Photosynthetic Agents of Carbonate Bioerosion in Endolithic Microbiomes

by

Daniel Roush

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved July 2020 by the
Graduate Supervisory Committee:

Ferran Garcia-Pichel, Chair
Ariel Anbar
Hinsby Cadillo-Quiroz
Huansheng Cao

ARIZONA STATE UNIVERSITY

August 2020

ABSTRACT

Cyanobacteria and algae living inside carbonate rocks (endoliths) have long been considered major contributors to bioerosion. Some bore into carbonates actively (euendoliths); others simply inhabit pre-existing pore spaces (cryptoendoliths). While naturalistic descriptions based on morphological identification have traditionally driven the field, modern microbial ecology has shown that this approach is insufficient to assess microbial diversity or make functional inferences. I examined endolithic microbiomes using 16S rRNA genes and lipid-soluble photosynthetic pigments as biomarkers, with the goal of reassessing endolith diversity by contrasting traditional and molecular approaches. This led to the unexpected finding that in all 41 littoral carbonate microbiomes investigated around Isla de Mona (Puerto Rico) and Menorca (Spain) populations of anoxygenic phototrophic bacteria (APBs) in the phyla Chloroflexi and Proteobacteria, were abundant, even sometimes dominant over cyanobacteria. This was not only novel, but it suggested that APBs may have been previously misidentified as morphologically similar cyanobacteria, and opened questions about their potential role as euendoliths. To test the euendolithic role of photosynthetic microbes, I set a time-course experiment exposing virgin non-porous carbonate substrate in situ, under the hypothesis that only euendoliths would be able to initially colonize it. This revealed that endolithic microbiomes, similar in biomass to those of mature natural communities, developed within nine months of exposure. And yet, APB populations were still marginal after this period, suggesting that they are secondary colonizers and not euendolithic. However, elucidating colonization dynamics to a sufficiently accurate level of molecular identification among cyanobacteria required the development of a curated cyanobacterial 16S rRNA gene reference database and web tool, Cydrasil. I could then detect that the pioneer euendoliths were in a novel cyanobacterial clade (named UBC),

immediately followed by cyanobacteria assignable to known euendoliths. However, as bioerosion proceeded, a diverse set of likely cryptoendolithic cyanobacteria colonized the resulting pore spaces, displacing euendoliths. Endolithic colonization dynamics are thus swift but complex, and involve functionally diverse agents, only some of which are euendoliths. My work contributes a phylogenetically sound, functionally more defined understanding of the carbonate endolithic microbiome, and more specifically, Cydrasil provides a user-friendly framework to routinely move beyond morphology-based cyanobacterial systematics.

DEDICATION

To Danny, Debbie, Beth, Jennifer, Robyn

ACKNOWLEDGMENTS

I would first like to express my deep gratitude for every single person that has helped me throughout my time here at ASU. Everyone who has been my second pair of eyes, lent me an ear or shoulder, and put up with my characteristic morning grumpiness, I thank you.

Ferran. Man, we made it. From getting stitched up on the beach of a deserted island to defending my dissertation, it has been quite the ride. Thank you for never pulling a punch and pushing me to do my very best. I have grown greatly as both a person and a scientist under your tutelage, and I am eternally grateful for the opportunity.

To my committee, thank you for your patience, understanding and interest throughout my time here at ASU. Being able to pivot my final chapter into focused coding and bioinformatics project was some of the most fun I have ever had, and I thank you for letting me follow my passions to complete my dissertation.

Dad, thanks for always picking up the phone when I gave you a call. You listened to every single one of my rants and gave me good advice at every turn. I am not sure if I would have kept it together this long without our biweekly chats. Beyond that, from an early age you kept me on a track to achieve something like this. I will never forget our Saturday morning fishing trips, or those chess matches.

FGP lab members past and present, thank you for always being present in my academic and person life, through good days and bad. I learned a lot from all of you, be it techniques, languages, or traditions. Normally you would say cheers or salud here, but I think I would rather ask if anyone wants a Manzana.

I would like to specially thank Aralcy Garrástazu for both repairing my injured legs on Isla de Mona, helping with tile collection, and also being our liaison and guide for both Puerto Rico trips.

To all my friends, all over the world. Ya'll rock. Thanks for listening, playing games, telling stories, and doing all the things you do to keep me sane over the years. I could not have done it without each and every one of you.

Thank you all.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1. Background

1.1. The Endolithic Habitat

The endolithic habitat consists of the cracks, crevices, and pore spaces found within the interior of rocks. Considered an extreme environment, the endolithic habitat is colonized by a wide range of microbial life forms that can scavenge necessary nutrients or utilize available sources of energy. For example, sulfur- and iron-oxidizing chemoautotrophic communities inhabit cracks in basaltic rocks on the seafloor, where oxygenated deep-sea water meets minerals rich in reduced iron and sulfur compounds [1,2]; microbes there obtain energy from the respective redox reactions. Phototrophic endoliths can be found in sedimentary rock outcrops [3–7] or on the interior of naturally translucent rocks [8,9] but are confined to shallow depths where light can penetrate. Endoliths are classified into three types: those that actively dissolve the substrate and bore their way into it (euendoliths), those that inhabit extant pore spaces or pore spaces excavated by euendoliths (cryptoendoliths), and those that colonize natural cracks and crevices (chasmoendoliths). Organisms found growing on the outside of the rock are deemed epiliths [10] (Figure 1). Phototrophic endolithic microbiomes are ecologically important in extreme ecosystems where plant growth is restricted, where they can become the main source of primary production, such as in deserts rock outcrops [8,9,11–13], polar biomes [8,9,14,15], and marine coastal environments [6,7,16–23]. In carbonates for example, endolithic communities can reach high productivity rates, upwards of 100 g C km^{-2} d^{-1} [24] or, by extrapolation, several billion tons of carbon per year globally [25]. This high primary productivity even can support food webs of specialized metazoan grazers like hard-toothed invertebrates and fish [26].

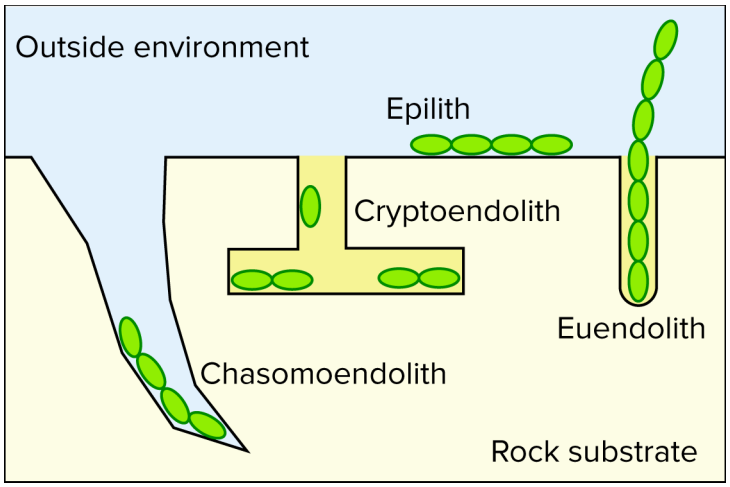

Figure 1. Diagram of different endolithic lifestyles adapted from Golubic et al. 1981 [10].

Of the endolithic habitats, those on carbonate rocks (Figure 2) have been the focus of most investigations due to their ecological relevancy. Carbonate formations are common and found across a wide range of ecosystems, perhaps the most prominent being coral reefs. Reef systems are comprised of the naturally porous aragonite skeletons of coral polyps and are home to a many endolithic organisms [21,27,28]. However, reefs only comprise 1.4% of the exposed carbonate globally. In fact, carbonate outcrops cover 2 x 10^7 square kilometers [29], or some 15%, of the ice-free land surface. In addition to their geographical relevance, carbonates (and calchophosphatic minerals), if sufficiently sunlit, are susceptible to the bioerosive activity of photosynthetic euendolithic microbes [19] (Figure 3). As bioerosion proceeds, new pore spaces are created that allow for general endolith growth, but also can lead to the destruction of corals, fisheries, and karstic coasts [27,30–34]. Newly exposed carbonates are quickly colonized by euendoliths, with complete colonization occurring in under a year [35]. The recently elucidated mechanism of boring found in the euendolithic cyanobacterium *Mastigocoleus testarum* BC008 [36] suggests that colonization and bioerosion rates will only increase in the future as anthropogenic increase in atmospheric CO_2 will enhance

carbonate dissolution [23]. Traces of euendolithic lifestyles as either fossilized organisms or their microborings (trace fossils) are abundantly found in the fossil record, dating as far back as the Proterozoic [37–39]. Euendolithic fossils patterns are commonly used in paleobathymetric reconstruction [40] and as indicators of coral growth [41] in paleobiology.

Figure 2. Typical euendolith carbonate habitats. (**a**) Dolomite field at low tide in the intertidal zone of Churchill, Manitoba, Canada. (**b**) Terrestrial travertine boulders at Tonto Natural Bridge, Arizona, USA. (**c**) Broken chip in an exposed fossil reef carbonate platform from Isla de Mona, Puerto Rico, showing a thin line of green endolithic growth beneath the surface. (**d**) Photograph of Isla de Mona carbonate platform showing limestone-dolomite stratification and characteristic intertidal notch.

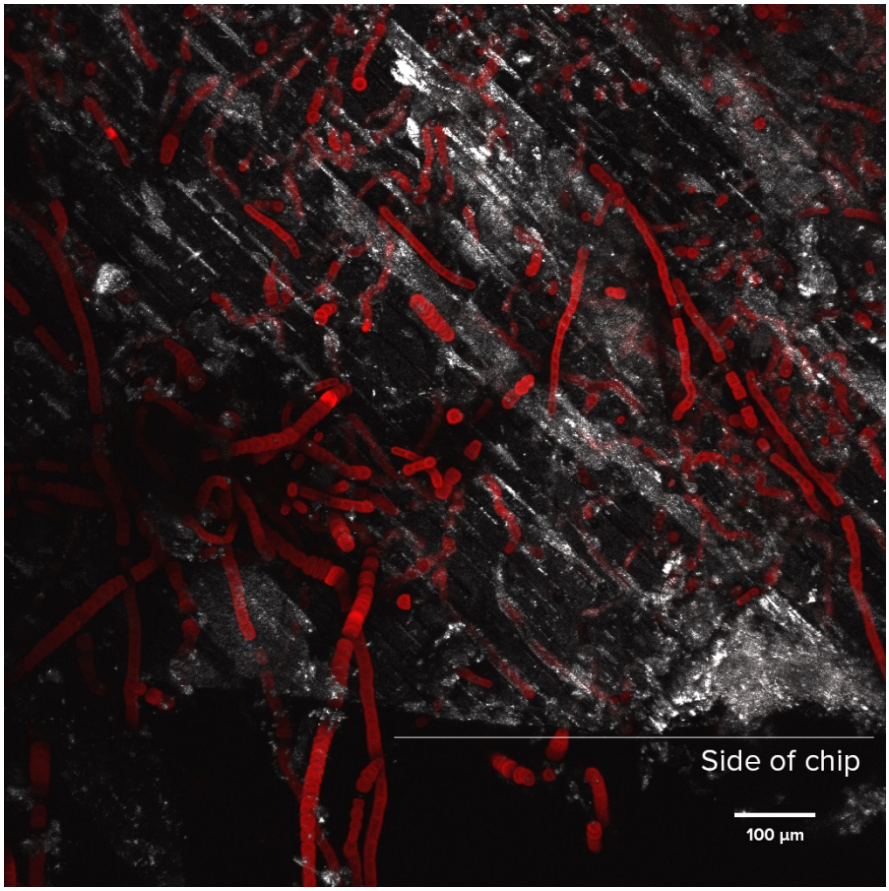

Figure 3. Top down laser scanning confocal microscopy image of a euendolithic cyanobacterium, *Brasilonema* sp. The organisms were allowed to bore into a piece of Icelandic spar (optically clear calcium carbonate) and then imaged by autofluorescence using laser excitation at 405 nm. Chl *a* autofluorescence was captured between 660 nm and 690 nm. The grey/white artifact is the reflection of the surface of the spar.

1.2. Community Structure of the Carbonate Endolithic Microbiome

For almost two centuries, scientists have provided descriptive studies on (eu)endolithic microbiomes that span the microbiological, ecological, and geological fields [42–50]. These studies, dependent upon microscopic inspection and morphological descriptions, have described photosynthetic carbonate euendoliths among eukaryotes (green and red algae) and prokaryotes (cyanobacteria). Photosynthetic euendoliths inhabit calcareous and calcophosphatic minerals [51], only

as deep into the rock as light is available for photosynthesis, which typically does not go beyond a few millimeters. The most common algal endoliths are in the genera *Ostreobium* and *Phaeophila* [52] (Chlorophyta), as well as the rhodophyte *Porphyra* [53]. As for cyanobacteria, the pseudofilamentous, baeocyte formers various species of *Hyella* [54] and *Solentia* [55], the simple filamentous *Plectonema terebrans* [52], and the complex heterocystous *Mastigocoleus testarum* [6] are all common. Importantly, because these genera were assigned using only by morphological criteria, it is extremely likely that they obscure the true genetic and phylogenetic diversity of phototrophic endoliths. The first study of carbonate endolithic using simple molecular techniques for diversity assessment[6], certainly suggests as much. The recent rapid advancement of sequencing technology and subsequent application to microbial ecology has had much to offer recently in this regard: with its application to endolithic microbiomes it has become apparent that biodiversity has been underestimated, although there is also some level of coherence between DNA-based surveys and the traditional literature with regards to cyanobacteria [6,7]. Concerning the missed biodiversity, high-throughput sequencing studies have shown that endolithic microbiomes are more metabolically complex than predicted. For example, Horath and Bachofen [5] found the predicted green algae and cyanobacteria in dolomite outcrops in the Alps, but an additional large contribution to microbiome diversity from chemotrophic bacteria and archaea. Similarly, findings from a large range of ecosystems and host rocks including Antarctic sandstones [14], terrestrial limestones, sandstones and granites [56,57], as well as halites in the Atacama desert [58] all confirmed that the diversity of endolithic microbial populations extends well beyond cyanobacteria and algae. This metabolic complexity was also made apparent in the first high-throughput, multi-sample survey of endolithic microbiomes from intertidal outcrops [59]. We found that morphology-based studies of intertidal

carbonates can underrepresent cyanobacterial diversity estimates by factors of 10 to 100 [6,20,59], and that the level of diversity found was comparable to that of other microbiomes such as those of biological soil crusts [60] or sedimentary microbialites [61]. We found that littoral carbonates contained representatives of a variety of microbial metabolisms, from fermenters to sulfur oxidizers. As for phototrophic organisms, these habitats contained plenty of algae and cyanobacteria as expected, but surprisingly many phylotypes potentially assignable to anoxygenic phototrophic bacteria (APBs) were also present, and significantly so. There was nary a precedent for this. APBs had only been mentioned twice as components of the endolithic microbiome: once indirectly through detection of APB photopigments [62], and once as opportunistic colonizers of dead coral skeletons [63].

1.3. Anoxygenic Phototrophic Bacteria

Anoxygenic phototrophic bacteria are a polyphyletic metabolic group (a functional guild) that has representatives in six different bacterial phyla. Historically, APBs have been classified into five groups; the green sulfur bacteria in the phylum *Chlorobi*, the green non-sulfur bacteria of the phylum *Chloroflexi*, the purple sulfur bacteria in the Gammaproteobacteria, the purple non-sulfur bacteria of the Alpha- and Betaproteobacteria, and the *Heliobacteria* within the Firmicutes [64]. Recently new guild members have been discovered among the Alpha- and Betaproteobacteria [65] (the so called, aerobic APBs, and single species within the Acidobacteria [66,67] and the Gemmatimonadetes [68]. APBs occupy a rather narrow range of habitats: green and purple sulfur bacteria are limited to anaerobic locations in which there is a ready supply of an electron donor, typically hydrogen sulfide, like the anoxic bottom of meromictic lakes, ocean sediments, hot springs [69–71], microbial mats [72], stratified marine and

estuarine waters [73,74], and subglacial lakes [75]. Green non-sulfur bacteria have been found in hot springs, hypersaline microbial mats, and in some marine sediments [76]. Purple non-sulfur bacteria occur in a somewhat wider range of habitats, including the open ocean and Antarctic lakes, but they rarely are the dominant phototroph within any given community [77].

Anoxygenic phototrophy shares many of the same molecular underpinnings as oxygenic photosynthesis. Both rely on chlorins (bacteriochlorophylls or chlorophylls) as the main photopigments, organized around molecularly similar membrane-bound photosystems [64]. This may shed light as to why APBs were overlooked as endoliths, as the primary photopigments have overlapping absorbance spectra (Figure 4). In the case of anoxygenic photosynthesis, a wide range of primary reductants are available for autotrophy, such as H_2, NO_2^-, H_2S, Fe^{2+} [64]. Alternatively, some APBs conduct photoheterotrophy, consuming available organics as a carbon source while using light to generate ATP [64]. Yet still, other APBs are facultative photoautotrophs, acting mostly as photoheterotrophs when reductant is unavailable, as is the case of *Chloroflexus aurantiacus* from hot springs, known to obtains organic carbon and oxygen from thermophilic *Synechococcus* cyanobacteria thriving in close proximity [76].

