelaceae COI-5P sequences available on GenBank were downloaded into Geneious v. 9.1.5 (Kearse et al. 2012) and aligned with the data generate by PCR for *Harveyella mirabilis* and *Leachiella pacifica*. This alignment was subjected to maximum likelihood analysis using RAxML with the GTR+I+G model

with 1,000 bootstrap replicates. Based upon the results of the preliminary tree, 18S, 28S, COI-5P, and *rbcL* data were compiled from GenBank for thirteen of the closest relatives to a weakly supported clade containing *Choreocolax polysiphoniae* Reinsch, *Harveyella mirabilis*, *Leachiella pacifica*, and GWS021225 (Table 1). The longest sequence for each gene from this compiled dataset was used as a query against local BLAST databases created from the *C. polysiphoniae*, *H. mirabilis*, *L. pacifica*, and GWS021225 assemblies, (Table 1). The 18S and 28S regions were extracted from the ribosomal cistrons identified in each assembly based upon its alignment to the query gene.

Each dataset was aligned using the MAFFT plugin in Geneious v. 9.1.5 and subsequently concatenated using Sequence Matrix v. 1.7.8 (Vaidya, Lohman, and Meier 2011). The concatenated dataset was then subjected to phylogenetic analysis using maximum likelihood, implemented in RAxML v. 8.2.2 (Stamatakis 2014) and Bayesian Inference using MrBayes v. 3.2.2 (Ronquist et al. 2012). For the maximum likelihood analysis the GTR+I+G model was used and bootstrap support values were calculated using 1,000 bootstrap replicates. For the Bayesian inference analysis, two Metropolis-coupled Markov chain Monte Carlo (MCMCMC) runs consisting of one cold chain and three hot chains were performed. Each run was sampled every 100 generations for 1,010,000 generations. After confirming the runs converged by checking to ensure that the average standard deviation of split frequencies was below 0.01, the trees were merged. The resulting tree and posterior probabilities were calculated from the 20,202 trees generated.

Plastid Genome Annotation

A 90,654 bp contig was identified as the plastid genome of *Harveyella mirabilis* from the previously described assembled Illumina MiSeq data. Open reading frame (ORF) prediction on the *H. mirabilis* plastid was done in Geneious Pro v6.1 and the resulting ORFs were manually annotated using GenBank and Pfam (Finn et al. 2010, 2015) databases. Functional annotations were assigned from the UniProt (The UniProt Consortium 2017) and KEGG databases (Kanehisa et al. 2016). Genes found in red algal plastid genomes that were missing from the *H. mirabilis* plastid were searched for using BLAST, against the plastid sequence and the genomic assemblies to verify their absence and check for evidence of transfer to another genetic compartment. The plastid genome sequence was submitted to the tRNAscan-SE online server v1.21 (Schattner, Brooks, and Lowe 2005) for identification of tRNA sequences and to MFannot (<http://megasun.bch.umontreal.ca/cgi-bin/mfannot/mfannotInterface.pl>) to identify rRNA sequences and confirm manual gene annotations.

Plastid Genome Comparative Analysis

The circular plastid genomes of the free-living florideophytes, *Calliarthron tuberculosum* (KC153978), *Ceramium cimbricum* (KR025491), *Ceramium japonicum* (KX284719), *Chondrus crispus* (HF562234), *Dasya binghamiae* (KX247284), *Laurencia sp.* JFC0032 (LN833431), *Grateloupia taiwanensis* (KC894740), *Gracilaria tenuistipitata* (AY673996), *Vertebrata lanosa* (KP308097), and the alloparasites *Choreocolax polysiphoniae* (KP308096) and *Harveyella mirabilis*

(XXXXXXXX) were arranged so their sequences began with the *ftsH* gene. Whole genome alignment was completed using the default settings for the progressiveMauve algorithm in the Mauve v2.3.1 (Darling et al. 2004).

RESULTS

Identification of a Cryptic Parasite

A maximum likelihood phylogenetic analysis of 823 Rhodomelaceae COI-5P sequences was completed using RAxML. This analysis provided weak support (bootstrap support of 23) for a clade containing the alloparasites *Choreocolax polysiphoniae*, *Harveyella mirabilis*, *Leachiella pacifica*, as well as 2 sequences labeled *Rhodomela* sp1Cal voucher GWS021225 (KM254767) and *Rhodomela* sp1Cal voucher GWS021347 (KM254267). Interestingly, these two GenBank sequences had bootstrap support of 100 as being sister to the sequence of *Leachiella pacifica* generated in this study (data not shown). Based upon these results, the specimens were reexamined and parasite galls were found on the host, which was subsequently identified as *Polysiphonia paniculata* based upon *rbcL* sequence data.

Monophyletic Alloparasite Clade

Although identification of parasites on the samples GWS021225 and GWS021347 was a direct result of resolving a clade of parasite sequences, bootstrap support for a monophyletic clade of alloparasites was weak and a subsequent Bayesian phylogeny failed to recover the same clade (data not shown). Based upon this original

analysis, COI-5P data from a reduced dataset of the parasites and 28 of their closest relatives was subjected to maximum likelihood analysis, which again recovered a poorly supported monophyletic alloparasite clade (data not shown). To resolve the issue of low statistical support, thirteen taxa that were continually allied to the alloparasite clade were selected based upon availability of published comparative data. Phylogenetic analysis was conducted on a concatenated dataset comprised of 18S, 28S, COI-5P and *rbcL* (Table 1) from the 13 free-living species, and the alloparasites *C. polysiphoniae*, *H. mirabilis*, *L. pacifica*, and the parasite on GWS021225 (from here on called *Leachiella* sp.). With the additional data beyond the COI-5P, both maximum likelihood and Bayesian inference recovered a monophyletic clade of red algal alloparasites with 75% bootstrap support and 0.99 posterior probability (Figure 1). *Choreocolax polysiphoniae* was the earliest branching parasite followed by *Harveyella mirabilis*, with *Leachiella pacifica* and *Leachiella* sp. being recovered as sister taxa with 100% bootstrap support/1.0 posterior probability (Figure 1).

Taxonomic Considerations

Choreocolax polysiphoniae was initially described by Reinsch (1875) from the Atlantic coast of North America as parasitic on *Polysiphonia fastigiata* (now *Vertebrata lanosa*). Specimens used the phylogenetic analysis presented here (Figure 1) were collected at Beavertail State Park in Jamestown, RI, USA infecting *Vertebrata lanosa* and are a strong match to the type description (Reinsch 1875). As the type collection cannot be located, we formally lectotypify *C. polysiphoniae* on image #49 accompanying the description in Reinsch (1875). Based upon the molecular analyses

here, which resolved a monophyletic clade of parasitic red algae within the Rhodomelaceae (Figure 1), we formally transfer *Choreocolax* to the Rhodomelaceae and recognize Choreocolacaceae as a synonym of this family. To adhere to the principle of monophyly, the genus *Harveyella*, based on the type and only species *H. mirabilis* and included in our analyses (Figure 1), is also transferred to the Rhodomelaceae.

Harveyella mirabilis Plastid Genome

The plastid genome of *Harveyella mirabilis* was assembled as a 90,654 kb circular molecule with 322x coverage. The plastid genome has an overall AT content of 76.5% and contains 84 protein coding genes, 3 rRNAs, and 23 tRNAs (Figure 2). Similar to the *Choreocolax polysiphoniae* plastid (Salomaki, Nickles, and Lane 2015), all genes related to photosynthesis have been lost with the exception of *petF* which has been demonstrated to be involved in electron transport in other metabolic pathways (Happe and Naber 1993; Jacobs et al. 2009). Genes involved in transcription/translation and fatty acid, amino acid, protein, isoprene biosynthesis remain conserved. As in the *C. polysiphoniae* plastid, *gltB* appears to be a vanishing pseudogene. BLAST similarity searches are able to find conserved homology, however the presence of stop-codons throughout the region suggests that the gene is likely no longer capable of being completely translated.

Plastid Genome Comparisons

A whole genome MAUVE alignment of the *H. mirabilis* with *C. polysiphoniae* and

nine representative free-living Florideophyceae plastid genomes identified 13 locally collinear blocks in the *H. mirabilis* genome that aligned with the free-living plastids (Figure 3). There were no rearrangements or inversions in the *H. mirabilis* plastid genome when compared to photosynthetic Rhodomelaceae taxa. When aligning *H. mirabilis* to the *C. polysiphoniae* plastid genome, 11 locally collinear blocks are identified, and several genome inversions and rearrangements are evident (Figure 4). Additionally, gene content varies slightly between the two parasite plastid genomes. *Harveyella mirabilis* retains *argB*, *carA*, *infB*, *rnz*, *rpl9*, *rpl17*, *rpl24*, *rpl32*, *rpl33*, *rpl34*, *rpoZ*, *rps18*, *rps20*, *ycf21*, which have all been lost in *C. polysiphoniae* (Table 2), while *C. polysiphoniae* maintains copies of *dnaB* and *fabH* which have both been lost in the *H. mirabilis* plastid genome.