When considering the potential role of APBs in carbonate endolithic ecosystems, one must pay particular attention to the relationship between metabolism and carbonate equilibrium. In the case of photosynthetic APBs, studies been shown that carbonate precipitation does occur in culture if the primary reductant is H_2S [78], or H_2 [79]. Bundeleva et al. [80] report that *Rhodovulum sp.,* using nitrite as the primary electron donor, acidifies the medium, inhibiting carbonate precipitation. Photoheterotrophic APB metabolism can also potentially affect carbonate equilibrium: *Rhodovulum* consuming acetate and lactate as sources of organics raises the medium pH, leading to carbonate

precipitation, but it does not apparently do so when grown on neutral sugars [80]. A similar result was found when using *Rubrivivax* isolates [81]. Another level of complexity is added when one takes into account APB metabolism at night, which, either through respiration, fermentation, or even chemolithothrophy [82], likely contributes to the acidification of their microenvironment potentially resulting in carbonate dissolution. APBs could therefore logically act as euendoliths in nature by virtue of their metabolism in the cases where this leads to acidification, potentially upending the long-established understanding of endolith ecology by broadening the pool of possible pioneer organisms and boring mechanisms.

Figure 4. Laser scanning confocal microscopy image of a mixed APB (*Chloroflexus aurantiacus*; thin filaments) and cyanobacterial (*Nostoc punctiforme*; large, defined multicellular filaments) culture. The culture was excited with a 405 nm laser and photosynthetic pigment autofluorescence (false green color) was captured between 660 nm and 800 nm, and shows broad overlap in the autofluorescence emission of APB (bacteriochlorophyll *a, c*) and cyanobacterial (chlorophyll *a*) photopigments.

1.4. Successional Dynamics in Carbonate Endolithic Microbiomes

Some of the first euendolithic colonization studies were undertaken by Le Campion-Alsumard and colleagues [83,84] who examined these communities by microscopic inspection, classifying endoliths by traditional morphological systems. Concerns for coral reef bioerosion in the 1990s spurred a rapid increase in such colonization studies [21,22,34,35,48,85,86]; they all applied the same morphological techniques as Le Campion-Alsumard, in attempts to understand and mitigate reef collapse. Due to the overwhelming support for coral research, much of our understanding of euendolithic colonization came from investigations of biogenic carbonates from coral skeletons, but because these substrates are naturally porous, it is not immediately obvious that the results should apply to carbonate outcrops in general. Those studies showed that euendolithic algae colonize new substrates in a matter of weeks, followed by successional dynamics over the ensuing months, with apparent maturity (relative stasis) reached after about a year. In contrast, only a few studies on hard mineral carbonate substrates exist. Kiene [34], Gektidis [35] and Chacón et al. [6] were able to show that non-porous substrates favored prokaryotic colonization resulting in euendolithic cyanobacteria, not algae, as dominant pioneers.

In spite of the fact that morphological characterization can severely underestimate microbial diversity [6,7], as discussed above, early contributions could still identify and characterize some patterns. Three morphotypical groups (morphotypes) of euendolithic cyanobacteria were distinguished. The first are the thin, filamentous, *Leptolyngbya*-like organisms typically identified as *Plectonema terebrans* (or *Leptolyngbya terebrans)* which are one of the most prolific euendolith morphotypes, exceeding 80% of total euendolithic biomass in some studies [21]. Unfortunately, there is simply no easy way to assign *P. terebrans*-like morphotypes with certainty to

cyanobacteria, as other thin, filamentous organisms exist that could be easily mistaken for it, such as Flexibacteria or some of the filamentous APBs, like the Chloroflexi. To make matters worse, the ability to detect cell color in optical microscopy vanishes as one approaches cell sizes around one micron, which is quite close to the cell width of all of these filaments, including *Plectonema terebrans* itself. The fact that no cultivated isolates of euendolithic *P. terebrans* exist, makes it close to impossible to assign unambiguously a 16S rRNA gene sequence or group of sequences to this morphotype. Some environmental sequences from endolithic microbiomes that have been assigned cyanobacterial genera with thin filament morphology (such as *Halomicronema* and *Leptolyngbya)* were tentatively suggested to represent *P. terebrans* [59], but this should be regarded as a working hypothesis, and the true nature of this euendolith remains elusive. The second morphotypical group is based on the species *Mastigocoleus testarum*. *M. testarum* has a complex, true-branching filamentous morphology, making it easily identifiable from microscopic examination. It has been recently re-described on the basis of a polyphasic approach based on strain BC008, showing congruency between molecular and traditional approaches [19], has served as a model to elucidate the physiological mechanism of boring [19,36,87], and is the first euendolith whose genome has been fully sequenced [88]. *Mastigocoleus testarum* is one of the earliest colonizers in soft carbonates, detected as early as one week after initial exposure [22,24,83]. A third, diverse group includes several members in the order Pleurocapsales (baeocyte formers) in the genera *Hyella, Solentia*, *Hormathonema*, and the recently described *Candidatus* Pleuronema. Members of the Pleurocapsales typically act as pioneer borers but can bore only to shallow depths and are easily preyed upon by grazers, leading to low abundance in mature communities [21,22,24,48].

1.5. Cyanobacterial Systematics

Table 1. Common terms in cyanobacterial systematics

Term	Definition
Systematics	A general term describing the study of organism diversity and relationships
Taxonomy	A classification system for organisms; an organized schematic for naming and grouping
Phylogeny	The evolutionary history of an organism
Nomenclature	A system of naming for a group of organisms
Taxon	An individual taxonomic group
Clade	A group of organisms that share a common ancestor

Though these morphotypes have been useful in painting a general picture of euendolithic succession, using morphological descriptions to classify cyanobacterial species comes with inherent uncertainty. Unlike all other prokaryotes, cyanobacterial taxonomy is governed by both the International Code of Nomenclature of Prokaryotes[89] (ICNP) and the International Code of Nomenclature for algae, fungi, and plants[90] (ICN). Historically, cyanobacteria were classified as blue green algae, and much of the early work on cyanobacterial classification utilized botanical principles, including identifying new isolates based upon their morphology. The application of microbial culture and molecular phylogeny techniques has since been beneficial in that it not only sped up the identification process, but also the detection and discovery of new organisms. However, the current use of molecular phylogeny techniques has at the same time resulted in the discovery of many "bad" taxa, usually remnants of the early taxonomic system that typically encompass morphologically simple forms. Not unfrequently researchers find that organisms that share a single genus name species, are phylogenetic very distant. Two examples of genera where this occurs commonly are *Leptolyngbya* [91] and *Microcoleus* [92]. This stems from the practice of delineating new taxa by comparing sequence based phylogenies from just a few sequences of known

organisms most closely related to the one being studied, since placing them into comprehensive phylogenies that could provide the big picture would require major efforts. More uncertainty arises from the recent push for the inclusion of non-phototrophic organisms in the Cyanobacterial phylum. Soo et al.[93,94] described two sibling clades, *Melainabacteria* and *Sericytochromatia*, that are the phylogenetically closest non-phototrophic prokaryotes to the previously defined phylum of Cyanobacteria. Though this classification has been contentious within the cyanobacteria research community (see Garcia-Pichel et al. 2019 [95]), their classification as Cyanobacteria has been propagated through the taxonomic databases, leading to confusion by researchers. In addition, other errors in the two most used taxonomic databases, Silva [96] and Greengenes [97], have been well documented [98–100], leading to spurious taxonomic assignments and error amplification.

To overcome these sources of uncertainty, researchers should optimally move away from a laissez faire approach based on sequence similarity algorithms and databases with broad criteria for sequence inclusion, to either manually curate the resulting amplicon sequences after traditional taxonomic assignment, or better, to use a complete phylogenetic perspective based upon curated, organism specific databases. Bioinformaticians have already taken steps to alleviate these issues by developing new algorithms that use the principle of phylogenetic placement. Phylogenetic placement algorithms represent a phylogenetic-accurate and efficient way to perform classification if done on trusted databases. These algorithms were developed to place query sequences (like those from an amplicon survey) onto a precalculated reference phylogenetic tree, inferred from a curated set of reference sequences[101]. Maximum-likelihood based programs such as PPLACER [102], RAxML-EPA [103], and EPA-ng [104] take any query sequence (or a set) supplied by the user, along with a reference package (precalculated

reference phylogenetic tree and alignment of all the curated sequences included in the tree), and generate a placement file (JPLACE) that contains the query sequences placed onto leaves of the reference tree with confidence values. Although phylogenetic placement is a robust method of classification, it has not been widely adopted perhaps due to the time investment required for the creation of reference packages, or the complexity in initial data analysis.

2. Research Objectives

My primary goal was to further the understanding of the carbonate endolithic microbiome using modern microbial ecology methods. Since only a handful of 16S rRNA gene sequences of euendoliths exist, I first wanted to determine a molecular baseline for microbial diversity and contrast the findings with older, naturalistic descriptions to clearly define endolithic genera. Then I would use those findings with time-course experiments to make functional determinations of endolithic lifestyles and characterize colonization dynamics.

3. Approach

The experimental focus of my dissertation was to apply modern high-throughput sequencing, robust bioinformatic methods, and whole community lipid-soluble photosynthetic pigment analysis to corroborate qualitative morphological descriptions with quantitative molecular approaches, and then extend those new characterizations to the assignment of euendolithic potential. I chose to work in coastal marine intertidal ecosystems instead of deep benthic zones to both increase available knowledge on hard carbonate ecology (benthic research mostly focuses on corals) and to address economically relevant bioerosion (euendolith bioerosion of carbonate coasts can lead to

tourism loss; boring of mollusk shells can cause oyster and mussel fishery collapse). Literature has shown that at least morphologically, eukaryotic algae tend to dominate colonization in deeper waters, and thus would not be as relevant in my experimental system, allowing me to focus on prokaryotic endoliths. Isla de Mona (Puerto Rico), and Menorca (Spain) were chosen as sampling locations due to their unique mineralogical composition. Both are carbonate platform islands, that contain limestones and dolomites, providing an opportunity to examine the effect of mineral substrate type on endolithic microbiome composition. It is also advantageous to work on Isla de Mona as it is a protected nature preserve. Only a handful of rangers and biologists inhabit the island during the year. This allowed for my time-course experiments to run undisturbed in a more easily accessible location right off the beach.

My first task was to molecularly determine endolithic community structure and compare it to morphological descriptions from past studies. Baseline endolith diversity was essential to inform further colonization experiments as there only existed three 16S rRNA sequences of euendoliths (encompassing only 2 of the 3 three major morphotypes) and the potential diversity of the community was unknown at the time. The findings of the general survey showed that photosynthetic diversity was much higher than expected. Complexity of this level required comprehensive bioinformatic approaches to identify potential euendoliths, including curated databases of cyanobacteria and all phototrophic APB groups, to not inadvertently eliminate any potential euendolith from analysis. These databases then informed my analyses during time-course colonization experiments and allowed me to identify previously unknown euendolithic groups, greatly expanding upon the understanding of carbonate euendolithic microbiomes.

4. Dissertation Structure

I have included an introductory chapter that reviews current literature and provides the background and framework for my dissertation. Two traditional research chapters follow, which are formatted as stand-alone published works. Third is a non-traditional dissertation product that includes a manuscript, a reference data package, and a web application that were developed and used in the production of the second research chapter. Finally, I have included a comprehensive conclusions chapter that summarizes my main findings and contains a discussion of potential avenues for future work.

4.1. Chapter 1: Dissertation Introduction

This chapter introduces the endolithic habitat, and photosynthetic euendolithic communities, with a special focus on intertidal carbonates. I then discuss anoxygenic phototrophic bacteria, both their habitats and metabolism. Finally, I review our current understanding of euendolithic succession and the state of cyanobacterial systematics.

4.2. Chapter 2: A New Niche for Anoxygenic Phototrophs as Endoliths (Published in Applied and Environmental Microbiology)

This data chapter describes the results of a high throughput 16S rRNA gene survey of intertidal carbonate microbiomes from Isla de Mona, Puerto Rico and Menorca, Spain. I describe the discovery of anoxygenic phototrophs as a major microbial player in these communities, at times exceeding the once thought dominant cyanobacteria.

4.3. Chapter 3: Succession and Colonization Dynamics of Endolithic Phototrophs within Intertidal Carbonates (Published in microorganisms)

This data chapter describes a time-course experiment used to identify if endolithic anoxygenic phototrophic bacteria were pioneer microborers or opportunistic secondary colonizers. Additionally, I wanted to ascertain which photosynthetic endolithic microbes were euendolithic using modern molecular methods. I exposed virgin non-porous carbonate substrates *in situ* in the intertidal zone of Isla De Mona, Puerto Rico under the hypothesis that only euendoliths would be able to initially colonize it. This revealed that endolithic microbiomes, similar in biomass to those of mature natural communities, developed within nine months of exposure. And yet, APB populations were still marginal after this period, suggesting that they are secondary colonizers and not euendolithic. However, as bioerosion proceeded, a diverse set of likely cryptoendolithic cyanobacteria colonized the resulting pore spaces, displacing euendoliths. Endolithic colonization dynamics are thus swift but complex, and involve functionally diverse agents, only some of which are euendoliths.

4.4. Chapter 4: Cydrasil 2, a curated 16S rRNA gene reference package and web app for cyanobacterial sequence placement (In prep, Scientific Data)

This non-traditional chapter describes the development of a database, reference package, and web tool that was essential to molecularly elucidating euendolithic colonization dynamics to among cyanobacteria. The chapter is comprised of multiple components. First, a data-descriptor style manuscript describing the creation, curation, and use of Cydrasil, a curated 16S rRNA gene reference package. Second, the reference package which is comprised of multiple documents that have been included in the

appendix. Finally, the web application, which provides a user-friendly interface to bioinformatics pipeline that utilizes the Cydrasil reference package.

4.5. Chapter 5: Dissertation Conclusions

The conclusions chapter summarizes my main findings of both data chapters and includes a short discussion regarding the application of Cydrasil. I then discuss possible avenues of future research, including preliminary findings of both terrestrial and arctic endolithic microbiomes.

References

1. Edwards, K.J.; Bach, W.; Rogers, D.R. Geomicrobiology of the ocean crust: A role for chemoautotrophic Fe-bacteria. *Biol. Bull.* **2003**, *204*, 180–185, DOI:10.2307/1543555.

2. Singer, E.; Chong, L.S.; Heidelberg, J.F.; Edwards, K.J. Similar microbial communities found on two distant seafloor basalts. *Front. Microbiol.* **2015**, *6*, 1–11, DOI:10.3389/fmicb.2015.01409.

3. Russell, N.C.; Edwards, H.G.M.; Survey, B.A.; Environment, N.; Cross, H.; Road, M.; Oet, C.C.B.; Wynn-Williams, D.D. FT-Raman spectroscopic analysis of endolithic microbial communities from Beacon sandstone in Victoria Land, Antarctica. *Antart. Sci.* **1998**, *10*, 63–74.

4. Budel, B.; Weber, B.; Kuhl, M.; Pfanz, H.; Sultemeyer, D.; Wessels, D. Reshaping of sandstone surfaces by cryptoendolithic cyanobacteria: bioalkalization causes chemical weathering in arid landscapes. *Geobiology* **2004**, *2*, 261–268, DOI:10.1111/j.1472-4677.2004.00040.x.

5. Horath, T.; Bachofen, R. Molecular characterization of an endolithic microbial community in dolomite rock in the central Alps (switzerland). *Microb. Ecol.* **2009**, *58*, 290–306, DOI:10.1007/s00248-008-9483-7.

6. Chacón, E.; Berrendero, E.; Garcia Pichel, F. Biogeological signatures of microboring cyanobacterial communities in marine carbonates from Cabo Rojo, Puerto Rico. *Sediment. Geol.* **2006**, *185*, 215–228, DOI:10.1016/j.sedgeo.2005.12.014.

7. Ramírez-Reinat, E.L.; Garcia-Pichel, F. Prevalence of Ca2+-ATPase-mediated carbonate dissolution among cyanobacterial euendoliths. *Appl. Environ. Microbiol.* **2012**, *78*, 7–13, DOI:10.1128/AEM.06633-11.

8. Friedmann, E.I.; Ocampo, R. Endolithic blue-green algae in the dry valleys: primary producers in the antartic desert ecosystem. *Science (80-.).* **1976**, *193*, 1247–1249.

9. Friedmann, E.I. Endolithic microorganisms in the antarctic cold desert. *Science* **1982**, *215*, 1045–1053, DOI:10.1126/science.215.4536.1045.

10. Stjepko Golubic, Imre Friedmann, Ju, S.; Friedmann, E.I.; Schneider, J. The lithobiontic ecological niche, with special reference to microorganisms. *J Sediment Res* **1981**, DOI:10.1306/212F7CB6-2B24-11D7-8648000102C1865D.

11. Wierzchos, J.; Ascaso, C.; McKay, C.P. Endolithic cyanobacteria in halite rocks from the hyperarid core of the Atacama Desert. *Astrobiology* **2006**, *6*, 415–422, DOI:10.1089/ast.2006.6.415.

12. Dong, H.; Rech, J.A.; Jiang, H.; Sun, H.; Buck, B.J. Endolithnic cyanobacteria in soil gypsum: Occurences in Atacama (Chile), Mojave (United States), and Al-Jafr Basin (Jordan) Deserts. *J. Geophys. Res. Biogeosciences* **2007**, *112*, 1–11, DOI:10.1029/2006JG000385.

13. Ziolkowski, L. a; Wierzchos, J.; Davila, A.F.; Slater, G.F. Radiocarbon evidence of active endolithic microbial communities in the hyperarid core of the Atacama Desert. *Astrobiology* **2013**, *13*, 607–616, DOI:10.1089/ast.2012.0854.

14. De La Torre, J.R.; Goebel, B.M.; Friedmann, E.I.; Pace, N.R. Microbial diversity of cryptoendolithic communities from the McMurdo Dry Valleys, Antarctica. *Appl Env. Microbiol* **2003**, *69*, 3858–3867, DOI:10.1128/AEM.69.7.3858.

15. Palmer, R.J.; Friedmann, E.I. Water relations and photosynthesis in the cryptoendolithic microbial habitat of hot and cold deserts. *Microb. Ecol.* **1990**, *19*, 111–118, DOI:10.1007/BF02015057.

16. Couradeau, E.; Roush, D.; Guida, B.S.; Garcia-Pichel, F. Diversity and mineral substrate preference in endolithic microbial communities from marine intertidal outcrops (Isla de Mona, Puerto Rico). *Biogeosciences* **2017**, *14*, 311–324, DOI:10.5194/bg-14-311-2017.

17. Golubić, S.; Pietrini, A.M.; Ricci, S. Euendolithic activity of the cyanobacterium Chroococcus lithophilus Erc. In biodeterioration of the Pyramid of Caius Cestius, Rome, Italy. *Int. Biodeterior. Biodegrad.* **2015**, *100*, 7–16, DOI:10.1016/j.ibiod.2015.01.019.

18. Kobluk, D.R.; Risk, M.J. Rate and nature of infestation of a carbonate substratum by a boring alga. *J. Exp. Mar. Bio. Ecol.* **1977**, *27*, 107–115, DOI:10.1016/0022-0981(77)90131-9.

19. Ramírez-Reinat, E.L.; Garcia-Pichel, F. Characterization of a marine cyanobacterium that bores into carbonates and the redescription of the genus

Mastigocoleus . *J. Phycol.* **2012**, *48*, 740–749, DOI:10.1111/j.1529-8817.2012.01157.x.

20. Roush, D.; Couradeau, E.; Guida, B.; Neuer, S.; Garcia-Pichel, F. A new niche for anoxygenic phototrophs as endoliths. *Appl. Environ. Microbiol.* **2018**, *84*, AEM.02055-17, DOI:10.1128/AEM.02055-17.

21. Tribollet, A. Dissolution of dead corals by euendolithic microorganisms across the northern Great Barrier Reef (Australia). *Microb. Ecol.* **2008**, *55*, 569–580, DOI:10.1007/s00248-007-9302-6.

22. Grange, J.S.; Rybarczyk, H.; Tribollet, A. The three steps of the carbonate biogenic dissolution process by microborers in coral reefs (New Caledonia). *Environ. Sci. Pollut. Res.* **2015**, *22*, 13625–13637, DOI:10.1007/s11356-014-4069-z.

23. Enochs, I.C.; Manzello, D.P.; Tribollet, A.; Valentino, L.; Kolodziej, G.; Donham, E.M.; Fitchett, M.D.; Carlton, R.; Price, N.N. Elevated Colonization of Microborers at a Volcanically Acidified Coral Reef. *PLoS One* **2016**, *11*, 1–16, DOI:10.1371/journal.pone.0159818.

24. Tribollet, A.; Langdon, C.; Golubic, S.; Atkinson, M. Endolithic microflora are major primary producers in dead carbonate substrates of Hawaiian coral reefs. *J. Phycol.* **2006**, *42*, 292–303, DOI:10.1111/j.1529-8817.2006.00198.x.

25. Guida, B.S.; Bose, M.; Garcia-Pichel, F. Carbon fixation from mineral carbonates. *Nat. Commun.* **2017**, *8*, 1–6, DOI:10.1038/s41467-017-00703-4.

26. Shachak, M.; Jones, C.G.; Granot, Y. Herbivory in rocks and the weathering of a desert. *Science (80-.).* **1987**, *236*, 1098–1099, DOI:10.1126/science.236.4805.1098.

27. Vogel, K.; Gektidis, M.; Golubic, S.; Kiene, W.E.; Radtke, G. Experimental studies on microbial bioerosion at Lee Stocking Island, Bahamas and One Tree Island, Great Barrier Reef, Australia: Implications for paleoecological reconstructions. *Lethaia* **2000**, *33*, 190–204, DOI:10.1080/00241160025100053.

28. Tribollet, A.; Payri, C. Bioerosion of the coralline alga Hydrolithon onkodes by microborers in the coral reefs of Moorea, French Polynesia. *Oceanol. Acta* **2001**, *24*, 329–342, DOI:10.1016/S0399-1784(01)01150-1.

29. Ford, D.; Williams, P. The Global Distribution Of Karst. In *Karst Hydrogeology and Geomorphology*; 2013; pp. 1–8 ISBN 9781118684986.

30. Ćurin, M.; Peharda, M.; Calcinai, B.; Golubić, S. Incidence of damaging endolith infestation of the edible mytilid bivalve Modiolus barbatus. *Mar. Biol. Res.* **2014**, *10*, 179–189, DOI:10.1080/17451000.2013.814793.

31. Pfister, C.A.; Meyer, F.; Antonopoulos, D.A. Metagenomic profiling of a microbial assemblage associated with the california mussel: A node in networks of carbon

and nitrogen cycling. *PLoS One* **2010**, *5*, 1–10, DOI:10.1371/journal.pone.0010518.

32. Kaehler, S. Incidence and distribution of phototrophic shell-degrading endoliths of the brown mussel *Perna perna* . *Mar. Biol.* **1999**, *135*, 505–514, DOI:10.1007/s002270050651.

33. Raghukumar, C.; Sharma, S.; Lande, V. Distribution and biomass estimation of shell-boring algae in the intertidal at Goa, India. *Phycologia* **1991**, *30*, 303–309, DOI:10.2216/i0031-8884-30-3-303.1.

34. Kiene, W.; Radtke, G.; Gektidis, M.; Golubić, S.; Vogel, K. Factors controlling the distribution of microborers in Bahamian Reef environments. In *Facies*; Schuhmacher, H., Kiene, W., Dullo, W.-C., Eds.; 1995; pp. 174–188.

35. Gektidis, M. Development of microbial euendolithic communities : The influence of light and time. *Bull. Geol. Soc. Denmark*. **1999**, *45*, 147–150.

36. Guida, B.S.; Garcia-Pichel, F. Extreme cellular adaptations and cell differentiation required by a cyanobacterium for carbonate excavation. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 5712–5717, DOI:10.1073/pnas.1524687113.

37. Knoll, A.H.; Golubic, S.; Green, J.; Swett, K. Organically preserved microbial endoliths from the late Proterozoic of East Greenland. *Nature* **1986**, *321*, 856–857, DOI:10.1038/321856a0.

38. Zhang, X. guang; Pratt, B.R. Microborings in Early Cambrian phosphatic and phosphatized fossils. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **2008**, *267*, 185–195, DOI:10.1016/j.palaeo.2008.06.015.

39. Radtke, G.; Glaub, I.; Vogel, K.; Golubic, S. isp. nov., Distribution, Variability and Biological Origin. *Ichnos* **2010**, *17*, 25–33, DOI:10.1080/10420940903358628.

40. Glaub, I. Paleobathymetric reconstructions and fossil microborings. *Bull. Geol. Soc. Denmark* **1999**, *45*, 143–146.

41. Elias, R.J.; Lee, D. Microborings and Growth in Late Ordovician Halysitids and Other Corals. *J. Paleontol.* **1993**, *67*, 922–934.

42. Kölliker, A. On the frequent occurrence of vegetable parasites in the hard structures of animals. *Proc. R. Soc. London* **1859**, *10*, 95–99, DOI:10.5962/bhl.title.103118.

43. Bornet, E.; Flahault, C. Note sur deux nouveaux genres d'algues perforantes. *J. Bot.* **1888**.

44. Duerden, J.E. Boring algae as agents in the disintegration of corals. *Bull. Am. Museum Nat. Hist.* **1902**, *889*, 323–332.

45. Ercegović, A. La végétation lithophytes sur les calcaires et les dolomites en Croatie. *Acta Bot. Croat.* **1925**.

46. Goiubic, S. Distribution, taxonomy, and boring patterns of marine endolithic algae. *Integr. Comp. Biol.* **1969**, DOI:10.1093/icb/9.3.747.

47. Budd, D.A.; Perkins, R.D. Bathymetric zonation and paleoecological significance of algal microborings in Puerto Rican shelf and slope sediments. *J. Sediment. Petrol.* **1980**, DOI:10.1306/212F7B17-2B24-11D7-8648000102C1865D.

48. Chazottes, V.; Campion-Alsumard, T. Le; Peyrot-Clausade, M. Bioerosion rates on coral reefs: interactions between macroborers, microborers and grazers (Moorea, French Polynesia). *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **1995**, *113*, 189–198, DOI:10.1016/0031-0182(95)00043-L.

49. Pantazidou, A.; Louvrou, I.; Economou-Amilli, A. Euendolithic shell-boring cyanobacteria and chlorophytes from the saline lagoon Ahivadolimni on Milos Island, Greece. *Eur. J. Phycol.* **2006**, DOI:10.1080/09670260600649420.

50. Wisshak, M.; Tribollet, A.; Golubic, S.; Jakobsen, J.; Freiwald, A. Temperate bioerosion: Ichnodiversity and biodiversity from intertidal to bathyal depths (Azores). *Geobiology* **2011**, DOI:10.1111/j.1472-4669.2011.00299.x.

51. Campbell, S.E. The modern distribution and geological history of calcium carbonate boring microorganisms. *Biominer. Biol. Met. Accumul.* **1983**, 99–104, DOI:10.1007/978-94-009-7944-4.

52. Perkins, R.D.; Tsentas, C.I. Microbial infestation of carbonate substrates planted on the St. Croix shelf, West Indies. *Bull. Geol. Soc. Am.* **1976**, *87*, 1615–1628, DOI:10.1130/0016-7606(1976)87<1615:MIOCSP>2.0.CO;2.

53. Campbell, S.E.; Cole, K. Developmental studies on cultured endolithic conchocelis (Rhodophyta). *Hydrobiologia* **1984**, *116/117*, 201–208, DOI:10.1007/BF00027666.

54. Brito, Â.; Ramos, V.; Seabra, R.; Santos, A.; Santos, C.L.; Lopo, M.; Ferreira, S.; Martins, A.; Mota, R.; Frazão, B.; et al. Culture-dependent characterization of cyanobacterial diversity in the intertidal zones of the Portuguese coast: A polyphasic study. *Syst. Appl. Microbiol.* **2012**, *35*, 110–119, DOI:10.1016/j.syapm.2011.07.003.

55. Foster, J.S.; Green, S.J.; Ahrendt, S.R.; Golubic, S.; Reid, R.P.; Hetherington, K.L.; Bebout, L. Molecular and morphological characterization of cyanobacterial diversity in the stromatolites of Highborne Cay, Bahamas. *ISME J.* **2009**, *3*, 573–587, DOI:10.1038/ismej.2008.129.

56. Walker, J.J.; Spear, J.R.; Pace, N.R. Geobiology of a microbial endolithic community in the Yellowstone geothermal environment. *Nature* **2005**, *434*, 1011–1014, DOI:10.1038/nature03447.

57. Walker, J.J.; Pace, N.R. Endolithic Microbial Ecosystems. *Annu. Rev. Microbiol.* **2007**, *61*, 331–347, DOI:10.1146/annurev.micro.61.080706.093302.

58. Crits-Christoph, A.; Gelsinger, D.R.; Ma, B.; Wierzchos, J.; Ravel, J.; Davila, A.; Casero, M.C.; DiRuggiero, J. Functional interactions of archaea, bacteria and viruses in a hypersaline endolithic community. *Environ. Microbiol.* **2016**, *18*, 2064–2077, DOI:10.1111/1462-2920.13259.

59. Couradeau, E.; Roush, D.; Guida, B.S.; Garcia-Pichel, F. Diversity and mineral substrate preference in endolithic microbial communities from marine intertidal outcrops (Isla de Mona, Puerto Rico). *Biogeosciences* **2017**, *14*, 311–324, DOI:10.5194/bg-14-311-2017.

60. Belnap, J.; Büdel, B.; Lange, O.L. Biological Soil Crusts: Characteristics and Distribution. In *Biological Soil Crusts: Structure, Function, and Management*; 2003; Vol. 150, pp. 3–30 ISBN 978-3-540-43757-4.

61. Couradeau, E.; Benzerara, K.; Moreira, D.; Gérard, E.; Kaźmierczak, J.; Tavera, R.; López-García, P. Prokaryotic and eukaryotic community structure in field and cultured microbialites from the alkaline Lake Alchichica (Mexico). *PLoS One* **2011**, *6*, DOI:10.1371/journal.pone.0028767.

62. Magnusson, S.H.; Fine, M.; Kühl, M. Light microclimate of endolithic phototrophs in the scleractinian corals *Montipora monasteriata* and *Porites cylindrica* . *Mar. Ecol. Prog. Ser.* **2007**, *332*, 119–128, DOI:10.3354/meps332119.

63. Yang, S.H.; Lee, S.T.M.; Huang, C.R.; Tseng, C.H.; Chiang, P.W.; Chen, C.P.; Chen, H.J.; Tang, S.L. Prevalence of potential nitrogen-fixing, green sulfur bacteria in the skeleton of reef-building coral *Isopora palifera* . *Limnol. Oceanogr.* **2016**, *61*, 1078–1086, DOI:10.1002/lno.10277.

64. Overmann, J.; Garcia-Pichel, F. The phototrophic way of life. In *The Prokaryotes: Prokaryotic Communities and Ecophysiology*; 2013; Vol. 9783642301, pp. 203–257 ISBN 9783642301230.

65. Koblížek, M. Ecology of aerobic anoxygenic phototrophs in aquatic environments. *FEMS Microbiol. Rev.* **2015**, *39*, 854–870, DOI:10.1093/femsre/fuv032.

66. Bryant, D.A.; Costas, A.M.G.; Maresca, J.A.; Chew, A.G.M.; Klatt, C.G.; Bateson, M.M.; Tallon, L.J.; Hostetler, J.; Nelson, W.C.; Heidelberg, J.F.; et al. *Candidatus* Chloracidobacterium thermophilum: An Aerobic Phototrophic Acidobacterium. *Science (80-.).* **2007**, *317*, 523–526, DOI:10.1126/science.1143236.

67. Ward, L.M.; McGlynn, S.E.; Fischera, W.W. Draft genome sequence of chloracidobacterium sp. CP2_5A, a phototrophic member of the phylum acidobacteria recovered from a Japanese hot spring. *Genome Announc.* **2017**, *5*, DOI:10.1128/genomeA.00821-17.

68. Zeng, Y.; Feng, F.; Medova, H.; Dean, J.; Kobli ek, M. Functional type 2 photosynthetic reaction centers found in the rare bacterial phylum Gemmatimonadetes. *Proc. Natl. Acad. Sci.* **2014**, *111*, 7795–7800, DOI:10.1073/pnas.1400295111.

69. Imhoff, J.F. The family Chlorobiaceae. In *The Prokaryotes: Other Major Lineages of Bacteria and The Archaea*; 2014; Vol. 9783642389, pp. 501–514 ISBN 9783642301230.

70. Imhoff, J.F. The family Chromatiaceae. In *The Prokaryotes: Gammaproteobacteria*; 2014; Vol. 9783642389, pp. 151–178 ISBN 9783642389221.

71. Oren, A. The family ectothiorhodospiraceae. *The Prokaryotes: Gammaproteobacteria* **2014**, *9783642389*, 199–222, DOI:10.1007/978-3-642-38922-1_248.

72. Nicholson, J.A.; Stolz, J.F.; Pierson, B.K. Structure of a microbial mat at Great Sippewisset Marsh, Cape Cod, Masschusetts. *Fems Microbiol. Ecol.* **1987**, *45*, 343–364.