DISCUSSION

Origin of Parasites

With 121 described parasites occurring across eight Florideophyceae orders (Salomaki and Lane 2014; Blouin and Lane 2015; Preuss, Nelson, and Zuccarello 2017), red algae appear more able to transition from autotrophy to parasitic lifestyles than any other eukaryotic lineage. The terms adelphoparasite and alloparasite have traditionally been used to describe parasites that infect hosts within their tribe/family, or in different tribes/families, respectively. The use of these terms has been questioned as molecular data have revealed that alloparasites, like adelphoparasites, infect close relatives rather than distantly related species (Zuccarello, Moon, and Goff 2004; Kurihara et al. 2010). Defining red algal parasites using taxonomy alone is no longer

appropriate (See conclusion chapter). Phylogenetic analysis using data presented here place *Choreocolax polysiphoniae*, *Harveyella mirabilis*, and two species of *Leachiella* firmly within the same family as their hosts further supporting abandonment of these terms for differentiating two types of red algal parasites.

The Demise of Exclusively Parasitic Families

Sturch (1926) initially described the Choreocolacaceae as a family of holoparasites, containing the genera *Choreocolax*, *Harveyella*, and *Holmsella*. However, more recent morphological investigation considered that the genus *Holmsella* was related to the parasites *Gelidiocolax* and *Pterocladophila*, and it was moved to the family Pterocladophilaceae in the Gracilariales (Fredericq and Hommersand 1990). Their observations were subsequently supported by molecular data generated with the specific aim of testing the phylogenetic affinities of parasites that Sturch had assigned to the Choreocolacaceae. This work demonstrated that *Holmsella australis* and *Holmsella pachyderma* formed a well-supported monophyletic clade within the Gracilariaceae (Zuccarello, Moon, and Goff 2004). Their molecular data also indicated that *Choreocolax* and *Harveyella* belong in the Ceramiales, leading the authors to question whether the Choreocolacaceae should continue to be recognized (Zuccarello, Moon, and Goff 2004). However, their use of 18S sequence data was insufficient to resolve the issue of a monophyletic clade for these parasites. Other authors have also noted that species in *Choreocolax*, *Harveyella*, and *Leachiella* have features aligning them to the Ceramiales, but again, taxonomic affinities among the parasite species and within this order remained uncertain (Goff

and Cole 1975; Kugrens 1982; Fredericq and Hommersand 1990). Our data confirm the findings of Zuccarello, Moon, and Goff (2004), placing *Choreocolax*, *Harveyella* and *Leachiella*, within the Rhodomelaceae. Furthermore, the phylogeny utilizing additional molecular markers provides strong support for a monophyletic clade containing *Choreocolax*, *Harveyella* and *Leachiella* (Figure 1), and supports the placing the Choreocolacaceae in synonymy with the Rhodomelaceae.

Cryptic Species

Until recently, the dogma behind red algal parasite evolution was the notion that parasites arise sympatrically as adelphoparasites and over time evolve and adapt to infect more distant hosts becoming alloparasites (Goff et al. 1996; Goff, Ashen, and Moon 1997). Based on molecular data demonstrating that alloparasites also infect members of their own family (Zuccarello, Moon, and Goff 2004; Kurihara et al. 2010), we now consider that these terms are not suitable for distinguishing parasites. The recognition of multiple monophyletic parasite clades, the *Holmsella* clade (Zuccarello, Moon, and Goff 2004), and the Rhodomelaceae clade uncovered here containing *Choreocolax*, *Harveyella*, and *Leachiella* (Figure 1), supports the idea of a sympatric origin of alloparasites with subsequent speciation as parasites adapt to their hosts, without passing through an adelphoparasite-like stage.

Previously, *Leachiella pacifica* was described from multiple hosts including members of the genus *Polysiphonia* and *Pterosiphonia* (Kugrens 1982; Zuccarello and West 1994). In their study on *L. pacifica* host specificity, Zuccarello and West (1994) found that parasite spores isolated from one host genus were unable to infect members

of the other host genus. Kugrens initially designated the *Leachiella pacifica* type specimen as a parasite infecting *Polysiphonia* spp. from Cattle Point, Friday Harbor, WA, USA (Kugrens 1982). Based upon this, here we are considering the specimen collected on *Polysiphonia hendryi* from Friday Harbor, WA, USA as true *Leachiella pacifica*. Subsequently, parasites infecting *Polysiphonia paniculata* (and several other species) have also been attributed to *L. pacifica* (Zuccarello and West 1994; Zuccarello, Moon, and Goff 2004). Molecular data here identify two distinct species of *Leachiella* that infect separate hosts but are otherwise difficult to distinguish. Due to the highly reduced morphology of these parasites, it seems likely that host identity may be the easiest way to identify parasites in the genus *Leachiella* to species.

Some parasitic red algae are reportedly capable of infecting a range of hosts [e.g., *Asterocolax* – see Goff, Ashen, and Moon (1997)], however molecular data are contradicting that notion. Rather, it appears that red algal parasite species have higher host specificity than previously believed, and instead, we have underestimated the amount of parasite diversity as a result of their reduced morphology. Morphology based studies report *Leachiella pacifica* infecting at least seven unique host species (Zuccarello and West 1994), further molecular analyses will surely uncover additional species in the genus *Leachiella*.

Parasite Plastids

The link between plastid origin in red algal parasites and their evolutionary relationship to their hosts may be central to our understanding of red algal parasite evolution. The *Harveyella mirabilis* plastid genome (Figure 2) represents the second

red algal parasite demonstrated to retain a reduced native plastid. Similar to *Choreocolax polysiphoniae*, the *H. mirabilis* plastid genome remains conserved for functions including amino acid, fatty acid, and protein biosynthesis, but has lost genes involved in building the light harvesting apparatus, photosystems I and II, and other photosynthesis related genes. Identifying a monophyletic clade of parasites (Figure 1) in which the two earliest branching members identified so far retain plastids, strongly suggests that plastids are also retained in species of *Leachiella*. Preliminary data generated in the Lane lab supports that hypothesis though work remains to completely sequence plastids from *Leachiella* species.

The *H. mirabilis* plastid gene order, with the exception of the missing photosynthesis genes, is conserved when compared to plastids of free-living Rhodomelaceae species (Figure 3). In contrast, we find that the *C. polysiphoniae* plastid has undergone greater gene loss than the *H. mirabilis* plastid, and a substantial amount of genome reorganization (Figure 4). *Harveyella mirabilis* retains *argB* and *carA*, which are involved in arginine biosynthesis processes, *rpoZ*, which promotes RNA polymerase assembly, nine genes involved in building ribosomes, and an uncharacterized hypothetical protein, all of which have been lost from the *C. polysiphoniae* plastid (Table 2). However, *C. polysiphoniae* does retain copies of *dnaB*, which is involved in DNA replication, and *fabH*, which is involved in fatty acid biosynthesis, both of which have been lost in *H. mirabilis*. Analysis of additional plastid genomes in this clade will provide greater insights into patterns of plastid genome evolution in red algal parasites.

Two competing hypotheses have been proposed for the origin of red algal parasites. Setchell (1918) suggested that a mutation in a spore causes parasites to arise sympatrically, while Sturch (1926) postulated that parasitic red algae start out as epiphytes that over time become endophytes and increasingly rely on the host for nutrition. It seems plausible that Setchell's origin hypothesis could explain the rise of the so-called adelphoparasites, which comprise the majority of known red algal parasite biodiversity. By evolving from their hosts, adelphoparasites could easily incorporate a copy of a genetically similar plastid as their own. Replacing a functionally reduced or non-functional native plastid with a host-derived plastid seems to be an easy mechanism for survival. Unfortunately, reliance on the host plastid also starts the newly evolved parasite down a path towards inevitable extinction. In order for a non-photosynthetic parasite to survive it still must retain compatibility for other plastid functions including amino acid and fatty acid biosynthesis. Because the host-derived plastid is newly acquired during each new infection, the host plastid experiences one set of evolutionary pressures while the parasite evolves and accumulates mutations of its own. Eventually the parasite will inevitably lose the ability to communicate with the host plastid as the parasite and host increasingly become genetically distinct. This leaves the parasite with two possibilities for survival, either to find another host with a compatible plastid or go extinct. Alternatively, the success of an alloparasite may be explained by Sturch's hypothesis. By evolving from a closely related epiphyte that is able to create secondary pit connections, the parasite may retain its own plastid and therefore enable its longevity and the opportunity to even speciate as the parasite adapts to new hosts. Therefore, what were previously

viewed as competing hypotheses to explain the evolution of red algal parasites, may each explain how different types of parasites arise.

Although organisms have transitioned from photosynthesis to other methods of nutrient acquisition numerous times across the tree of life, support for total plastid loss is rare (Gornik et al. 2015) with most cases resulting in highly reduced plastids retained for core functions (e.g. isoprenoid and fatty acid biosynthesis) (Wolfe, Morden, and Palmer 1992; Knauf and Hachtel 2002; de Koning and Keeling 2006) or in extreme cases the retention of the apicoplast in Apicomplexans. Red algal parasites traditionally considered adelphoparasites, like *Gracilariophila oryzoides* and *Gardneriella tuberifera*, still require a copy of the host plastid for some function that, to date, remain undetermined. Interestingly, in the *Gracilariophila oryzoides* genome and transcriptome, photosynthesis related nuclear-encoded plastid-targeted genes remain conserved while those same genes are absent from the *Choreocolax polysiphoniae* transcriptome (Salomaki and Lane, unpublished).

Data presented here further highlight the need to abandon the notion that taxonomy defines alloparasites. By placing the family Choreocolacaceae in synonymy with the Rhodomelaceae we are making steps to remove the artificial appearance of parasites infecting hosts in different families. Furthermore, we recognize that a major distinction between types of red algal parasites is the origin of the parasite plastids. While some parasites, including *Harveyella mirabilis* and *Choreocolax polysiphoniae*, retain a native plastid that evolves in concert with the parasite, others including *Gracilariophila oryzoides* and *Gardneriella tuberifera*, incorporate a copy of a host plastid when packaging spores, but will inevitably evolve and become incompatible

with the host plastid and blink out of existence. It is predicted that investigations of parasites that have traditionally referred to as alloparasites, like *Holmsella* and *Gelidiocolax*, will also provide evidence of plastid retention and monophyletic clades of parasites.

ACKNOWLEDGEMENTS

The authors thank Amanda Savoie for help identifying parasite sequence data in the UNB database, Jillian Freese for collecting specimens of *Harveyella mirabilis* and *Leachiella pacifica*, and Kristina Terpis for her support in the lab. Funding for this work was provided to by a Phycological Society of America Grant-in-Aid of Research to ES and by grant #1257472 from the National Science Foundation awarded to CL. This research is based in part upon work conducted using the Rhode Island Genomics and Sequencing Center which is supported in part by the National Science Foundation (MRI Grant No. DBI-0215393 and EPSCoR Grant Nos. 0554548 & EPS-1004057), the US Department of Agriculture (Grant Nos. 2002-34438-12688 and 2003-34438-13111), and the University of Rhode Island.

LITERATURE CITED

Blouin N. & Lane C. 2015. Red algae provide fertile ground for exploring parasite evolution. *Perspect. Phycol.* 3:11–19.

Broadwater S. T. & LaPointe E. a. 1997. Parasitic interactions and vegetative ultrastructure of *Choreonema thuretii* (Corallinales, Rhodophyta). *J. Phycol.*, 33:396–407.

- Darling A. C. E., Mau B., Blattner F. R. & Perna N. T. 2004. Mauve : Multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.*, 14:1394–1403.
- Feldmann G. & Feldmann J. 1958. Recherches sur quelques Floridees parasites. *Rev. Gen. Bot.*, 65:49–128.
- Finn R. D., Cogill P., Eberhardt R. Y., Eddy S. R., Mistry J., Mitchell A. L., Potter S. C., Punta M., Qureshi M., Sangrador-Vegas A., et al. 2015. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* 44:D279–D285.
- Finn R. D., Mistry J., Tate J., Cogill P., Heger A., Pollington J. E., Gavin O. L., Gunasekaran P., Ceric G., Forslund K., et al. 2010. The Pfam protein families database. *Nucleic Acids Res.* 38:D211-22.
- Fredericq S. & Hommersand M. 1990. Morphology and systematics of *Holmsella pachyderma* (Pterocladophilaceae, Gracilariales). *Br. Phycol. J.*, 25:39–51.
- Le Gall L. & Saunders G. W. 2010. DNA barcoding is a powerful tool to uncover algal diversity: A case study of the Phylloporaceae (Gigartinales, Rhodophyta) in the Canadian flora. *J. Phycol.*, 46:374–389.
- Goff L. J. 1982. The biology of parasitic red algae. *Prog. Phycol. Res.*, 1:289–369.
- Goff L. J., Ashen J. & Moon D. 1997. The evolution of parasites from their hosts: A case study in the parasitic red alga. *Evolution.* 51:1068–1078.
- Goff L. J. & Cole K. 1975. The biology of *Harveyella mirabilis* (Cryptonemiales, Rhodophyceae). II . Carposporophyte development as related to the taxonomic affiliation of the parasitic red alga, *Harveyella mirabilis*. *Phycologia*, 14:227–238.

- 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*. 7:1899–1911.
- 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.
- Gornik S. G., Febrimarsa, Cassin A. M., MacRae J. I., Ramaprasad A., Rchiad Z., McConville M. J., Bacic A., McFadden G. I., Pain A., et al. 2015. Endosymbiosis undone by stepwise elimination of the plastid in a parasitic dinoflagellate. *Proc. Natl. Acad. Sci.* 112:5767-5772.
- Happe T. & Naber J. D. 1993. Isolation, characterization and N-terminal amino acid sequence of hydrogenase from the green alga *Chlamydomonas reinhardtii*. *Eur. J. Biochem.* 214:475–81.
- Jacobs J., Pudollek S., Hemschemeier A. & Happe T. 2009. A novel, anaerobically induced ferredoxin in *Chlamydomonas reinhardtii*. *FEBS Lett.* 583:325–9.
- Kanehisa M., Sato Y., Kawashima M., Furumichi M. & Tanabe M. 2016. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.*, 44:D457–D462.
- Kearse M., Moir R., Wilson A., Stones-Havas S., Cheung M., Sturrock S., Buxton S., Cooper A., Markowitz S., Duran C., et al. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28:1647–1649.

- Knauf U. & Hachtel W. 2002. The genes encoding subunits of ATP synthase are conserved in the reduced plastid genome of the heterotrophic alga *Prototheca wickerhamii*. *Mol. Genet. Genomics* 267:492–7.
- de Koning A. P. & Keeling P. J. 2006. The complete plastid genome sequence of the parasitic green alga *Helicosporidium* sp. is highly reduced and structured. *BMC Biol.* 4:12.
- Kugrens P. 1982. *Leachiella pacifica*, Gen. et Sp. nov., a new parasitic red alga from Washington and California. *Am. J. Bot.*, 123:306–319.
- 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 *Ululania* (Ceramiales). *J. Phycol.* 46:580–590.
- 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.* 60:1–13.
- Reinsch P. F. 1875. *Contributiones ad Algologiam et Fungologiam*. (Weigel T. O. (ed.)). Nuremberg, Lipsiae.
- Ronquist F., Teslenko M., van der Mark P., Ayres D. L., Darling A., Höhna S., Larget B., Liu L., Suchard M. a & Huelsenbeck J. P. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61:539–42.
- Salomaki E. D. & Lane C. E. 2014. Are all red algal parasites cut from the same cloth? *Acta Soc. Bot. Pol.* 83:369–375.

- Salomaki E. D. & Lane C. E. 2016. Red algal mitochondrial genomes are more complete than previously reported. *Genome Biol. Evol.* 48:evw267.
- Salomaki E. D., Nickles K. R. & Lane C. E. 2015. The ghost plastid of *Choroecolax polysiphoniae*. *J. Phycol.*, 51:217–221.
- Saunders G. W. 1993. Gel purification of red algal genomic DNA: An inexpensive and rapid method for the isolation of polymerase chain reaction-friendly DNA. *J. Phycol.*, 29:251–254.
- Saunders G. W. 2005. Applying DNA barcoding to red macroalgae: a preliminary appraisal holds promise for future applications. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 360:1879–88.
- Saunders G. W. 2014. Long distance kelp rafting impacts seaweed biogeography in the Northeast Pacific: the kelp conveyor hypothesis. *J. Phycol.* 50:968–974.
- Saunders G. W. & Moore T. E. 2013. Refinements for the amplification and sequencing of red algal DNA barcode and RedToL phylogenetic markers: a summary of current primers, profiles and strategies. *Algae* 28:31–43.
- Schattner P., Brooks A. N. & Lowe T. M. 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res.*, 33:686–689.
- Setchell W. A. 1914. Parasitic Florideae I. *Univ. Calif. Publ. Bot.*, 6:1–34.
- Setchell W. A. 1923. Parasitic Florideae II. *Univ. Calif. Publ. Bot.*, 10:393–401.
- Stamatakis A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, **30**:1312–1313.