73. Manske, A.K.; Glaeser, J.; Kuypers, M.M.M.; Overmann, J. Physiology and phylogeny of green sulfur bacteria forming a monospecific phototrophic assemblage at a depth of 100 meters in the Black Sea. *Appl. Environ. Microbiol.* **2005**, *71*, 8049–8060, DOI:10.1128/AEM.71.12.8049-8060.2005.

74. Findlay, A.J.; Bennett, A.J.; Hanson, T.E.; Luther, G.W. Light-dependent sulfide oxidation in the anoxic zone of the chesapeake bay can be explained by small populations of phototrophic bacteria. *Appl. Environ. Microbiol.* **2015**, *81*, 7560–7569, DOI:10.1128/AEM.02062-15.

75. Lauro, F.M.; DeMaere, M.Z.; Yau, S.; Brown, M. V; Ng, C.; Wilkins, D.; Raftery, M.J.; Gibson, J.A.; Andrews-Pfannkoch, C.; Lewis, M.; et al. An integrative study of a meromictic lake ecosystem in Antarctica. *ISME J.* **2011**, *5*, 879–895, DOI:10.1038/ismej.2010.185.

76. Hanada, S. The phylum Chloroflexi, the family Chloroflexaceae, and the related phototrophic families Oscillochloridaceae and Roseiflexaceae. In *The Prokaryotes: Other Major Lineages of Bacteria and The Archaea*; 2014; Vol. 9783642389, pp. 515–532 ISBN 9783642301230.

77. Madigan, M.T.; Jung, D.O. An Overview of Purple Bacteria: Systematics, Physiology, and Habitats. *Purple Phototrophic Bact.* **2009**, *28*, 1–15, DOI:10.1007/978-1-4020-8815-5.

78. Castanier, S.; Le Métayer-Levrel, G.; Perthuisot, J.P. Ca-carbonates precipitation and limestone genesis - the microbiogeologist point of view. *Sediment. Geol.* **1999**, *126*, 9–23, DOI:10.1016/S0037-0738(99)00028-7.

79. Bosak, T.; Greene, S.E.; Newman, D.K. A likely role for anoxygenic photosynthetic microbes in the formation of ancient stromatolites. *Geobiology* **2007**, *5*, 119–126, DOI:10.1111/j.1472-4669.2007.00104.x.

80. Bundeleva, I.A.; Shirokova, L.S.; Bénézeth, P.; Pokrovsky, O.S.; Kompantseva, E.I.; Balor, S. Calcium carbonate precipitation by anoxygenic phototrophic bacteria. *Chem. Geol.* **2012**, *291*, 116–131, DOI:10.1016/j.chemgeo.2011.10.003.

81. Li, R.Y.; Fang, H.H.P. Hydrogen production characteristics of photoheterotrophic *Rubrivivax gelatinosus* L31. *Int. J. Hydrogen Energy* **2008**, *33*, 974–980, DOI:10.1016/j.ijhydene.2007.12.001.

82. de Wit, R.; van Gemerden, H. Chemolithotrophic growth of the phototrophic sulfur bacterium Thiocapsa roseopersicina. *FEMS Microbiol. Lett.* **1987**, *45*, 117–162, DOI:10.1016/0378-1097(87)90033-4.

83. Le Campion-Alsumard, T. Étude Expérimentale De La Colonisation D'Éclats De Calcite Par Les Cyanophycées Endolithes Marines. *Cah. Biol. Mar.* **1975**, *16*, 177–185.

84. Le Campion-Alsumard, T. Les cyanophycées endolithes marines--Systématique, ultrastructure, écologie et biodestruction. *Ocean. Acta* **1979**, *2*, 143–156.

85. Tribollet, A.; Golubic, S. Cross-shelf differences in the pattern and pace of bioerosion of experimental carbonate substrates exposed for 3 years on the northern Great Barrier Reef, Australia. *Coral Reefs* **2005**, *24*, 422–434, DOI:10.1007/s00338-005-0003-7.

86. Vogel, K.; Gektidis, M.; Golubic, S.; Kiene, W.E.; Radtke, G. Experimental studies on microbial bioerosion at Lee Stocking Island, Bahamas and One Tree Island, Great Barrier Reef, Australia: Implications for paleoecological reconstructions. *Lethaia* **2000**, *33*, 190–204, DOI:10.1080/00241160025100053.

87. Garcia-Pichel, F.; Ramirez-Reinat, E.; Gao, Q. Microbial excavation of solid carbonates powered by P-type ATPase-mediated transcellular Ca2+ transport. *Proc. Natl. Acad. Sci.* **2010**, *107*, 21749–21754, DOI:10.1073/pnas.1011884108.

88. Guida, B.S.; Garcia-Pichel, F. Draft Genome Assembly of a Filamentous Euendolithic (True Boring) Cyanobacterium, Mastigocoleus testarum Strain BC008. *Genome Announc.* **2016**, *4*, 1–2, DOI:10.1128/genomeA.01574-15.Copyright.

89. Parker, C.T.; Tindall, B.J.; Garrity, G.M. International code of nomenclature of Prokaryotes. *Int. J. Syst. Evol. Microbiol.* **2019**, DOI:10.1099/ijsem.0.000778.

90. Turland, N.J.; Wiersema, J.H.; Barrie, F.R.; Greuter, W.; Hawksworth, D.L.; Herendeen, P.S.; Knapp, S.; Kusber, W.-H.; Li, D.-Z.; Marhold, K.; et al. *International Code of Nomenclature for algae, fungi, and plants (Shenzhen Code) adopted by the Nineteenth International Botanical Congress Shenzhen, China, July 2017*; 2018; ISBN 9783946583165.

91. Stoyanov, P.; Moten, D.; Mladenov, R.; Dzhambazov, B.; Teneva, I. Phylogenetic relationships of some filamentous cyanoprokaryotic species. *Evol. Bioinforma.* **2014**, *10*, 39–49, DOI:10.4137/EBo.s13748.

92. Boyer, S.L.; Johansen, J.R.; Flechtner, V.R.; Howard, G.L. Phylogeny and genetic variance in terrestrial Microcoleus (Cyanophyceae) species based on sequence analysis of the 16S rRNA gene and associated 16S-23S its region. *J. Phycol.* **2002**, *38*, 1222–1235, DOI:10.1046/j.1529-8817.2002.01168.x.

93. Soo, R.M.; Skennerton, C.T.; Sekiguchi, Y.; Imelfort, M.; Paech, S.J.; Dennis, P.G.; Steen, J.A.; Parks, D.H.; Tyson, G.W.; Hugenholtz, P. An expanded genomic representation of the phylum cyanobacteria. *Genome Biol. Evol.* **2014**, *6*, 1031–1045, DOI:10.1093/gbe/evu073.

94. Soo, R.M.; Hemp, J.; Hugenholtz, P. Evolution of photosynthesis and aerobic respiration in the cyanobacteria. *Free Radic. Biol. Med.* **2019**, *140*, 200–205, DOI:10.1016/j.freeradbiomed.2019.03.029.

95. Garcia-Pichel, F.; Zehr, J.P.; Bhattacharya, D.; Pakrasi, H.B. What's in a name? The case of cyanobacteria. *J. Phycol.* **2020**, *56*, 1–5, DOI:10.1111/jpy.12934.

96. Quast, C.; Pruesse, E.; Yilmaz, P.; Gerken, J.; Schweer, T.; Yarza, P.; Peplies, J.; Glöckner, F.O. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **2013**, *41*, 590–596, DOI:10.1093/nar/gks1219.

97. DeSantis, T.Z.; Hugenholtz, P.; Larsen, N.; Rojas, M.; Brodie, E.L.; Keller, K.; Huber, T.; Dalevi, D.; Hu, P.; Andersen, G.L. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **2006**, *72*, 5069–5072, DOI:10.1128/AEM.03006-05.

98. Edgar, R. Taxonomy annotation and guide tree errors in 16S rRNA databases. *PeerJ* **2018**, *2018*, DOI:10.7717/peerj.5030.

99. Park, S.-C.; Won, S. Evaluation of 16S rRNA Databases for Taxonomic Assignments Using a Mock Community. *Genomics Inform.* **2018**, *16*, e24, DOI:10.5808/gi.2018.16.4.e24.

100. Lydon, K.A.; Lipp, E.K. Taxonomic annotation errors incorrectly assign the family Pseudoalteromonadaceae to the order Vibrionales in Greengenes: Implications for microbial community assessments. *PeerJ* **2018**, *2018*, DOI:10.7717/peerj.5248.

101. Felsenstein, J. Evolutionary trees from DNA sequences: A maximum likelihood approach. *J. Mol. Evol.* **1981**, DOI:10.1007/BF01734359.

102. Matsen, F.A.; Kodner, R.B.; Armbrust, E.V. pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics* **2010**, DOI:10.1186/1471-2105-11-538.

103. Berger, S.A.; Krompass, D.; Stamatakis, A. Performance, accuracy, and web server for evolutionary placement of short sequence reads under maximum likelihood. *Syst. Biol.* **2011**, *60*, 291–302, DOI:10.1093/sysbio/syr010.

104. Barbera, P.; Kozlov, A.M.; Czech, L.; Morel, B.; Darriba, D.; Flouri, T.; Stamatakis, A. EPA-ng: Massively Parallel Evolutionary Placement of Genetic Sequences. *Syst. Biol.* **2019**, *68*, 365–369, DOI:10.1093/sysbio/syy054.

CHAPTER 2

A NEW NICHE FOR ANOXYGENIC PHOTOTROPHS AS ENDOLITHS

Published in *Applied and Environmental Microbiology,* **2018**

Coauthors have acknowledged the use of this manuscript in my dissertation

Authors:

Daniel Roush, Estelle Couradeau, Brandon Guida, Susanne Neuer, and Ferran Garcia-Pichel

Abstract: Anoxygenic phototrophic bacteria (APBs) occur in a wide range of aquatic habitats, from hot springs to freshwater lakes and intertidal microbial mats. Here we report the discovery of a novel niche for APBs: endoliths within marine littoral carbonates. In a study of 40 locations around Isla de Mona, Puerto Rico, and Menorca, Spain, 16S rRNA high-throughput sequencing of endolithic community DNA revealed the presence of abundant phylotypes potentially belonging to well-known APB clades. An *ad hoc* phylogenetic classification of these sequences allowed us to refine the assignments more stringently. Even then, all locations contained such putative APBs, often reaching a significant proportion of all phototrophic sequences. In fact, in some 20% of samples, their contribution exceeded that of oxygenic phototrophs, previously regarded as the major type of endolithic microbe in carbonates. The communities contained representatives of APBs in the Chloroflexales, various Proteobacterial groups, and Chlorobi. The most abundant phylotypes varied with geography: on Isla de Mona, *Roseiflexus* and *Chlorothrix*-related phylotypes dominated, whereas those related to

Erythrobacter were the most common in Menorca. The presence of active populations of APBs was corroborated through an analysis of photopigments: bacteriochlorophylls were detected in all samples, bacteriochlorophyll *c* and *a* being most abundant. We discuss the potential metabolism and geomicrobial roles of endolithic APBs. Phylogenetic inference suggests that APBs may be playing a role as photoheterotrophs, adding biogeochemical complexity to our understanding of such communities. Given the global extent of coastal carbonate platforms, they likely represent a very large and unexplored habitat for APBs.

---

1. Introduction

Over the past two hundred years, naturalists have extensively studied the endolithic habitat within intertidal carbonates. Evidence from as early as the mid-1800s from Agassiz, reported by Duerden [1] and Kölliker [2] describe the presence of vegetable [sic] parasites within mollusk shells and corals. These descriptions eventually extended to a range of substrates and settings, including marine carbonates [3–7], terrestrial limestones and marbles [8], corals [9,10] and microbialites [11]. Much of the work focused on the boring algae and cyanobacteria, known as euendoliths, that can actively penetrate the carbonate substrate to establish a home within the solid rock. The crypto- and chasmoendoliths, which colonize pore spaces and cracks [12] respectively, have received less attention but are undoubtedly common. Endolithic communities play significant bio-erosive roles in the natural environment [13,14], can become pests of bivalve fisheries [15–17] and, judging by the presence of microfossils in the rock record, have been active in their roles since the Precambrian [18,19].

The deployment of early molecular methods for community fingerprinting (clone libraries, DGGE) provided expanded accounts of marine and terrestrial endolithic communities of carbonates as well as other substrates [7,20–22]. They revealed that the

endolithic habitat can harbor complex communities of microbes, with important heterotrophic components, particularly when the substrate rock is naturally porous, or when it has been made porous through excavation by euendoliths. This level of complexity was clear in the first high-throughput, multi-sample survey of community diversity from intertidal outcrops [23], which we conducted on Isla de Mona, Puerto Rico. Results of that survey made two things apparent: the diversity and complexity of these communities had been dramatically underreported in the literature, and they could host a potentially wide range of metabolic niches previously unrecognized in this environment [23]. The level of diversity found in endolithic communities was comparable to that of other microbial communities such as biological soil crusts [24] and microbialites [25], containing representatives of a variety of microbial metabolisms, from fermenters to sulfur oxidizers. Among these one could discern many phylotypes potentially allied with anoxygenic phototrophic bacteria (APB).

Anoxygenic phototrophic bacteria are a phylogenetically widespread, metabolic guild distributed among in six different bacterial phyla. Canonically, APBs have been delineated into five groups; the green sulfur bacteria in the phylum Chlorobi, the green non-sulfur bacteria of the phylum Chloroflexi, the purple sulfur bacteria of the Gammaproteobacteria, the purple non-sulfur bacteria of the Alpha- and Betaproteobacteria, and the *Heliobacteria* within the Firmicutes [26]. Over the past two decades, new APBs have been discovered, including the aerobic anoxygenic phototrophic bacteria found within the Alpha- and Betaproteobacteria [27], along with one representative each in the Acidobacteria [28] and the Gemmatimonadetes [29]. Some groups of APBs are typically found within a narrow range of habitats: green and purple sulfur bacteria are limited to locations in which there is a ready supply of an electron donor, typically hydrogen sulfide, like the anoxic bottom of meromictic lakes, ocean

sediments, hot springs [30–32], microbial mats [33], stratified marine and estuarine waters [34,35], subglacial lakes [36]. Green non-sulfur bacteria have been found in hot springs, hypersaline microbial mats, and in some marine sediments [37]. Purple non-sulfur bacteria occur in a wider range of habitats, including the open ocean and Antarctic lakes, but they rarely are the dominant phototroph within a community [38]. Here we report that the endolithic habitat must be now added to this short list.

2. Materials and Methods

2.1. Sampling Collection

Samples from intertidally exposed hard carbonate rock were collected from Isla de Mona (18.0867° N, 67.8894° W), a small (11 by 7 km) carbonate island 66 km west of Puerto Rico having obtained permits from the Departamento de Recursos Naturales y Ambientales (Commonwealth of Puerto Rico) and from Menorca (39° 58′ 0″ N, 4° 5′ 0″ E), a populated carbonate-rich island 200 km east the Iberian Peninsula (Figure 1). The study did not involve endangered or protected species. Rock samples were broken off from large boulders or cliff walls using a geological hammer, which was washed in seawater near the respective sampling site. Most samples were collected within the intertidal notch typical of carbonate cliffs, but a few were collected by SCUBA diving below tidal ranges (only at Mona, K samples). Initial samples were then aliquoted with ethanol-sterilized chisel and hammer. Samples were divided for mineralogical analysis (kept air dried), and either preserved in 70% ethanol (Mona) or frozen at -80□C (Menorca) for eventual DNA and lipid-soluble pigment extraction, then shipped at room temperature (Mona) or in liquid nitrogen (Menorca), reaching the laboratory in less than a week, and stored frozen at -80°C until analysis.