- The UniProt Consortium. 2017. UniProt: the Universal Protein knowledgebase. *Nucleic Acids Res.* 45:D158–D169.
- Vaidya G., Lohman D. J. & Meier R. 2011. SequenceMatrix: Concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics*, **27**:171–180.
- Wetherbee R. & Quirk H. 1982. The fine structure of secondary pit connection formation between the red algal allopasite *Holmsella australis* and its red algal host *Gracilaria furcellata*. *Protoplasma*, 176:166–176.
- Wolfe K. H., Morden C. W. & Palmer J. D. 1992. Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. *Proc. Natl. Acad. Sci. U. S. A.* 89:10648–52.
- Zuccarello G. C., Moon D. & Goff L. J. 2004. A phylogenetic study of parasitic genera placed in the family Choreocolacaceae (Rhodophyta). *J. Phycol.* 40:937–945.
- Zuccarello G. C. & West J. A. 1994. Genus and race specificity in the red algal parasite *Leachiella pacifica* (Choreocolacaceae, Rhodophyta). *Phycologia*, 33:213–218.

Table 1. GenBank Sequence accession numbers for taxa used in phylogenetic analyses (Figure 2).

	18S	28S	COI-5P	rbcL
<i>Amansia fimbriifolia</i>	HM582913	HQ422079	HM582889	-
<i>Amansia glomerata</i>	HM582909	HQ422225	HQ422913	-
<i>Bostrychia arbuscula</i>	-	AY920894	KM502795	AY920845
<i>Bostrychia intricata</i>	-	KM502850	KM502801	AY920846
<i>Bostrychia vaga</i>	-	KM502842	KM502791	KM502817
<i>Choreocolax polysiphoniae</i>	This Study	This Study	This Study	-
<i>Harveyella mirabilis</i>	This Study	This Study	This Study	-
<i>Heterodasya mucronata</i>	-	-	KC567673	KF367797
<i>Leachiella pacifica</i>	This Study	This Study	This Study	-
<i>Leachiella</i> sp. GWS021225	This Study	This Study	KM254767	-
<i>Neorhodomela larix</i>	AY617140	-	KM254241	GQ252553
<i>Neorhodomela oregona</i>	-	-	KM254304	GQ252556
<i>Odonthalia dentata</i>	-	JX572172	JX571960	KU564463
<i>Odonthalia floccosa</i>	AY617141	-	KM254276	GQ252492
<i>Osmundaria obtusiloba</i>	-	HQ422003	HQ422818	-
<i>Rhodomela confervoides</i>	AY617145	KX145642	KX258842	KX146197
<i>Rhodomela lycopodioides</i>	-	This Study	HM916516	KU564489

Table 2. *Harveyella mirabilis* plastid genes and their function, which have been lost from the *Choreocolax polysiphoniae* plastid genome.

Gene Name	Biological Process and Molecular Function
<i>argB</i>	Amino acid biosynthesis; ATP binding, acetylglutamate kinase activity
<i>carA</i>	Pyrimidine metabolism; ATP binding, carbamoyl-phosphate synthase (glutamine-hydrolyzing) activity
<i>infB</i>	Translation; Translation initiation, GTP binding
<i>rpl9</i>	Translation; structural constituent of the 50S ribosome
<i>rpl17</i>	Translation; structural constituent of the 50S ribosome
<i>rpl24</i>	Translation; structural constituent of the 50S ribosome
<i>rpl32</i>	Translation; structural constituent of the 50S ribosome
<i>rpl33</i>	Translation; structural constituent of the 50S ribosome
<i>rpl34</i>	Translation; structural constituent of the 50S ribosome
<i>rpoZ</i>	Transcription; DNA binding
<i>rps18</i>	Translation; structural constituent of the 30S ribosome
<i>rps20</i>	Translation; structural constituent of the 30S ribosome
<i>ycf21</i>	Uncharacterized hypothetical protein

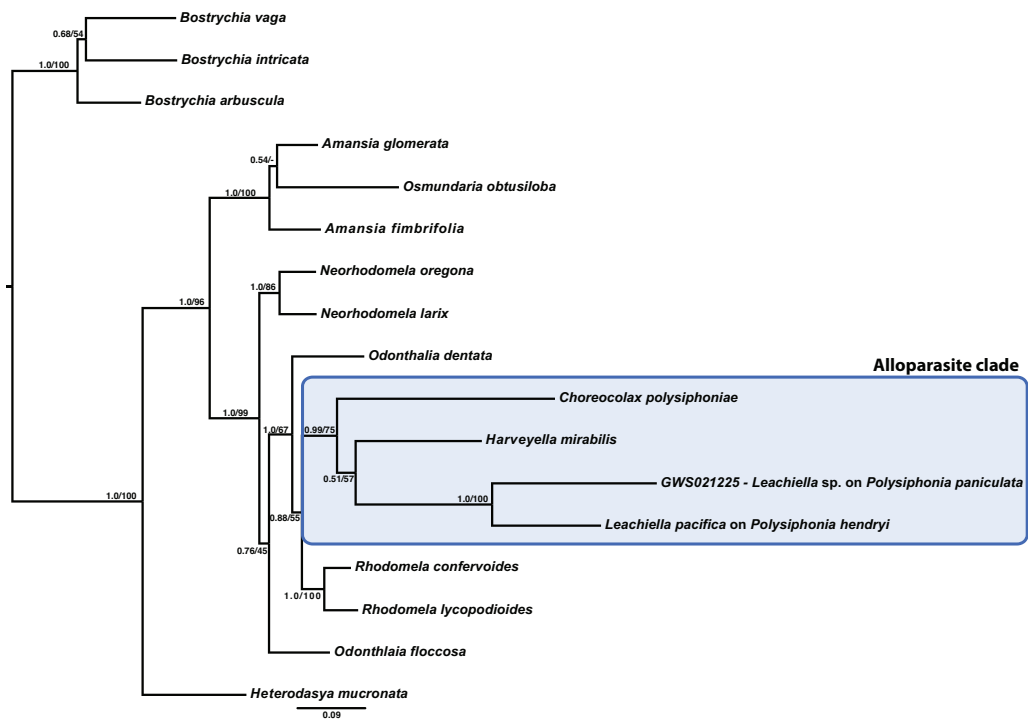


Figure 1. Maximum likelihood phylogeny based on concatenated 18S, 28S, COI-5P, and *rbcL* data as shown in Table 1, of the parasites *Choreocolax polysiphoniae*, *Harveyella mirabilis*, *Leachiella pacifica*, an undescribed species of *Leachiella* found on *Polysiphonia paniculata*, and thirteen close relatives. This phylogenetic analysis supports a monophyletic clade of parasites arising within the Rhodomelaceae. Support values shown as Bayesian posterior probability/maximum likelihood bootstrap.

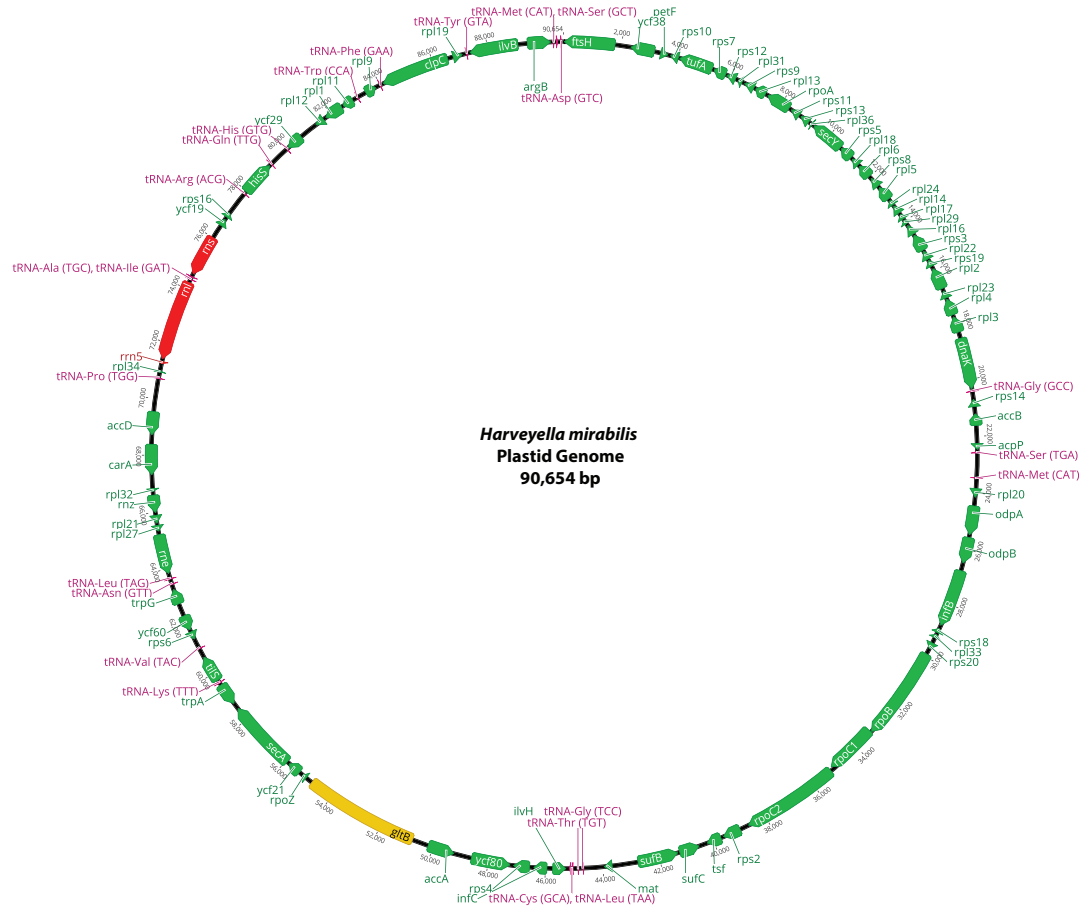


Figure 2. The plastid genome of the parasitic red alga, *Harveyella mirabilis* is 90,654 basepairs and contains 84 protein coding genes (Green), the 5S, 16S, and 23S rRNAs (Red), and 23 tRNAs (Pink). All genes involved with photosynthetic functions, except *petF*, have been lost. The *ftsH* gene is truncated but may still be transcribed, however *gltB* is a non-functional pseudogene (Yellow).

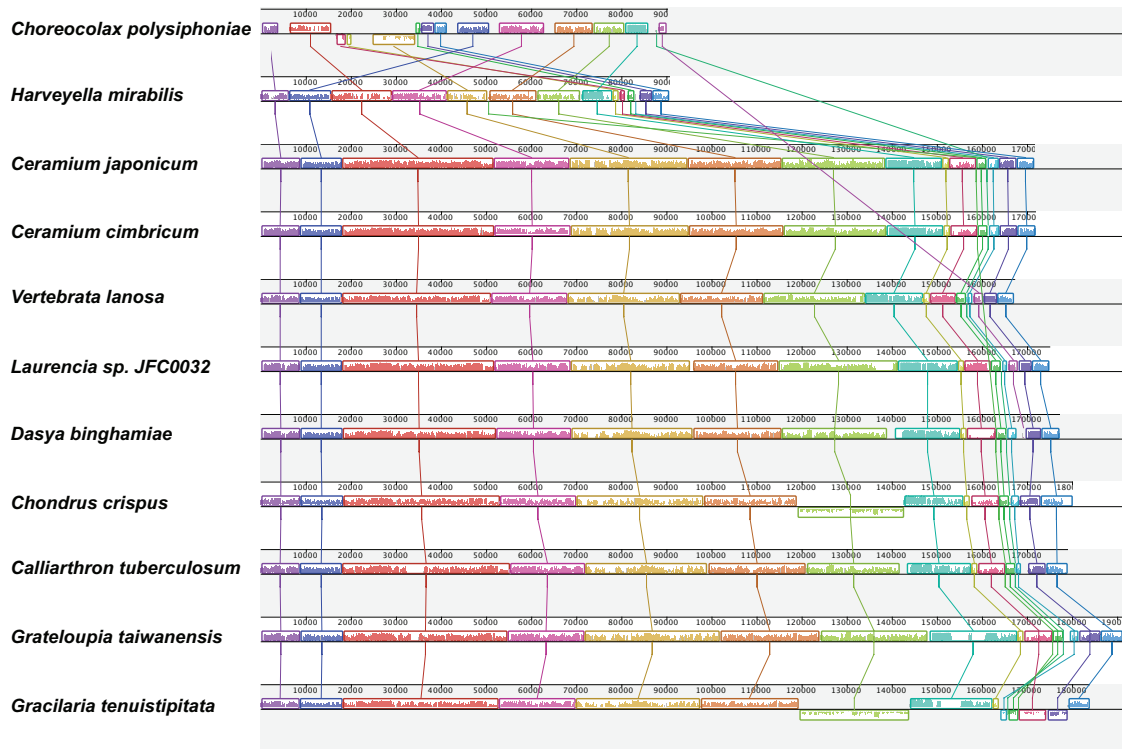


Figure 3. Mauve alignment of the parasites *Choreocolax polysiphoniae* (top) and *Harveyella mirabilis* (second from top) with all published Rhodomelaceae plastid genomes as well as those from select representatives of other Florideophyceae families (*Dasya binghamiae*, Dasyaceae; *Chondrus crispus*, Gigartinaceae; *Calliarthron tuberculosum*, Corallinaceae; *Grateloupia taiwanensis*, Halymeniaceae; *Gracilaria tenuistipitata*, Gracilariaceae). This alignment identifies 16 locally collinear blocks (LCBs) among the selected plastid genomes and demonstrates that even with the loss of photosynthesis genes overall synteny is shared between the parasite *H. mirabilis* and other plastid genomes, while *C. polysiphoniae* has undergone several genome rearrangements.

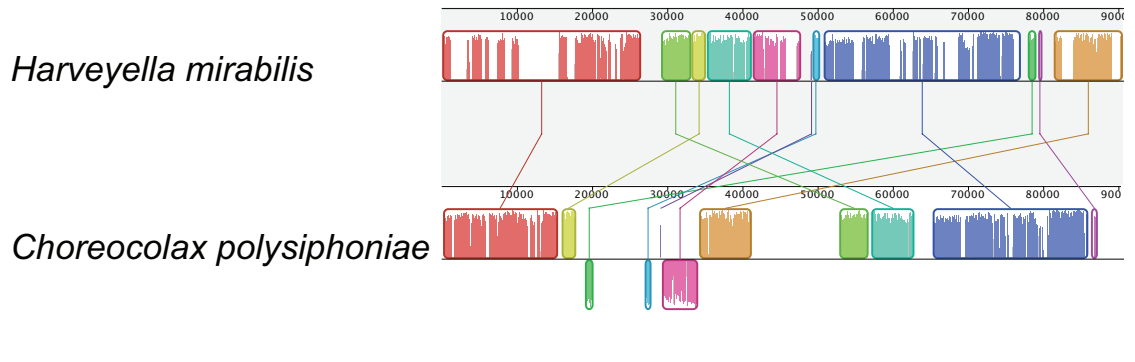


Figure 4. Mauve alignment of the plastid genomes from the parasites *Harveyella mirabilis* (top) and *Choreocolax polysiphoniae* (bottom). This alignment identifies 11 locally collinear blocks (LCBs) among these parasite plastid genomes, highlighting the high level of genome fragmentation and rearrangements evident in the two parasite plastids.

CHAPTER 5

Molecular Data Provides Support to Update Terminology for Distinguishing Parasitic Red Algae

by

Eric D. Salomaki¹, and Christopher E. Lane¹

is formatted for submission to the Journal of Phycology

¹Department of Biological Sciences, University of Rhode Island, Kingston, RI 02881.

Abstract:

The first parasitic red alga was described in 1845 with twenty-one additional species being described before the turn of the century. Today, more than 120 red algal parasites have been described from eight different orders within the Florideophyceae. Red algal parasites are predominately unpigmented and appear as erumpent pustules or irregular growths on their free-living red algal hosts. Despite their diminutive nature, parasitic red algae share morphological characteristics with other close relatives, allowing them to be assigned to tribes or families from their initial discovery. Historically parasites that have infected hosts in the same family have been considered ‘adelphoparasites’, while those that infect hosts in different families or tribes are called ‘alloparasites’. Recently, molecular data has demonstrated that species that are considered alloparasites and infect hosts outside of their family are exceedingly rare. Therefore, using a taxonomic, rather than biologic definition to classify parasites is no longer suitable. All ‘adelphoparasites’ that have been investigated have lost their native plastid and instead hijack a copy of the host plastid when packaging their spores. Functionally reduced native plastids were recently sequenced from the ‘alloparasites’ *Choreocolax polysiphoniae* and *Harveyella mirabilis*. Characteristics including previously described developmental differences and plastid origin can distinguish two distinct types of parasites. Here we describe the close link between plastid retention and developmental differences in red algal parasites. Based upon these observations, we propose the terms archaeoplastic parasites and neoplastic parasites to distinguish parasitic red algae based on a biological characteristic rather than relying on terms derived from a taxonomic framework.