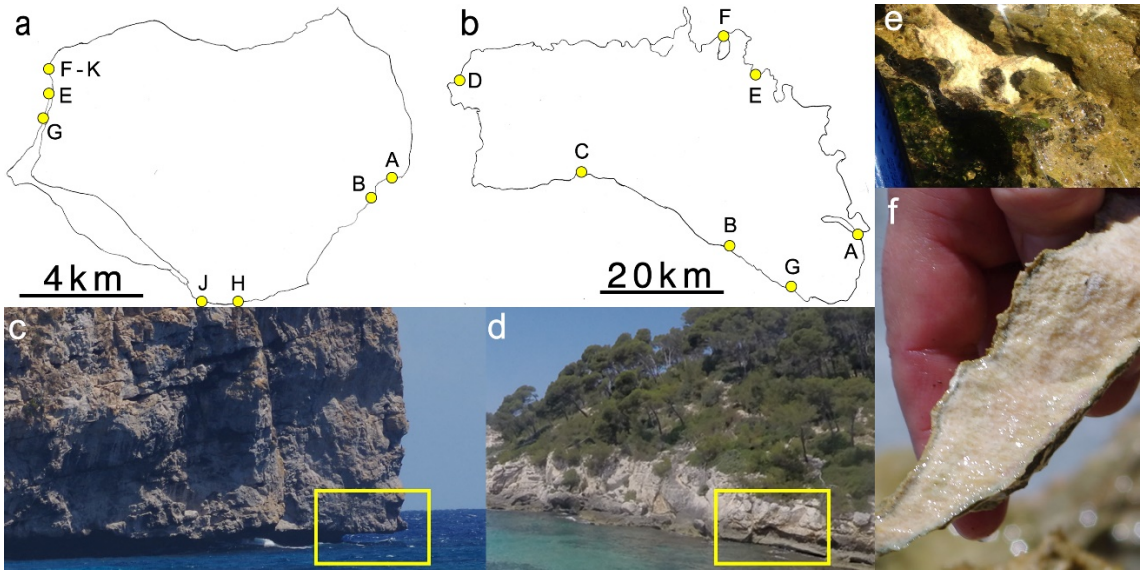

Figure 1. Sampling locations (yellow dots) in Isla de Mona (**a**) and Menorca (**b**), chosen primarily by accessibility, and to encompass a variety of substrate mineral composition. (**c**) Isla de Mona limestone/dolostone cliffs, with characteristic intertidal notch (box). (**d**) Limestone cliffs from Menorca showing intertidal notch (box). (**e**) Close up of typical Menorca sampling site, showing chipped off area with a pale coloration. (**f**) A sample rock chip of Isla de Mona raised-reef, showing green endolithic growth below the surface.

2.2. X-ray Diffraction

Subsamples were ground into a fine powder with a small amount of 100% ethanol. XRD patterns were collected using a Panalytical X'Pert Pro diffractometer mounted in the Debye-Scherrer configuration with a CuKα monochromatic X-Ray source. Continuous scan mode was used in the 10–90° 2θ range to collect data. Relative mineral composition was identified using the X'Pert High Score plus software with a Rietvald refinement using default parameters.

2.3. Confocal Microscopy

Selected subsamples were dissolved over a period of 48 h in 200 mM, pH 5 EDTA (for Calcite or Aragonite) or 250 mM pH 5 CDTA (for Dolomite) [39]. The resulting

liquid was then filtered using a GE black PC 0.8 μM filter. The filter was placed onto a drop of immersion oil on a microscope slide, and covered with a second drop, over which, a cover slip was placed, and kept at -20°C until imaging. Images were collected using a Leica SP5 confocal microscope at 1024 x 1024 resolution, minimum line average of 1, and a scan rate maximum of 400 hz. Samples were excited using a 405 nm wavelength laser. Innate Chl a emission (the "red" channel) was collected between 660 nm and 690 nm. Near-infrared emission (the "green" channel) was collected between 740 nm and 800 nm. A maximum intensity z projection of each channel was then visualized using FIJI [40].

2.4. Endolithic Community DNA Extraction

Samples for DNA extraction were brushed aggressively with sterile toothbrushes and sterilized MilliQ water to remove any epilithic biomass, which was already very sparse. Samples were chipped to get rid of the deep rock, and to obtain smaller volumes of the top several mm, where the endolithic phototrophic biomass was conspicuous by color. To ensure a consistent sampling effort, we measured and cut a piece of a chip with a surface area of 8 cm2. The pieces were then ground using sterile mortars as described in Wade and Garcia-Pichel 2003 [39], and 0.5 g of the sample were placed into the bead tube of a MoBio PowerPlant Pro kit (Mo Bio Laboratories, Inc., Carlsbad, CA, USA) for DNA extraction. We followed the protocol provided with one exception: prior to the first lysis step, we homogenized the bead tubes horizontally on a vortex at 2,200 rev/min for 10 minutes, and added 7 freeze-thaw cycles to ensure better disruption of the endolithic cells. DNA in the extract was quantified using a Qubit 2.0 HS DNA Kit (Life Technologies, Carlsbad, CA, USA).

2.5. 16S rRNA Gene Library Preparation and Illumina Sequencing

The V3 - V4 variable region of the 16S rRNA gene was targeted for amplification using PCR primers 341F (5'-CCTACGGGNGGCWGCAG) [41] and 806R (5'-GGACTACVSGGGTATCTAAT) [42] with a barcoded forward primer. PCR amplification was performed using the HotStartTaq Plus Master Mix Kit (Qiagen, USA) under the following parameters: 94°C for 3 minutes, followed by 28 cycles of 94°C for 30 seconds, 53°C for 40 seconds and 72°C for 1 minute, followed by a final 5 min elongation step at 72°C. The PCR product was then further purified and pooled to generate a single DNA library using the Illumina TruSeq DNA library preparation protocol. The library was sequenced using an Illumina MiSeq following the manufacturer's guidelines. The library preparation, sequencing paired ends assembly, and first quality trimming (with phred score of Q25 cutoff) were performed commercially (MrDNALab, Shallowater, TX, USA).

2.6. Bioinformatic Pipeline

Paired sequences were then processed using the QIIME 1.9 analysis pipeline [43]. First, chimera sequences were detected and removed utilizing the VSEARCH de novo chimera checking algorithm [44]. Next, we ran the split_libraries.py script using default parameters (removal of barcodes, removal of sequences below 200bp, removal of sequences containing homopolymer runs longer than 6) to prep the dataset for Operational Taxonomic Unit (OTU) picking, for which we used pick_open_reference_otus.py script with modified parameters (see supplementary information for full parameter file). Briefly, OTUs were clustered at 97% using SortMeRNA [45] and SUMAclust [46] , taxonomy was assigned using SortMeRNA and reference Green Genes 13_8 release database [47]. Singletons were removed after OTU picking.

2.7. Reference Phylogenetic Tree Building and OTU Placement

To determine which OTUs were likely phototrophic, we imposed the rule that for a given OTU to be considered phototrophic, it would have to phylogenetically belong to a clade that was composed of only known phototrophs. To facilitate this task, we constructed reference trees for each of the bacterial groups present in our samples known to contain phototrophic clades. Trees contained only sequences obtained from bona fide cultured isolates of known metabolism. Representative sequences for each tree were obtained from the SILVA SSU database [48], aligned using MAFFT [49] and Guidance2 [50]. The Guidance2 alignment with low scoring columns removed for each group was then used for all further analysis. Reference trees were constructed on the CIPRES high performance computing cluster [51] using the RAxML-HPC2 [52] workflow with the ML+Thorough bootstrap (1000 bootstraps) method and the GTRGAMMA model. In all we constructed 9 trees representing the following taxa: *Chloroflexiaceae*, *Chlorobiaceae*, *Rhodospirilliaceae*, *Rhodocyclaceae*, *Comamonadaceae*, *Rhodobacteraceae*, *Rhizobiales*, *Erythrobacteraceae*, and *Chromatiales*.

Next, we filtered the OTU table to include all OTUs that had been previously (see above) automatically assigned to taxa that contained known phototrophic bacteria. OTU representative sequences were then aligned to the appropriate reference alignment using PaPaRa [53], and placed into the edges of the respective reference trees using the Evolutionary Placement Algorithm (based on the maximum-likelihood model) feature of RAxML8 [54]. Placement trees were visualized using the ITOL3 website [55]. An OTU was considered a likely phototroph if it was placed within a phototroph clade (node) on the respective reference phylogenetic tree with better than 70% certainty.

2.8. Pigment Extraction and Analysis

Lipid soluble pigments were extracted as follows: 3 g of powdered sample (same sample as used for DNA) was suspended in a 7:2 Acetone:Methanol mixture, and sonicated twice for 30 s in an ice bath in the dark. Extracts were centrifuged at 2100 G for 10 m, decanted, and the supernatant filtered through a 0.22 μm nylon filter. These steps were repeated until the supernatant was devoid of color. The resulting filtered supernatant was then dried under a N_2 stream in the dark and then resuspended in 100% HPLC-grade acetone. HPLC analysis was conducted on an Agilent 1100 with an online photodiode array detector, using the protocol of Frigaard et al. [56] on a Novapak C18 3.9 x 300 mm (60 Å pore size, 4 μm particles) column. The gradient was composed of solvent A (methanol:acetonitrile: water, 42:33:25 by vol) and solvent B (methanol:acetonitrile:ethyl acetate, 50:20:30 by vol), and elution was performed as follows: at the time of injection 30% B, linear increase to 100% in 52 min, constant for 15 min, and a return to 30% in 2 min. The flow rate was 1.0 ml min^{-1}, and the column temperature was 30°C. Pigment identification was done by comparison of retention time and spectrum against true standards of Chl *a* and Bchl *a* from Sigma Aldrich. All other pigments were identified from known spectra [57] and from extracts of Chloroflexus auranticus grown anaerobically. Injected pigment mass was calculated from the chromatogram using the equation $m = FA (em\ d)^{-1}$, where m is the mass of BChl or Chl in milligrams, *F* is the solvent flow rate (1 mL min^{-1}), A is the peak area (in Au), e_m is the extinction coefficient in L mg^{-1} cm^{-1}, and d is the path length of the PDA detector (1 cm). Extinction coefficients were taken from Ley *et al.* 2006. We then converted the masses to mg m^{-2} using a per-sample surface area to volume ratio.

2.9. Data Availability

Isla de Mona sequences are deposited under GenBank KT972744-KT981874. Menorca sequences are deposited under GenBank BioProjectID PRJNA396581.

3. Results

3.1 Incidence and Composition of Endolithic APBs

Our molecular study, which included 30 discrete intertidal carbonate outcrop samples obtained around Isla de Mona, PR, and 11 obtained around Menorca, Spain (Figure 1), showed that APB were abundantly and universally present in this environment. Sequence reads belonging to operational taxonomic units (OTUs) that were classifiable as likely phototrophs (oxygenic or anoxygenic) accounted for a large percentage of the total, on average 21%. As expected from the literature, oxygenic phototrophs (largely cyanobacteria, but also some algae) were either dominant or very well represented (Figure 2). A significant proportion of phototrophic sequence reads could be classified as belonging to one of several groups of APBs, according to our stringent phylogenetic placement criteria (OTU sequence was placed within a clade formed exclusively by known, cultivated phototrophs with >70% confidence; see Materials and Methods). Such APB sequences were found in every sample, varying from 0.8% to 89% of total phototrophs. On average, they accounted for 30% of the total phototroph count. In close to one fifth of the samples, APBs were more abundant than oxygenic phototrophs

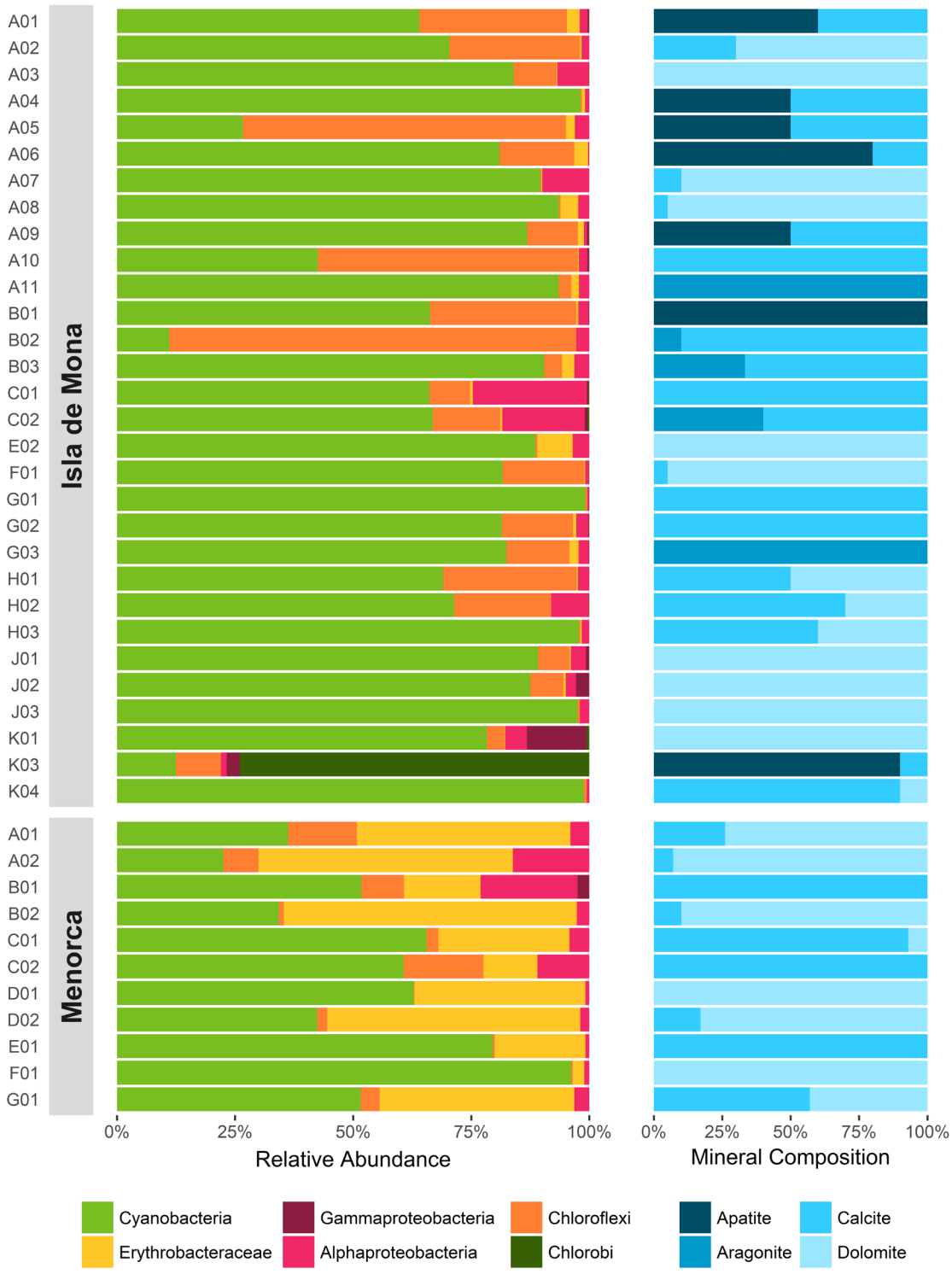

Figure 2. Distribution of 16S rRNA gene reads (left) among phototrophic bacterial groups in endolithic communities from Isla de Mona and Menorca, based on phylogenetic taxon assignment, and the corresponding mineralogical

composition of their substrate based on XRD analyses (right). Sample identifiers are to the very left, their initial capital letters corresponding to the site of origin in Figure 1.

Phylogenetic analyses at family-level resolution of these APBs, revealed that two distinct groups were both common and abundant: the *Chloroflexaceae* (green non-sulfur bacteria) and the *Erythrobacteriaceae* (in the alpha-proteobacteria). Other groups that were well represented included the *Chlorobiaceae*, various purple non-sulfur proteobacteria, and the *Chromatiaceae* in the Gammaproteobacteria. Samples from Isla de Mona supported relatively larger populations of *Chloroflexaceae*, while the endolithic microbiome in Menorca presented a larger relative abundance of *Erythrobacteriaceae*. While most samples were taken in the intertidal, for a few samples in Isla de Mona that were collected in the sub-tidal zone (3.5, 4.6, and 9.1 m deep) we detected a large relative abundance of *Chlorobiaceae* and *Chromatiaceae*, which were absent or very rare in intertidal samples. At this level of phylogenetic resolution, we could not detect any obvious influence of the mineralogical composition of the substrate (which included calcite, aragonite, dolomite and apatite) on the APB community composition (Figure 2). No phototrophic representatives of the betaproteobacteria, the Heliobacteria, the Acidobacteria, or the Gemmatimonadetes were detected.

3.2 Major Endolithic APB Phylotypes

Interestingly, only a few OTUs (defined as groups of sequences that are 97% self-similar) made up the vast majority of all sequences attributable to APBs (Table 1). Two of these OTUs were by far the most abundant: OTU 31154957, a representative of the genus *Roseiflexus* [58], and OTU 582344, a representative of *Erythrobacter* sp. strain NAP1 [59]. Both were present in all samples. Other OTUs in these two families were also quite important. In some samples, we found that OTU 112750, attributable to *Prosthecochloris* in the Chlorobi, were dominant in relative abundance among the APB phylotypes.

Table 1. Major putative APB OTUs within intertidal carbonates.