A History of Red Algal Parasites

The first red algal parasite described was *Microcolax botryocarpa* (Hooker & Harvey) F. Schmitz in 1845 (Schmitz and Falkenberg 1897, Preuss et al. 2017) and over the next 55 years, another twenty-one red algal parasites were described, largely as a result of work by Reinsch & Schmitz (Reinsch 1875, Schmitz and Falkenberg 1897, Preuss et al. 2017). Compared to their free-living relatives, parasitic red algae are greatly reduced in size and generally lacking pigmentation. Despite their diminutive nature, parasitic red algae were taxonomically assigned to families based upon morphological features as early as the late 19th and early 20th century (Reinsch 1875, Schmitz and Falkenberg 1897). Due to their relative lack of morphological characteristics, the use of molecular data has provided a new lens through which to examine phylogenetic affiliations of parasitic red algae (Chapter 4). The most recent counts identify 121 described species of red algal parasites (Blouin and Lane 2015, Preuss et al. 2017), however molecular data has demonstrated crypsis is common (see independent evolutions of *Asterocolax gardneri* in Goff et al. 1997, and Chapter 4) and most parasites have not been examined with molecular methods.

Molecular and morphological analyses indicate that multiple independent evolutionary events have given rise to parasites throughout Florideophyceae (Goff et al. 1996, 1997, Zuccarello et al. 2004). In the early 20th century, two hypotheses arose to explain the origins of red algal parasites. First, Setchell (1918) noted that approximately 80% of the parasites he was studying infected hosts in the same taxonomic family. Based upon his observations, he proposed that parasites originated as carpospores or tetraspores of their host, which had undergone a mutation causing it

to have reduced photosynthetic capabilities (Setchell 1918). Later, Sturch (1926) proposed that rather than evolving sympatrically, parasitic red algae started out as small epiphytes of their hosts that eventually would penetrate cortical cells of the host, becoming an endophyte. Once established as an endophyte, the alga would adopt mechanisms to obtain nutrients from the host, becoming increasingly reliant on its host, solidifying an irreversible path towards parasitism (Sturch 1926). Feldmann and Feldmann (1958) furthered Sturch's hypothesis, adding that epiphytes that are closely related to their hosts are more likely to be successful creating connections with hosts cells and are therefore more likely to complete a transition to parasitism. Molecular analysis has supported Setchell's initial hypothesis (Goff et al. 1996), however the possibility that not all parasites arise through the same mechanisms has also been suggested (Salomaki and Lane 2014) and perhaps the Setchell and Sturch hypotheses both hold true.

As previously noted (e.g. Setchell 1918), the Feldmann's (1958) recognized a distinction between parasites that were closely related to their hosts and those that infected hosts outside of their family. Based upon their observations, the term adelphoparasite (adelpho- is Greek for 'kin') was applied to those parasites that infect hosts within their same family/tribe and the term alloparasite (allo- is Greek for other) was adopted to describe parasites that infect hosts from other families (Feldmann and Feldmann 1958). These terms were widely accepted as a means of discussing parasites and, to date, the ratio of adelphoparasites to alloparasites (5:1) has remained quite similar to Setchell's initial observations (1918). Goff and colleagues later proposed

that all red algal parasites begin as adelphoparasites, and over time diversify and/or adapt to infect new hosts as they develop into alloparasites (Goff et al. 1996, 1997).

Redefining Red Algal Parasites

Since their adoption by Feldmann and Feldmann (1958), the terms adelpho- and alloparasite have been used to distinguish red algal parasites. Phylogenetic studies have cast doubt on the utility of these terms as data show that even alloparasites predominately infect with the same family (Goff et al. 1996, Zuccarello et al. 2004, Kurihara et al. 2010, Chapter 4). However, differences between these two groups of parasites remain, including their developmental patterns as they infect their hosts, and the origins of organelles. Plastids remain essential for the survival of red algal parasites however the origin of their plastid differs and appears to be intimately linked with other developmental characteristics discussed here. Therefore, we recommend using the term archaeoplastic parasite to refer to those parasites that retain a native plastid (formerly alloparasites), and neoplastic parasite (formerly adelphoparasites) to discuss those that hijack a host plastid rather than retain their own copy.

It remains plausible that an archaeoplastic parasite will be found infecting a close relative or even a sister species, especially if they originate via sympatric speciation, as Setchell (1918) hypothesized. Rather than distinguishing types of parasites by their evolutionary relationships to their hosts, it seems more meaningful to use biological distinctions between groups that can be easily examined using widely available tools. Many developmental characteristics separate the two groups of parasites including the ability to vegetatively grow between host cells (Goff and

Coleman 1985, Goff and Zuccarello 1994, Salomaki and Lane 2014), the location of parasite DNA replication (Goff and Coleman 1984, 1987, Salomaki and Lane 2014), and in a phylogenetic construct, the ability to successfully speciate (Zuccarello et al. 2004, Chapter 4). However, there are flaws in each of these options including the difficulty to assess parasite vegetative growth and location of DNA replication as well as the need for appropriate comparative data for phylogenetic analyses.

Developmental differences

As technology progressed, new tools became available to investigate the cellular interactions between red algal parasites and their hosts. The association between parasitic red algae and their hosts is facilitated, at least in part, by the ability of red algae to form cell-to-cell fusions between adjacent, non-daughter cells (Sturch 1899, 1926, Feldmann and Feldmann 1958). With modern microscopy, the structure and formation of cell fusions between parasites and their hosts was determined (Kugrens and West 1973, Goff and Cole 1976a, Wetherbee and Quirk 1982a, 1982b, Wetherbee et al. 1984). Furthermore, these cell-to-cell fusions were proposed to serve as the mechanism by which nutrients are transported from the host to parasite cells (Wetherbee and Quirk 1982b). With the use of epifluorescence microscopy, Goff and colleagues went on to establish their importance in the infection process by demonstrating that red algal parasites are able to transfer their nuclei and organelles into host cells via cell fusions (Goff and Coleman 1984, 1985, 1987, Goff and Zuccarello 1994).

In addition to recognizing the role of cell fusions in the infection process, sophisticated microscopy advanced our understanding of how parasitic red algae spread throughout their hosts and characterized the host responses. In a groundbreaking series of manuscripts, Goff and Cole described in great detail the biology of the archaeplastid parasite *Harveyella mirabilis* (Reinsch) F. Schmitz et Reinke, including its development, structure, and nutrient acquisition from its host *Odonthalia flocossa* (Esper) Falkenberg (Goff and Cole 1973, 1975, 1976a, 1976b, Goff 1976, 1979a, 1979b). These investigations provided a framework to more intimately understand the development and interactions of a range of red algal parasites.

Aside from the formation of cell fusions that initiate host infection, important differences have been observed between neoplastic parasites and archaeplastid parasites regarding their developmental patterns subsequent spread throughout their hosts (reviewed in Salomaki and Lane 2014, Freese and Lane 2017). First, archaeplastid parasites including *Choreocolax*, *Harveyella*, and *Holmsella* spread throughout their hosts by producing mitotically dividing rhizoidal filaments that grow between host cells (Sturch 1899, 1926, Goff and Cole 1976a, Fredericq and Hommersand 1990). In contrast, the neoplastic parasites *Gracilariophila oryzoides*, *Gardneriella tuberifera*, and *Janczewskia gardneri* infect their hosts directly and spread from host cell to host cell rarely creating their own rhizoidal filaments (Goff and Coleman 1987a, Goff and Zuccarello 1994). Additionally, while neoplastic parasites appear to transform infected host cells and undergo nuclear replication solely within host cells (Goff and Coleman 1987, Goff and Zuccarello 1994), archaeplastid

parasites appear to only undergo DNA replication in their rhizoidal filaments and infected host cells maintain a 1:1 ratio of parasite nuclei to cell fusions (Goff and Coleman 1984, 1985, 1987). These key developmental differences have furthered support for distinguishing types of parasites. However, the terms adelpho- and alloparasite, which specifically refer to taxonomic relatedness to their hosts, does not reflect the differing biology of the organisms.

Organellar Origins

The combination of microscopy with molecular tools has substantially advanced our understanding of red algal parasites and their interactions with their hosts. In a study investigating the origins of organelles from three neoplastic parasites, *Gracilariophila oryzoides*, *Gardneriella tuberifera*, and *Plocamiocolax pulvinata*, Goff and Coleman (1995) demonstrated that the parasites retain their own mitochondria, however all three had lost their native plastid and instead incorporate a dedifferentiated host plastid when packaging their spores. Molecular investigations including deep genomic sequencing, have confirmed that the neoplastic parasites *Faucheocolax attenuata*, *Gracilariophila oryzoides*, *Gardneriella tuberifera*, *Janczewskia gardneri*, and *Plocamiocolax pulvinata* do not retain a native plastid (Goff and Coleman 1995, Salomaki and Lane, unpublished).

Recently, plastid genomes have been completely sequenced from the archaeoplastic parasites *Choreocolax polysiphoniae* and *Harveyella mirabilis*. Both plastids are greatly reduced in coding capacity compared to the plastids of free-living photosynthetic red algae, having lost all genes related to photosynthesis (Salomaki et

al. 2015, Chapter 4). Interestingly, genes involved in amino acid and fatty acid biosynthesis, iron-sulfur cluster synthesis, as well as transcription and translation are conserved (Salomaki et al. 2015, Chapter 4). Aside from the loss of photosynthesis genes, the *H. mirabilis* plastid retains a high level of synteny with plastids of closely related free-living red algae, however the *C. polysiphoniae* plastid has experienced extensive genome reorganization (Figure 1). Furthermore, the discovery that some parasites retain functionally reduced plastids, while others rely on a host-derived plastid, provides another physical characteristic for differentiating between two types of parasites rather than relying on evolutionary relationships.

Phylogenetic investigations have definitively demonstrated that archaeoplasic parasites rarely infect hosts in a different family or tribe (Zuccarello et al. 2004, Kurihara et al. 2010). A recent analysis recognizes that the archaeoplasic parasites *C. polysiphoniae*, *H. mirabilis*, and two species of *Leachiella*, form a monophyletic clade in which all species retain their native plastid (Chapter 4). Plastid origin remains a remarkable difference between types of parasites and is easy to investigate with molecular tools these days. Furthermore, plastid origin is a meaningful biological characteristic that can distinguish between types of red algal parasites regardless of relationship to their host. In light of these results, the terms alloparasite and adelphoparasite, as originally described by Feldmann and Feldmann (1958), create more confusion than they allay (Zuccarello et al. 2004, Kurihara et al. 2010, Chapter 4).

Aside from the previously discussed developmental differences including the location of parasite DNA replication and mechanism of spreading throughout the host, plastid origin appears to be central to differences observed in red algal parasite evolution. One hypothesis is that the observed developmental differences between the two groups of parasites are linked to disparities in their plastid origin. Parasites that maintain a native plastid are in essence, a complete red alga in their own right, whereas those that incorporate a host derived plastid are borrowing parts of another organism they remain compatible with. The retention of native organelles means the plastid remains under the same selective pressures as the rest of the parasite genome, and therefore, retains its ability to function for amino acid, fatty acid, protein, and isoprene biosynthesis while losing genes for photosynthesis. Those parasites that retain their own plastid are capable of evolutionary processes typical of other formerly photosynthetic parasites, such as Apicomplexans. A native plastid would therefore enable those parasites maintain control over organellar evolution and presumably be more successful adapting to new hosts and speciating to form monophyletic clades. This hypothesis can be tested in future studies by investigating the organellar origins of parasites in the genus *Holmsella*.

Holmsella pachyderma was previously placed in the Choreocolacaceae based upon its developmental patterns within its host, until 1990 when it was transferred to the Gracilariaceae along with *Holmsella australis*, based upon morphological characteristics (Fredericq and Hommersand 1990). A subsequent phylogenetic analysis including *H. australis* and *H. pachyderma* found the species to form a monophyletic clade within the Gracilariaceae (Zuccarello et al. 2004). To date no

work has been completed on the plastid origin of members of the genus *Holmsella*, however if they do retain a native plastid, *Holmsella* would represent a second successful evolution of parasites retaining a native plastid that forms mitotically dividing rhizoidal filaments as a means of vegetative growth throughout their hosts, and forms a monophyletic clade of parasites. Further investigation is warranted to determine the taxonomic affiliation of parasite genera *Gelidiocolax* and *Pterocladiphila*, which have also been placed within the Gracilariaceae based on morphological characteristics.

The origin of parasites that have lost their native plastid and instead hijack a plastid from their host, though well studied, remains uncertain. The ability of florideophycean algae to easily form cellular fusions with adjacent cells would enable the spore with a non-functional plastid to germinate and utilize genetically similar plastids. By exploiting a host plastid, the newly evolved parasite is rescued from what otherwise would have been a lethal mutation. The parasite nucleus and mitochondrion could continue spreading from cell-to-cell of its newly acquired host via cell fusions, transforming the host cells as described in detail by Goff and colleagues (Goff and Coleman 1985, Goff and Zuccarello 1994). While providing a means of short-term survival for the newly transitioned parasite, co-opting a host plastid as their own would also establish these parasites as evolutionary dead ends. All red algae rely on a plastid for essential functions beyond photosynthesis including amino acid and fatty acid biosynthesis. By retaining a new plastid from its host in each new infection cycle, the parasite loses control of organelle evolution. While the parasite nucleus and mitochondrion will remain under a unique selective pressures, the host and host plastid

will be experiencing different selective pressures. Over time this will render the parasite incapable of communicating with the host plastids, rendering it unable to grow and reproduce in a diverging host.

In addition to the organelles themselves, nuclear-encoded, organellar-targeted, genes are also the subjects of ongoing investigations in the evolution of parasitism in red algae. Comparative analyses of the transcriptomes from the neoplastic parasite, *Gracilariophila oryzoides* and the archaeoplastic parasite *Choreocolax polysiphoniae*, revealed distinct differences in the expression of nuclear-encoded plastid-targeted photosynthesis genes involved in the carotenoid biosynthesis pathway. This research indicates that all genes with the exception of geranylgeranyl phosphate synthase (GGPS) are not being transcribed in the *C. polysiphoniae* transcriptome, while genes in the carotenoid biosynthesis pathway are transcribed by *G. oryzoides* (Figure 2) (Salomaki & Lane, unpublished). It is reasonable that *C. polysiphoniae* would no longer maintain selective pressure on photosynthesis related genes since it retains a plastid that is no longer capable of photosynthesizing. However, the finding that *G. oryzoides* not only retains functional copies, but also is expressing plastid-targeted genes involved in the carotenoid biosynthesis pathway suggests that selection is acting to conserve these genes. Their expression indicates the proteins are being targeted to the host-derived plastid as a means of controlling and utilizing the host plastid as its own. Analyses remain ongoing to determine the nature and extent of nuclear-encoded, plastid-targeted, genes in both parasites, and this research direction appear to hold promise in further explaining the interactions between red algal parasites and their hosts.

Literature Cited

- Blouin, N. & Lane, C. 2015. Red algae provide fertile ground for exploring parasite evolution. *Perspect. Phycol.* 3:11–9.
- Feldmann, G. & Feldmann, J. 1958. Recherches sur quelques Floridees parasites. *Rev. Gen. Bot.* 65:49–128.
- Fredericq, S. & Hommersand, M. 1990. Morphology and systematics of *Holmsella pachyderma* (Pterocladophilaceae, Gracilariales). *Br. Phycol. J.* 25:39–51.
- Freese, J.M. & Lane, C.E. 2017. Parasitism finds many solutions to the same problems in red algae (Florideophyceae, Rhodophyta). *Mol. Biochem. Parasitol.* 214:105–11.
- Goff, L.J. 1976. The Biology of *Harveyella mirabilis* (Cryptonemiales, Rhodophyceae). V. Host Responses to Parasite Infection. *J. Phycol.* 12:313–28.
- Goff, L.J. 1979a. The Biology of *Harveyella mirabilis* (Cryptonemiales, Rhodophyceae). VI. Translocation of Photoassimilated ¹⁴C. *J. Phycol.* 15:82–7.
- Goff, L.J. 1979b. The Biology of *Harveyella mirabilis* (Cryptonemiales, Rhodophyceae). VII. Structure and Proposed Function of Host Penetrating Cells. *J. Phycol.* 15:87–100.
- Goff, L.J., Ashen, J. & Moon, D. 1997. The evolution of parasites from their hosts: A case study in the parasitic red alga. *Evolution.* 51:1068–78.
- Goff, L.J. & Cole, K. 1975. The biology of *Harveyella mirabilis* (Cryptonemiales, Rhodophyceae). II. Carposporophyte development as related to the taxonomic affiliation of the parasitic red alga, *Harveyella mirabilis*. *Phycologia.* 14:227–38.