OTU	Classification[a]	% of Total Phototrophic Sequences	
		Isla de Mona	*Menorca*
3114957	*Roseiflexus* sp.	**12.49**	0.71
NROTU1	*Chlorothrix* sp.	**8.26**	-
582344	*Erythrobacter* sp. NAP1	0.47	**29.05**
NROTU6	*Erythrobacter* sp. NAP1	0.49	**6.48**
112750[b]	*Prosthecochloris* sp.	**2.17**	-

[a]Classification assigned by placement in phylogenetic tree
[b]Present in one sample as 57.57% of the total phototrophic sequences

3.3 Differential Distribution with Geography

The most abundant APB phylotypes throughout Isla de Mona intertidal samples were members of the genera *Roseiflexus*, *Chlorothrix*, and *Chloroflexus* in the phylum Chloroflexi (Figure 3). In stark contrast, the most abundant APBs in Menorca samples were the aerobic APB *Erythrobacter* NAP1. Yet, the latter phylotype could also be found in Isla de Mona, albeit making up a small contribution to the APB guild, and, conversely, the Chloroflexi OTUs that dominated communities in Isla de Mona, also contributed to the communities of Menorca. Additionally, a similar differential incidence could be detected for the purple non-sulfur members of the *Rhodovulum* and *Rubrimonas*, the latter being more abundant in Isla de Mona, and the former in Menorca.

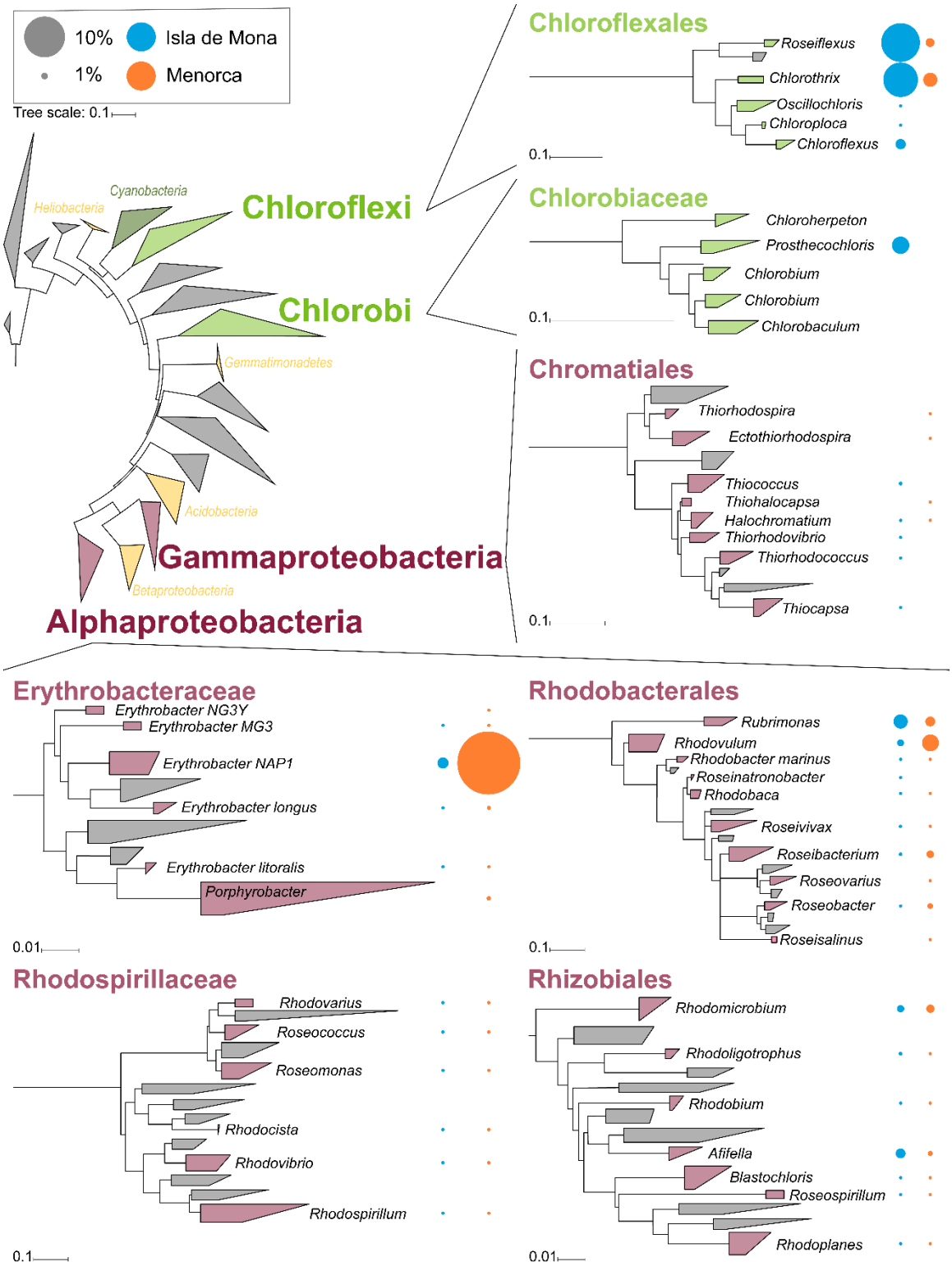

Figure 3. Phylogenetic distribution of endolithic APB phylotypes detected in marine carbonates on Isla de Mona and Menorca. Phyla in the Bacterial

phylogenetic tree (upper left) known to contain phototrophs are shown in color. Detailed sub-trees for each of such phylum that found APB representatives in our survey are shown as enlargements. Circles of variable area to the right of individual clades in these sub-trees represent the average percent of total phototrophic sequences assignable to the clade in Isla de Mona (orange) or Menorca (blue). All trees were constructed using maximum likelihood algorithms.

3.4 Endolithic Morphological Diversity

Confocal microscopy using autofluorescence in the visible and near infrared (NIR) of preparations in which the carbonate substrate had been dissolved away [39] was used in an attempt to visually confirm the presence of APBs, and to gauge morphological and pigmentation diversity. Communities showed a range of morphological diversity. At one end, some samples displayed apparently monospecific beds of filamentous cyanobacteria, with abundant Chl *a* and little NIR fluorescence (Figure 4a). Other samples, however, contained heterogenous mixtures of unicells, clusters, and thin filaments, with varying fluorescence profiles, including cells that were only fluorescent in the NIR (Figure 4b; clearly APBs), and cells that both fluorescent in the NIR and visible, which probably correspond to chlorophyll *d* and *f* containing cyanobacteria.

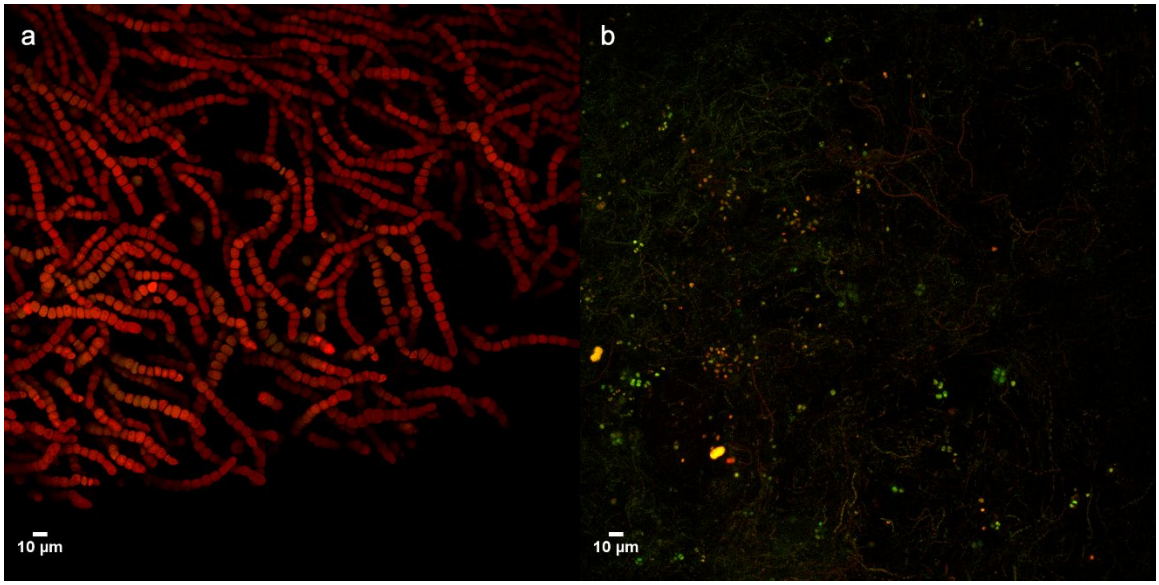

Figure 4. Laser scanning confocal microscopy images of endolith communities after dissolution of their carbonate substrate. Chips were dissolved to liberate biomass from the surrounding mineral, and then imaged with a 405 nm laser. Chl *a* emission (false color red channel) was measured between 660-690 nm and near-infrared emission (false color green channel) was captured in 740- 800 nm range. Images show a wide diversity of morphotypes within and between samples, varying from virtually monotypic beds of red-fluorescing cyanobacteria (**a**; likely *Mastigocoleus*), to diverse assemblages of unicellular, colonial, and filamentous types showing both red and NIR emission (**b**).

3.5 Photosynthetic Pigments in Endolithic Communities

To confirm beyond molecular signatures and confocal microscopy if active anoxygenic phototrophy was occurring within the endolithic microbiome, we analyzed lipid-soluble extracts by HPLC with online diode array spectroscopy detection, to isolate and identify major photopigments. As expected, we found concentrations of Chl *a* within the typical range that had been determined elsewhere in these systems [60,61]. But we also found bacteriochlorophylls (Tables 2 and S1): Bchl *c* (as various isomers, [62]) was detected in every sample, Bchl *a* was detected in eight samples from Isla de Mona and in all samples from Menorca, and Bchl *d* was detected in two samples from Isla de Mona and four samples from Menorca. Bacteriochlorophyll concentrations were generally

much higher in Isla de Mona than in Menorca. The average ratio of Chl *a* to total Bchl on Isla de Mona was 1.9 (Table 2), while on Menorca, the ratio was one order of magnitude higher, at 11.03 (Table 2). Interestingly, chlorophylls *b*, *d*, and *f* were also detected in low quantities in samples from both Isla de Mona and Menorca (Table S1).

Table 2. Detected pigments from intertidal carbonates. Average concentration and range in mg m^{-2}.

	Isla de Mona n = 29			Menorca n = 11			Total n = 40		
Pigment	Samples Detected	Avg. Conc.	Conc. Range	Samples Detected	Avg. Conc.	Conc. Range	Samples Detected	Avg. Conc.	Conc. Range
Chl *a*	29 (100%)	8.95	0.11 – 52.56	11 (100%)	10.14	0.71 – 21.95	40 (100%)	9.28	0.11 – 52.56
Chl *b*	6 (21%)	0.04	0.00 – 0.27	11 (100%)	0.58	0.01 – 1.69	17 (43%)	0.19	0.00 – 1.69
Chl *d*	3 (10%)	0.01	0.00 – 0.16	11 (100%)	0.12	0.01 – 0.56	14 (35%)	0.04	0.00 – 0.56
Chl *f*	3 (10%)	0.01	0.00 – 0.22	9 (82%)	0.19	0.00 – 0.61	12 (30%)	0.06	0.00 – 0.61
BChl *a*	8 (28%)	< 0.01	0.00 - < 0.01	11 (100%)	0.53	0.01 – 1.33	19 (48%)	0.15	0.00 – 1.33
BChl *c*	29 (100%)	8.58	0.14 – 41.66	11 (100%)	1.93	0.02 – 8.19	40 (100%)	6.75	0.02 – 41.66
BChl *d*	2 (7%)	< 0.01	0.00 – 0.12	4 (36%)	0.03	0.00 – 0.14	6 (14%)	0.01	0.00 – 0.14

4. Discussion

The prevalence of bacterial families associated with anoxygenic phototrophy within the endolithic communities became evident upon close examination. Admittedly, however, assigning metabolism to phylotypes based on automated taxonomic assignment, with often poorly curated databases, carries some uncertainties. In order to make our assignments stringent, we constructed our own curated databases and phylogenetic trees (available under http://itol.embl.de/shared/dwroush), and counted as "phototrophs" only OTUs that would fall *within* clades formed *exclusively* by known, cultivated phototrophs with >70% confidence. This strict assignment, in fact, likely led to

an under-estimation of the relative abundance and number of phylotypes of APBs, in that we could have excluded any APBs that were close but not within APB clades, and obviously could not have detected any APBs with no known cultured representatives. However, while likely conservative, our results give us confidence in our finding that APBs are indeed a widespread and significant component of endolithic communities. The fact that this component could have been missed during almost two centuries of research is perplexing. It is possible that the spectral overlap of some chlorophylls and bacteriochlorophylls in extracted form may have disguised these biomarkers [57]. Perhaps the shared morphological characteristics of small thin filamentous cyanobacteria such as *Halomicronema* [63] or *Plectonema terebrans* [15] and *Chloroflexi* [64], rendered them hard to discern under the microscope. And yet, an *ad hoc* literature review returned some corroborating evidence for the presence of APBs in endolithic communities from coral skeletons: spectroscopy revealed absorption peaks in the IR, attributable to the presence of bacteriochlorophylls [65], and Yang et al. [66] report directly the presence of populations of *Prosthecochloris* spp. (Chlorobi).

Bacteriochlorophylls are diagnostic biomarkers for APBs as they are integral to the reaction centers and antenna complexes at the core of their phototrophic capacity [67]. We first examined samples using confocal microscopy, looking for the characteristic profile of near-infrared fluorescence associated with APBs. Though many morphotypes contained varying levels of NIR fluorescence, this evidence was insufficient, in that it could also be attributed to the tail of Chl *d* or Chl *f* fluorescence. Still, some cells were exclusively fluorescent in the NIR, indicating an abundant presence of bacteriochlorophylls. HPLC pigment composition analysis, however, offered a more direct way of identification, showing beyond doubt their presence in all samples. Consistent with the dominance of Chloroflexi in Isla de Mona, Bchl c, a characteristic

photopigment of the Chloroflexi [37] was by far the most abundant bacteriochlorophyll present. Conversely, Bchl *a,* which is characteristic of the Erythrobacteraceae [68] was more abundant in Menorca, consistent with the dominance by *Erythrobacter*. The detection of Bchl *d*, a primary photopigment of *Chlorobi* and some *Chloroflexi* [69,70] was also expected given the abundance of *Chlorobi* in sample K003. However, the concentrations of total Bchl did not correlate well in absolute terms with our molecular tallies, suggesting that we could have missed novel APB populations with our stringent phylogenetic litmus test. Additionally, our differential ability to detect Bchl *a* in each site due to differences in the storage protocols may have also played a role. In any event, these analyses confirmed the presence and breadth of APBs.

While it is clear that the geographical extent of our sampling is insufficient to establish biogeographical patterns of distribution, the switch in intertidal endolith APB dominance between Isla de Mona and Menorca involving *Roseiflexus/Chlorothrix* on the one side and *Erythrobacter* on the other was internally consistent and quite significant. It will be interesting to determine in future studies if the pattern holds in other locations with larger biogeographical provinces, but in the interim, a potential ecophysiological explanation could be put forward. It is known that in terrestrial environments, temperature can in fact drive biogeographic patterns of microbial phototroph distribution [71] and our two sites experience rather different temperature regimes. Isla de Mona has a minimum yearly seawater temperature of 25°C, while in Menorca, winter temperature can dip down to 13°C [72]. A literature review shows that the minimal reported temperature for growth in marine *Chloroflexi* is 18°C [64], whereas it can be as low as 10°C for *Erythrobacter* at [68], and purple non-sulfur alphaproteobacteria can grow at temperatures as low as 5°C [73]. This suggests that temperature may be a significant factor in determining the composition of APBs in intertidal carbonates.

Crucial to establishing the functional impact of APBs on endolithic communities and their geochemical impact on carbonates, is to determine their metabolism *in situ*. Because most of the major APB OTUs in our survey (i.e. in Table 1) are allied with taxa known to act as photoheterotrophs in nature [37,59,74], and because of the absence of an obvious source of electron donors in our samples, we hypothesize that endolithic APBs likely conduct photoheterotrophy as their predominant metabolic function as endoliths, generating ATP through photophosphorylation, and consuming organic compounds including neutral and acidic sugars produced by cyanobacteria [75] as their source of carbon. Considering that diffusion limitation is one of the most important constraints in endolithic habitats [76], photoheterotrophic APBs could add a component of endolithic element cycling, consuming excess sugars, fermentation byproducts and even molecular oxygen [26], along with the release of CO_2 back into the environment. Furthermore, photoheterotrophy has a demonstrable effect on carbonate geochemistry; *Rhodovulum* growing photoheterotrophically on acetate and lactate raised external pH and precipitated carbonate, but it did not do so when grown on neutral sugars [77]. Similar results [78] were obtained with *Rubrivivax* isolates.

Even though cyanobacteria have a mineral substrate preference at the single OTU level [23], we did not detect any such same preference within APBs. This apparent independence of mineral substrate would be consistent with the notion that APBs are not actively carrying out carbonate dissolution, but rather depend on the boring action of cyanobacteria for endolithic space, a hypothesis that will require direct experimentation to formally test.

In summary, we have identified APBs as important endoliths of marine carbonates, with Chloroflexi (*Roseiflexus* and *Chlorothrix*), Erythrobacter (*Erythrobacter sp.* NAP1) and purple non-sulfur alphaproteobacteria as the most

important types. Endolithic APBs could potentially play important metabolic roles in these communities, and, in turn, exert geomicrobial effects on coastal carbonates.

It is of interest to compare the relevance of this new habitat for APBs to existing ones. Our samples had a depth-integrated average biomass of some 7 mg Bchl m^{-2}, which is much less than observed in microbial mats [860 mg Bchl m^{-2} [33]] or lake blooms [some 500 mg m^{-2} [79]], but much more than found in the open ocean [0.1 mg m^{-2} [80]]. When these areal densities are multiplied by the global extent of the respective habitats considered [81,82], it becomes clear that endolithic APB biomass constitutes potentially a significant reservoir, slightly upwards of 10^5 kg of Bchl globally if our survey is representative of most outcrops. This reservoir is much larger than that in microbial mats (some 80 kg Bchl), or in the open ocean (3×10^4 kg Bchl), and similar in magnitude to that of lake blooms (1×10^5 kg Bchl; assuming that as much of 1/10 by surface of all lakes stratify and are sufficiently eutrophic to support these blooms). Considering these simple calculations, the shallow interior of carbonates must be regarded as a globally major reservoir of APBs biomass.

Acknowledgements

This work was funded by National Science Foundation grant EAR 1224939 to F.G.-P., and by a Marie Skłodowska-Curie IOF grant awarded to E.C. We would like to thank the staff in ASU's Goldwater Materials Science Facility for analytical support and A. Garrástazu for field support and medical assistance.

Supplementary Materials

Table S1. Phototrophic pigment distribution and concentration (mg m^{-2}) within intertidal carbonates.

Site	Sample	Chl *a*	Chl *b*	Chl *d*	Chl *f*	BChl *a*	BChl *c*	BChl *d*	Chl:BChl
Isla de Mona	A001	11.63	-	-	-	-	3.42	-	3.40
	A002	11.44	0.20	-	-	-	26.36	-	0.44
	A003	2.58	0.10	-	-	-	3.86	-	0.69
	A004	4.56	0.21	0.01	-	-	7.67	-	0.62
	A005	1.02	-	-	-	-	1.27	-	0.81
	A006	7.78	-	-	-	-	2.98	-	2.61
	A007	1.74	-	-	-	-	1.52	-	1.14
	A008	18.66	-	-	-	-	8.77	-	2.13
	A009	5.02	0.09	-	-	-	3.31	-	1.54
	A010	0.12	-	-	-	-	0.14	-	0.74
	A011	3.33	-	-	-	-	18.28	-	0.18
	B001	3.80	-	-	-	Detected	7.12	-	0.53
	B002	1.07	-	-	-	Detected	5.05	-	0.21
	B003	1.92	-	-	-	-	3.04	-	0.63
	C001	18.67	-	-	-	Detected	12.26	-	1.52
	C002	1.39	-	-	-	-	0.74	-	1.88
	E002	7.98	-	-	-	-	2.89	-	2.76
	F001	28.32	0.16	-	-	-	41.66	-	0.68
	G001	52.56	-	-	-	-	15.62	-	3.36
	G002	19.14	-	-	-	Detected	16.82	-	1.14
	G003	1.89	-	-	0.03	Detected	4.14	-	0.46
	H001	7.98	0.27	0.01	0.05	-	19.64	-	0.42
	H002	8.96	-	-	-	Detected	3.28	-	2.73
	H003	12.88	-	-	-	-	9.52	-	1.35
	J001	9.65	-	0.16	0.22	Detected	14.21	-	0.71
	J002	0.84	-	-	-	-	1.10	-	0.77
	K001	5.95	-	-	-	Detected	0.29	-	20.70
	K003	0.90	-	-	-	-	2.98	-	0.27
	K004	8.01	-	-	-	-	10.83	-	0.74
Menorca	A001	3.25	0.07	0.03	0.09	0.08	0.02	0.10	17.14
	A002	0.71	0.45	0.07	-	0.37	8.17	-	0.14
	B001	12.25	1.69	0.09	0.32	1.10	1.48	0.03	5.49
	B002	17.05	1.07	0.09	0.37	1.07	1.06	-	8.71
	C001	21.90	0.43	0.19	0.32	1.32	2.68	-	5.71
	C002	10.62	0.83	0.03	0.04	0.09	0.78	0.14	11.43
	D001	14.32	0.91	0.04	0.05	0.06	1.16	-	12.49
	D002	17.52	0.59	0.18	0.61	0.80	0.67	-	12.84
	E001	8.88	0.01	0.02	0.01	0.03	0.32	0.06	21.47
	F001	1.93	0.05	0.01	-	0.01	0.06	-	31.86
	G001	2.82	0.27	0.56	0.32	0.87	4.73	-	0.71

References

1. Duerden, J.E. Boring algae as agents in the disintegration of corals. *Bull. Am. Museum Nat. Hist.* **1902**, *889*, 323–332.

2. Kölliker, A. On the frequent occurrence of vegetable parasites in the hard structures of animals. *Proc. R. Soc. London* **1859**, *10*, 95–99, DOI:10.5962/bhl.title.103118.

3. Ramírez-Reinat, E.L.; Garcia-Pichel, F. Characterization of a marine cyanobacterium that bores into carbonates and the redescription of the genus *Mastigocoleus* . *J. Phycol.* **2012**, *48*, 740–749, DOI:10.1111/j.1529-8817.2012.01157.x.

4. Golubić, S.; Le Campion-Alsumard, T. Boring behavior of marine blue-green algae *Mastigocoleus testarum* Lagerheim and *Kyrtuthrix dalmatica* Ercegović, as a taxonomic character. *Aquat. Sci.* **1973**, *35*, 157–161, DOI:10.1007/BF02502070.

5. Carreiro-Silva, M.; Kiene, W.E.; Golubic, S.; McClanahan, T.R. Phosphorus and nitrogen effects on microbial euendolithic communities and their bioerosion rates. *Mar. Pollut. Bull.* **2012**, *64*, 602–613, DOI:10.1016/j.marpolbul.2011.12.013.

6. Chacón, E.; Berrendero, E.; Garcia Pichel, F. Biogeological signatures of microboring cyanobacterial communities in marine carbonates from Cabo Rojo, Puerto Rico. *Sediment. Geol.* **2006**, *185*, 215–228, DOI:10.1016/j.sedgeo.2005.12.014.

7. Ramírez-Reinat, E.L.; Garcia-Pichel, F. Prevalence of Ca2+-ATPase-mediated carbonate dissolution among cyanobacterial euendoliths. *Appl. Environ. Microbiol.* **2012**, *78*, 7–13, DOI:10.1128/AEM.06633-11.

8. Macedo, M.F.; Miller, A.Z.; Dionísio, A.; Saiz-Jimenez, C. Biodiversity of cyanobacteria and green algae on monuments in the Mediterranean Basin: An overview. *Microbiology* **2009**, *155*, 3476–3490, DOI:10.1099/mic.0.032508-0.

9. Godinot, C.; Tribollet, A.; Grover, R.; Ferrier-Pagès, C. Bioerosion by euendoliths decreases in phosphate-enriched skeletons of living corals. *Biogeosciences* **2012**, *9*, 2377–2384, DOI:10.5194/bg-9-2377-2012.

10. Tribollet, A. Dissolution of dead corals by euendolithic microorganisms across the northern Great Barrier Reef (Australia). *Microb. Ecol.* **2008**, *55*, 569–580, DOI:10.1007/s00248-007-9302-6.

11. Reid, R.P.; Foster, J.S.; Radtke, G.; Golubic, S. Modern marine stromatolites of Little Darby Island, exuma archipelago, Bahamas: Environmental setting, accretion mechanisms and role of euendoliths. *Lect. Notes Earth Sci.* **2011**, *131*, 77–89, DOI:10.1007/978-3-642-10415-2.

12. Stjepko Golubic, Imre Friedmann, Ju, S.; Friedmann, E.I.; Schneider, J. The

lithobiontic ecological niche, with special reference to microorganisms. *J Sediment Res* **1981**, DOI:10.1306/212F7CB6-2B24-11D7-8648000102C1865D.

13. Hutchings, P.A. Biological destruction of coral reefs. *Coral Reefs* 1986, *4*, 239–252.

14. Campbell, S.E. The modern distribution and geological history of calcium carbonate boring microorganisms. *Biominer. Biol. Met. Accumul.* **1983**, 99–104, DOI:10.1007/978-94-009-7944-4.

15. Tribollet, A.; Veinott, G.; Golubic, S.; Dart, R. Infestation of the North American freshwater mussel Elliptio complanata (Head Lake, Canada) by the euendolithic cyanobacterium *Plectonema terebrans* Bornet et Flahault. *Arch. Hydrobiol. Suppl. Algol. Stud.* **2008**, *128*, 65–77, DOI:10.1127/1864-1318/2008/0128-0065.

16. Ćurin, M.; Peharda, M.; Calcinai, B.; Golubić, S. Incidence of damaging endolith infestation of the edible mytilid bivalve Modiolus barbatus. *Mar. Biol. Res.* **2014**, *10*, 179–189, DOI:10.1080/17451000.2013.814793.

17. Kaehler, S. Incidence and distribution of phototrophic shell-degrading endoliths of the brown mussel *Perna perna* . *Mar. Biol.* **1999**, *135*, 505–514, DOI:10.1007/s002270050651.

18. Campbell, S.E. Precambrian endoliths discovered. *Nature* **1982**, *299*, 429–431, DOI:10.1038/299429a0.

19. Knoll, A.H.; Golubic, S.; Green, J.; Swett, K. Organically preserved microbial endoliths from the late Proterozoic of East Greenland. *Nature* **1986**, *321*, 856–857, DOI:10.1038/321856a0.

20. Horath, T.; Bachofen, R. Molecular characterization of an endolithic microbial community in dolomite rock in the central Alps (switzerland). *Microb. Ecol.* **2009**, *58*, 290–306, DOI:10.1007/s00248-008-9483-7.

21. Walker, J.J.; Spear, J.R.; Pace, N.R. Geobiology of a microbial endolithic community in the Yellowstone geothermal environment. *Nature* **2005**, *434*, 1011–1014, DOI:10.1038/nature03447.

22. Walker, J.J.; Pace, N.R. Endolithic Microbial Ecosystems. *Annu. Rev. Microbiol.* **2007**, *61*, 331–347, DOI:10.1146/annurev.micro.61.080706.093302.

23. Couradeau, E.; Roush, D.; Guida, B.S.; Garcia-Pichel, F. Diversity and mineral substrate preference in endolithic microbial communities from marine intertidal outcrops (Isla de Mona, Puerto Rico). *Biogeosciences* **2017**, *14*, 311–324, DOI:10.5194/bg-14-311-2017.

24. Belnap, J.; Büdel, B.; Lange, O.L. Biological Soil Crusts: Characteristics and Distribution. In *Biological Soil Crusts: Structure, Function, and Management*; 2003; Vol. 150, pp. 3–30 ISBN 978-3-540-43757-4.

25. Couradeau, E.; Benzerara, K.; Moreira, D.; Gérard, E.; Kaźmierczak, J.; Tavera, R.; López-García, P. Prokaryotic and eukaryotic community structure in field and cultured microbialites from the alkaline Lake Alchichica (Mexico). *PLoS One* **2011**, *6*, DOI:10.1371/journal.pone.0028767.

26. Overmann, J.; Garcia-Pichel, F. The phototrophic way of life. In *The Prokaryotes: Prokaryotic Communities and Ecophysiology*; 2013; Vol. 9783642301, pp. 203–257 ISBN 9783642301230.

27. Koblížek, M. Ecology of aerobic anoxygenic phototrophs in aquatic environments. *FEMS Microbiol. Rev.* **2015**, *39*, 854–870, DOI:10.1093/femsre/fuv032.

28. Bryant, D.A.; Costas, A.M.G.; Maresca, J.A.; Chew, A.G.M.; Klatt, C.G.; Bateson, M.M.; Tallon, L.J.; Hostetler, J.; Nelson, W.C.; Heidelberg, J.F.; et al. *Candidatus* Chloracidobacterium thermophilum: An Aerobic Phototrophic Acidobacterium. *Science (80-.).* **2007**, *317*, 523–526, DOI:10.1126/science.1143236.

29. Zeng, Y.; Feng, F.; Medova, H.; Dean, J.; Kobli ek, M. Functional type 2 photosynthetic reaction centers found in the rare bacterial phylum Gemmatimonadetes. *Proc. Natl. Acad. Sci.* **2014**, *111*, 7795–7800, DOI:10.1073/pnas.1400295111.

30. Imhoff, J.F. The family Chlorobiaceae. In *The Prokaryotes: Other Major Lineages of Bacteria and The Archaea*; **2014**; Vol. 9783642389, pp. 501–514 ISBN 9783642301230.

31. Imhoff, J.F. The family Chromatiaceae. In *The Prokaryotes: Gammaproteobacteria*; **2014**; Vol. 9783642389, pp. 151–178 ISBN 9783642389221.

32. Oren, A. The family ectothiorhodospiraceae. *The Prokaryotes: Gammaproteobacteria* **2014**, *9783642389*, 199–222, DOI:10.1007/978-3-642-38922-1_248.

33. Nicholson, J.A.; Stolz, J.F.; Pierson, B.K. Structure of a microbial mat at Great Sippewisset Marsh, Cape Cod, Masschusetts. *Fems Microbiol. Ecol.* **1987**, *45*, 343–364.

34. Manske, A.K.; Glaeser, J.; Kuypers, M.M.M.; Overmann, J. Physiology and phylogeny of green sulfur bacteria forming a monospecific phototrophic assemblage at a depth of 100 meters in the Black Sea. *Appl. Environ. Microbiol.* **2005**, *71*, 8049–8060, DOI:10.1128/AEM.71.12.8049-8060.2005.

35. Findlay, A.J.; Bennett, A.J.; Hanson, T.E.; Luther, G.W. Light-dependent sulfide oxidation in the anoxic zone of the chesapeake bay can be explained by small populations of phototrophic bacteria. *Appl. Environ. Microbiol.* **2015**, *81*, 7560–7569, DOI:10.1128/AEM.02062-15.

36. Lauro, F.M.; DeMaere, M.Z.; Yau, S.; Brown, M. V; Ng, C.; Wilkins, D.; Raftery, M.J.; Gibson, J.A.; Andrews-Pfannkoch, C.; Lewis, M.; et al. An integrative study

of a meromictic lake ecosystem in Antarctica. *ISME J.* **2011**, *5*, 879–895, DOI:10.1038/ismej.2010.185.

37. Hanada, S. The phylum Chloroflexi, the family Chloroflexaceae, and the related phototrophic families Oscillochloridaceae and Roseiflexaceae. In *The Prokaryotes: Other Major Lineages of Bacteria and The Archaea*; **2014**; Vol. 9783642389, pp. 515–532 ISBN 9783642301230.

38. Madigan, M.T.; Jung, D.O. An Overview of Purple Bacteria: Systematics, Physiology, and Habitats. *Purple Phototrophic Bact.* **2009**, *28*, 1–15, DOI:10.1007/978-1-4020-8815-5.

39. Wade, B.; Garcia-Pichel, F. Evaluation of DNA Extraction Methods for Molecular Analyses of Microbial Communities in Modern Calcareous Microbialites. *Geomicrobiol. J.* **2003**, *20*, 549–561, DOI:10.1080/01490450390249460.

40. Schindelin, J.; Arganda-Carreras, I.; Frise, E.; Kaynig, V.; Longair, M.; Pietzsch, T.; Preibisch, S.; Rueden, C.; Saalfeld, S.; Schmid, B.; et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **2012**, *9*, 676–682, DOI:10.1038/nmeth.2019.

41. Muyzer, G.; De Waal, E.; Uitterlinden, A. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain Reaction-Amplified Genes Coding for 16S rRNA. *Appl. Enviromental Microbiol.* **1993**, *59*, 695–700.

42. Caporaso, J.G.; Lauber, C.L.; Walters, W.A.; Berg-Lyons, D.; Lozupone, C.A.; Turnbaugh, P.J.; Fierer, N.; Knight, R. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci.* **2011**, *108*, 4516–4522, DOI:10.1073/pnas.1000080107.

43. Caporaso, J.G.; Kuczynski, J.; Stombaugh, J.; Bittinger, K.; Bushman, F.D.; Costello, E.K.; Fierer, N.; Peña, A.G.; Goodrich, J.K.; Gordon, J.I.; et al. QIIME allows analysis of high- throughput community sequencing data. *Nat. Publ. Gr.* **2010**, *7*, 335–336, DOI:10.1038/nmeth0510-335.

44. Rognes, T.; Flouri, T.; Nichols, B.; Quince, C.; Mahé, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **2016**, *4*, e2584, DOI:10.7717/peerj.2584.

45. Kopylova, E.; Noé, L.; Touzet, H. SortMeRNA: Fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **2012**, *28*, 3211–3217, DOI:10.1093/bioinformatics/bts611.