- Goff, L.J. & Cole, K.M. 1973. The biology of *Harveyella mirabilis* (Cryptonemiales, Rhodophyceae). I. Cytological investigations of *Harveyella mirabilis* and its host, *Odonthalia floccosa*. *Phycologia*. 12:237–45.
- Goff, L.J. & Cole, K.M. 1976a. The biology of *Harveyella mirabilis* (Cryptonemiales, Rhodophyceae). III. Spore germination and subsequent development within the host *Odonthalia floccosa* (Cermiales, Rhodophyceae). *Can. J. Bot. Can. Bot.* 54:268–80.
- Goff, L.J. & Cole, K.M. 1976b. The biology of *Harveyella mirabilis* (Cryptonemiales, Rhodophyceae). IV. Life history and phenology. *Can. J. Bot. Can. Bot.* 54:281–92.
- Goff, L.J. & Coleman, A.W. 1984. Transfer of nuclei from a parasite to host. *Proc. Natl. Acad. Sci.* 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. 1987. Nuclear transfer from parasite to host: A new regulatory mechanism of parasitism. *Ann. N. Y. Acad. Sci.* 503:402–23.
- 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.
- Kugrens, P. & West, J.A. 1973. The ultrastructure of an alloparasitic red alga *Choreocolax polysiphoniae*. *Phycologia*. 12:175–86.

- 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 *Ululania* (Ceramiales). *J. Phycol.* 46:580–90.
- 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.* 60:1–13.
- Reinsch, P.F. 1875. *Contributiones ad Algologiam et Fungologiam*. Lipsiae, Nuremberg. 61 pp.
- Salomaki, E.D. & Lane, C.E. 2014. Are all red algal parasites cut from the same cloth? *Acta Soc. Bot. Pol.* 83:369–75.
- Salomaki, E.D. & Lane, C.E. 2016. Red Algal Mitochondrial Genomes are More Complete than Previously Reported. *Genome Biol. Evol.* 48:evw267.
- Salomaki, E.D., Nickles, K.R. & Lane, C.E. 2015. The ghost plastid of *Choreocolax polysiphoniae*. *J. Phycol.* 51:217–21.
- Schmitz, F. & Falkenberg, P. 1897. *Rhodomelaceae*.
- Setchell, W.A. 1918. Parasitism Among the Red Algae. *Proc. Am. Philosophical Soc.* 57:155–72.
- Sturch, H.H. 1899. *Harveyella mirabilis* (Schmitz & Reinke). *Ann. Bot.* 8:83–102.
- Sturch, H.H. 1926. *Choreocolax polysiphoniae*, Reinsch. *Ann. Bot.* XL.
- Wetherbee, R. & Quirk, H. 1982a. The Fine Structure of Secondary Pit Connection Formation Between the Red Algal Alloparasite *Holmsella australis* and Its Red Algal Host *Gracilaria furcellata*. *Protoplasma.* 176:166–76.

Wetherbee, R. & Quirk, H. 1982b. The Fine Structure and Cytology of the Association Between the Parasitic Red Alga *Holmsella australis* and Its Red Algal Host *Gracilaria furcellata*. *Protoplasma*. 165:153–65.

Wetherbee, R., Quirk, H.M., Mallett, J.E. & Ricker, R.W. 1984. The Structure and Formation of Host-Parasite Pit Connections Between the Red Algal Alloparasite *Harveyella mirabilis* and Its Red Algal Host *Odonthalia floccosa*. *Protoplasma*. 73:62–73.

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.

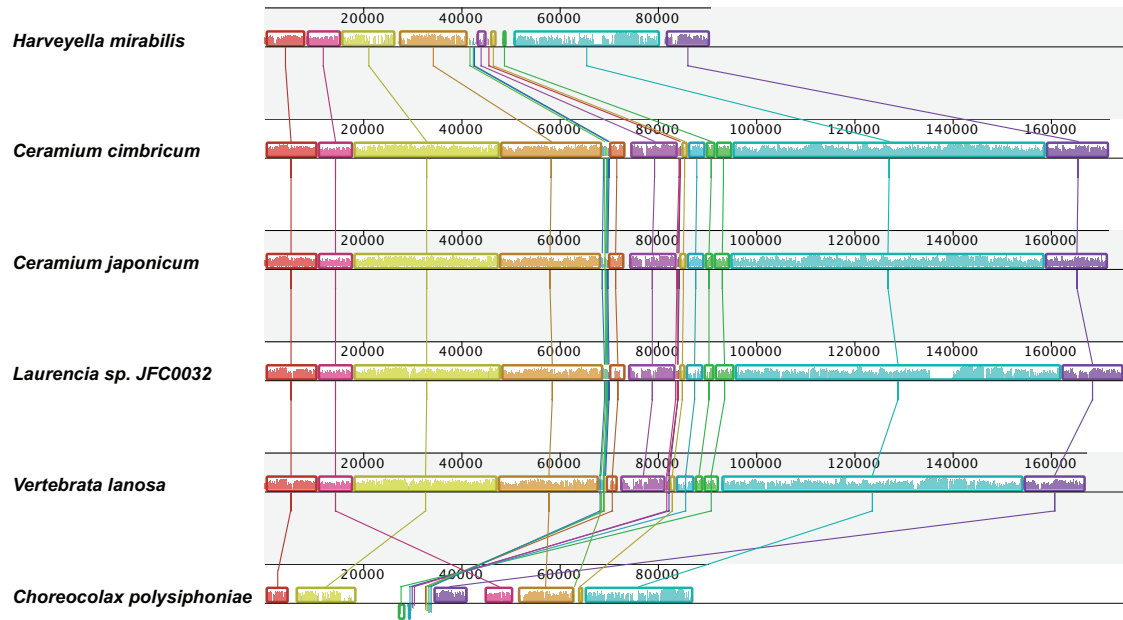


Figure 1. Mauve alignment of the parasites *Harveyella mirabilis* (top) and *Choreocolax polysiphoniae* (bottom) with all published Rhodomelaceae plastid genomes. This alignment identifies 12 locally collinear blocks (LCBs) among the selected plastid genomes and demonstrates that even with the loss of photosynthesis genes overall synteny is shared between the parasite *H. mirabilis* and other plastid genomes, while *C. polysiphoniae* has undergone several genome rearrangements.

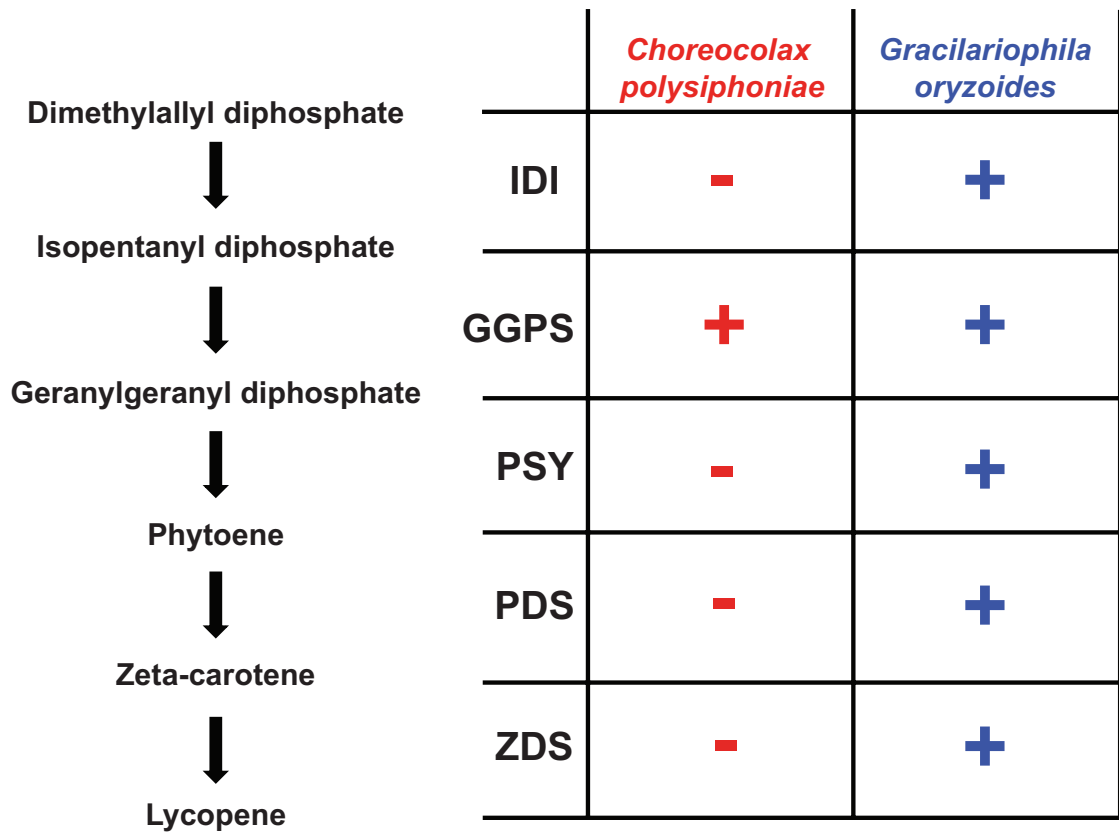


Figure 2. Comparison of carotenoid biosynthesis related genes between the red algal parasite *Choreocolax polysiphoniae* and *Gracilariophila oryzoides*.