46. Mercier, C.; Boyer, F.; Bonin, A.; Coissac, E. SUMATRA and SUMACLUST: fast and exact comparison and clustering of sequences Available online: http://metabarcoding.org/sumatra.

47. DeSantis, T.Z.; Hugenholtz, P.; Larsen, N.; Rojas, M.; Brodie, E.L.; Keller, K.; Huber, T.; Dalevi, D.; Hu, P.; Andersen, G.L. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **2006**, *72*, 5069–5072, DOI:10.1128/AEM.03006-05.

48. Quast, C.; Pruesse, E.; Yilmaz, P.; Gerken, J.; Schweer, T.; Yarza, P.; Peplies, J.; Glöckner, F.O. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **2013**, *41*, 590–596, DOI:10.1093/nar/gks1219.

49. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780, DOI:10.1093/molbev/mst010.

50. Sela, I.; Ashkenazy, H.; Katoh, K.; Pupko, T. GUIDANCE2: Accurate detection of unreliable alignment regions accounting for the uncertainty of multiple parameters. *Nucleic Acids Res.* **2015**, *43*, W7–W14, DOI:10.1093/nar/gkv318.

51. Miller, M.A.; Pfeiffer, W.; Schwartz, T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *2010 Gatew. Comput. Environ. Work. GCE 2010* **2010**, DOI:10.1109/GCE.2010.5676129.

52. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313, DOI:10.1093/bioinformatics/btu033.

53. Berger, S.A.; Stamatakis, A. Aligning short reads to reference alignments and trees. *Bioinformatics* **2011**, *27*, 2068–2075, DOI:10.1093/bioinformatics/btr320.

54. Berger, S.A.; Krompass, D.; Stamatakis, A. Performance, accuracy, and web server for evolutionary placement of short sequence reads under maximum likelihood. *Syst. Biol.* **2011**, *60*, 291–302, DOI:10.1093/sysbio/syr010.

55. Letunic, I.; Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **2016**, *44*, W242–W245, DOI:10.1093/nar/gkw290.

56. Frigaard, N.U.; Takaichi, S.; Hirota, M.; Shimada, K.; Matsuura, K. Quinones in chlorosomes of green sulfur bacteria and their role in the redox-dependent fluorescence studied in chlorosome-like bacteriochlorophyll *c* aggregates. *Arch. Microbiol.* **1997**, *167*, 343–349, DOI:10.1007/s002030050453.

57. Frigaard, N.U.; Larsen, K.L.; Cox, R.P. Spectrochromatography of photosynthetic pigments as a fingerprinting technique for microbial phototrophs. *FEMS Microbiol. Ecol.* **1996**, *20*, 69–77, DOI:10.1016/0168-6496(96)00005-0.

58. Hanada, S.; Takaichi, S.; Matsuura, K.; Nakamura, K. *Roseiflexus castenholzii* gen. nov., sp. nov., a thermophilic, filamentous, photosynthetic bacterium that lacks chlorosomes. *Int. J. Syst. Evol. Microbiol.* **2002**, *52*, 187–193, DOI:10.1099/00207713-52-1-187.

59. Koblízek, M.; Janouškovec, J.; Oborník, M.; Johnson, J.H.; Ferriera, S.; Falkowski, P.G. Genome sequence of the marine photoheterotrophic bacterium *Erythrobacter* sp. Strain NAP1. *J. Bacteriol.* **2011**, *193*, 5881–5882, DOI:10.1128/JB.05845-11.

60. Tribollet, A.; Langdon, C.; Golubic, S.; Atkinson, M. Endolithic microflora are major primary producers in dead carbonate substrates of Hawaiian coral reefs. *J. Phycol.* **2006**, *42*, 292–303, DOI:10.1111/j.1529-8817.2006.00198.x.

61. Raghukumar, C.; Sharma, S.; Lande, V. Distribution and biomass estimation of shell-boring algae in the intertidal at Goa, India. *Phycologia* **1991**, *30*, 303–309, DOI:10.2216/i0031-8884-30-3-303.1.

62. Caple, M.B.; Chow, H.; Strouse, C.E. Photosynthetic pigments of green sulfur bacteria. The esterifying alcohols of bacteriochlorophylls *c* from *Chlorobium limicola* . *J. Biol. Chem.* **1978**, *253*, 6730–6737.

63. Abed, R.M.M.; Garcia-Pichel, F.; Hernández-Mariné, M. Polyphasic characterization of benthic, moderately halophilic, moderately thermophilic cyanobacteria with very thin trichomes and the proposal of *Halomicronema excentricum* gen. nov., sp. nov. *Arch. Microbiol.* **2002**, *177*, 361–370, DOI:10.1007/s00203-001-0390-2.

64. Pierson, B.K.; Valdez, D.; Larsen, M.; Morgan, E.; Mack, E.E. Chloroflexus-like organisms from marine and hypersaline environments: Distribution and diversity. *Photosynth. Res.* **1994**, *41*, 35–52, DOI:10.1007/BF02184144.

65. Magnusson, S.H.; Fine, M.; Kühl, M. Light microclimate of endolithic phototrophs in the scleractinian corals *Montipora monasteriata* and *Porites cylindrica* . *Mar. Ecol. Prog. Ser.* **2007**, *332*, 119–128, DOI:10.3354/meps332119.

66. Yang, S.H.; Lee, S.T.M.; Huang, C.R.; Tseng, C.H.; Chiang, P.W.; Chen, C.P.; Chen, H.J.; Tang, S.L. Prevalence of potential nitrogen-fixing, green sulfur bacteria in the skeleton of reef-building coral *Isopora palifera* . *Limnol. Oceanogr.* **2016**, *61*, 1078–1086, DOI:10.1002/lno.10277.

67. Blankenship, R.E. Reaction centers and electron transport pathways in anoxygenic phototrophs. In *Molecular Mechanisms of Photosynthesis*; 2008; pp. 89–110 ISBN 9780470758472.

68. Tonon, L.A.C.; Moreira, A.P.B.; Thompson, F. The family Erythrobacteraceae. In *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria*; 2014; Vol. 9783642301, pp. 213–235 ISBN 9783642301971.

69. Dubinina, G.A.; Gorlenko, V.M. [New filamentous photosynthesizing green bacteria with gas vacuoles]. *Mikrobiologiia* **1975**, *44*, 511–517.

70. Gorlenko, V.M. A new phototrophic green sulphur bacterium–*Prosthecochloris aestuarii* nov. gen. nov. spec. *Z. Allg. Mikrobiol.* **1970**, *10*, 147–149, DOI:10.1002/jobm.19700100207.

71. Garcia-Pichel, F.; Loza, V.; Marusenko, Y.; Mateo, P.; Potrafka, R.M. Temperature Drives the Continental-Scale Distribution of Key Microbes in Topsoil Communities. *Science (80-.).* **2013**, *340*, 1574–1577, DOI:10.1126/science.1236404.

72. Reynolds, R.W.; Smith, T.M.; Liu, C.; Chelton, D.B.; Casey, K.S.; Schlax, M.G. Daily high-resolution-blended analyses for sea surface temperature. *J. Clim.* **2007**, *20*, 5473–5496, DOI:10.1175/2007JCLI1824.1.

73. Karr, E.A.; Sattley, W.M.; Jung, D.O.; Madigan, M.T.; Achenbach, L.A. Remarkable diversity of phototrophic purple bacteria in a permanently frozen Antarctic lake. *Appl. Environ. Microbiol.* **2003**, *69*, 4910–4914, DOI:10.1128/AEM.69.8.4910-4914.2003.

74. Madigan, M.T.; Jung, D.O.; Karr, E. a; Sattley, W.M.; Achenbach, L. a.; van der Meer, M.T.J. Diversity of Anoxygenic Phototrophs in Contrasting Extreme Environments. *Geotherm. Biol. Geochemistry YNP* **2005**, *1*, 203–220.

75. De Philippis, R.; Vincenzini, M. Exocellular polysccharides from cyanobacteria and their possible applications. *FEMS Microbiol. Rev.* **1998**, *22*, 151–175, DOI:10.1111/j.1574-6976.1998.tb00365.x.

76. Palmer, R.J.; Friedmann, E.I. Water relations and photosynthesis in the cryptoendolithic microbial habitat of hot and cold deserts. *Microb. Ecol.* **1990**, *19*, 111–118, DOI:10.1007/BF02015057.

77. Bundeleva, I.A.; Shirokova, L.S.; Bénézeth, P.; Pokrovsky, O.S.; Kompantseva, E.I.; Balor, S. Calcium carbonate precipitation by anoxygenic phototrophic bacteria. *Chem. Geol.* **2012**, *291*, 116–131, DOI:10.1016/j.chemgeo.2011.10.003.

78. Li, R.Y.; Fang, H.H.P. Hydrogen production characteristics of photoheterotrophic *Rubrivivax gelatinosus* L31. *Int. J. Hydrogen Energy* **2008**, *33*, 974–980, DOI:10.1016/j.ijhydene.2007.12.001.

79. Gich, F.; Garcia-Gil, J.; Overmann, J. Previously unknown and phylogenetically diverse members of the green nonsulfur bacteria are indigenous to freshwater lakes. *Arch. Microbiol.* **2002**, *177*, 1–10, DOI:10.1007/s00203-001-0354-6.

80. Lami, R.; Cottrell, M.T.; Ras, J.; Ulloa, O.; Obernosterer, I.; Claustre, H.; Kirchman, D.L.; Lebaron, P. High abundances of aerobic anoxygenic photosynthetic bacteria in the South Pacific Ocean. *Appl. Environ. Microbiol.* **2007**, *73*, 4198–4205, DOI:10.1128/AEM.02652-06.

81. Ford, D.; Williams, P. The Global Distribution Of Karst. In *Karst Hydrogeology and Geomorphology*; **2013**; pp. 1–8 ISBN 9781118684986.

82. Garcia-Pichel, F.; Belnap, J.; Neuer, S.; Schanz, F. Estimates of global cyanobacterial biomass and its distribution. *Arch. Hydrobiol. Suppl. Algol. Stud.* **2003**, *109*, 213–227, DOI:10.1127/1864-1318/2003/0109-0213.

CHAPTER 3

SUCCESSION AND COLONIZATION DYNAMICS OF ENDOLITHIC PHOTOTROPHS WITHIN INTERTIDAL CARBONATES

Published in *microoganisms*, **2020**

Coauthors have acknowledged the use of this manuscript in my dissertation

Authors:

Daniel Roush and Ferran Garcia-Pichel

Abstract: Photosynthetic endolithic communities are common in shallow marine carbonates, contributing significantly to their bioerosion. Cyanobacteria are well known from these settings, where a few are euendoliths, actively boring into the virgin substrate. Recently, anoxygenic phototrophs were reported as significant inhabitants of endolithic communities, but it is unknown if they are euendoliths or simply colonize available pore spaces secondarily. To answer this and to establish the dynamics of colonization, nonporous travertine tiles were anchored onto intertidal beach rock in Isla de Mona, Puerto Rico, and developing endolithic communities were examined with time, both molecularly and with photopigment biomarkers. By 9 months, while cyanobacterial biomass and diversity reached levels indistinguishable from those of nearby climax communities, anoxygenic phototrophs remained marginal, suggesting that they are secondary colonizers. Early in the colonization, a novel group of cyanobacteria (unknown boring cluster, UBC) without cultivated representatives, emerged as the most common euendolith, but by 6 months, canonical euendoliths such as *Plectonema* (*Leptolyngbya*)

sp., *Mastigocoleus* sp., and Pleurocapsalean clades displaced UBC in dominance. Later, the proportion of euendolithic cyanobacterial biomass decreased, as nonboring endoliths outcompeted pioneers within the already excavated substrate. Our findings demonstrate that endolithic cyanobacterial succession within hard carbonates is complex but can attain maturity within a year's time.

---

1. Introduction

The endolithic microbiome of intertidal carbonate rocks has been the subject of intensive study since the 1800s [1,2], with a main focus on the characterization of bioerosive agents within these communities. The agents, boring organisms referred to as euendoliths, excavate the rock substrate and create pore spaces for their own growth. Le Campion-Alsumard and colleagues [3,4] first examined succession and colonization by microscopic inspection in order to better understand the ecological principles that drive euendolith community formation. As concern for coral destruction rose in the 1990s, others [5–11] applied the same procedures to understand these dynamics and mitigate bioerosion in reef ecosystems. These studies on porous, biogenic carbonates from coral skeletons showed swift initial colonization by euendolithic algae, with successional changes occurring within months and communities reaching maturity after a year. Kiene [10], Gektidis [9] and Chacón et al. [12] examined hard mineral carbonates as well, finding that euendolithic cyanobacteria, not algae, were the dominant boring organisms there and that hard substrates led to more diverse cyanobacterial populations than those of corals.

Although early research was informative in identifying and characterizing major euendolithic players, the use of morphological characterization alone has been found to underestimate microbial diversity in endolithic cyanobacterial communities [12,13].

Indeed, in the case of marine carbonate communities, high-throughput amplicon sequencing has demonstrated that morphology-based studies can underrepresent cyanobacterial diversity estimates by factors of 10 to 100 [12,14,15]. Early research identified three major morphotypical groups of euendolithic cyanobacteria. One is represented by the thin, filamentous, *Leptolyngbya*-like organisms most commonly assigned to *Plectonema terebrans* (or *Leptolyngbya terebrans),* which are typically one of the most abundant euendolith morphotypes, at times exceeding 80% of total euendolithic biovolume [6]. Unfortunately, no 16S rRNA gene sequence of *P. terebrans* has been obtained from cultures, making it impossible to identify it with certainty in molecular surveys. Environmental sequences best matching *Halomicronema* and *Leptolyngbya* species have been tentatively suggested to represent the elusive *P. terebrans* [14]. The second group corresponds to the species *Mastigocoleus testarum,* which is characterized by a complex, true-branching filamentous morphology, making it easily identifiable from microscopic examination. It has been recently redescribed on the basis of a polyphasic approach based on strain BC008, showing congruency between molecular and traditional approaches [16], has served as a model to elucidate the physiological mechanism of boring [16–18], and is the first euendolith whose genome has been fully sequenced [19]. *Mastigocoleus testarum* is one of the earliest colonizers in soft carbonates, being found as early as one week after initial exposure [3,11,20]. A third, diverse group includes several members of the order Pleurocapsales in the genera *Hyella, Solentia*, *Hormathonema*, and the recently described *Candidatus* Pleuronema. Members of the Pleurocapsales typically act as pioneer borers but can bore only to shallow depths and are easily preyed upon by grazers, leading to low abundance in mature communities [6,7,11,20].

Through comprehensive, high-throughput molecular surveys, we recently found a diverse phototrophic community in the endolithic habitat of coastal hard carbonates, which included four distinct anoxygenic phototrophic bacterial (APB) groups. The most dominant APBs were members of the Chloroflexales (green nonsulfur bacteria) [21–23] and *Erythrobacter* (aerobic anoxygenic phototrophs) [24,25]. APBs could comprise upwards of 80% of the total phototroph community [15] in some samples. Our findings broadened the known habitats for APBs and suggested that some microscopic characterizations of endolithic thin filamentous organisms (*Plectonema*-like) may have in fact been APBs. Thus, APBs could be euendolithic in nature, potentially upending the long established understanding of endolith ecology by broadening the pool of possible pioneer organisms and boring mechanisms.

Therefore, to provide new molecular insights into euendolith colonization and succession and to attempt to answer questions that arose from our prior work, we set up a colonization experiment in the intertidal zone of Playa Ulvero, Isla de Mona, Puerto Rico. We anchored nonporous travertine (a dense, compact form of calcium carbonate) tiles onto beach rock 5 m from shore and collected samples every 3 months over a 9-month period. Our study had four specific aims: (1) to elucidate APB colonization timing to identify if APBs are pioneer organisms with the ability to bore; (2) to examine cyanobacterial euendolith colonization and succession using molecular methods; (3) to measure the colonization dynamics of the *Leptolyngbya*-like (*Plectonema*), *Mastigocoleus*-like, and Pleurocapsalean euendolithic cyanobacterial groups; and (4) to compare colonization progress to previously described steady-state climax communities of similar geological composition and geographic location in order to gauge community maturity.

2. Materials and Methods

2.1. Tile Placement and Sample Collection

Commercial, 4 inches wide, 1.5 inches thick travertine square tiles, were anchored onto intertidal beach rock some 5 m from the high-tide shoreline at Playa Uvero on Isla de Mona, Puerto Rico, (18°03′36.2″N 67°54′21.8″W) (Figure 1) after having received permits from the Departamento de Recursos Naturales y Ambientales (Commonwealth of Puerto Rico). Tiles were fastened to the beach rock using a combination of Red Head 5” × 3/8” 316 Stainless Steel Wedge Anchors and JB Weld Waterweld putty. Three tiles were sacrificially collected every three months, air-dried and shipped, reaching the laboratory in less than a week, and then stored on arrival at −80°C until analysis.

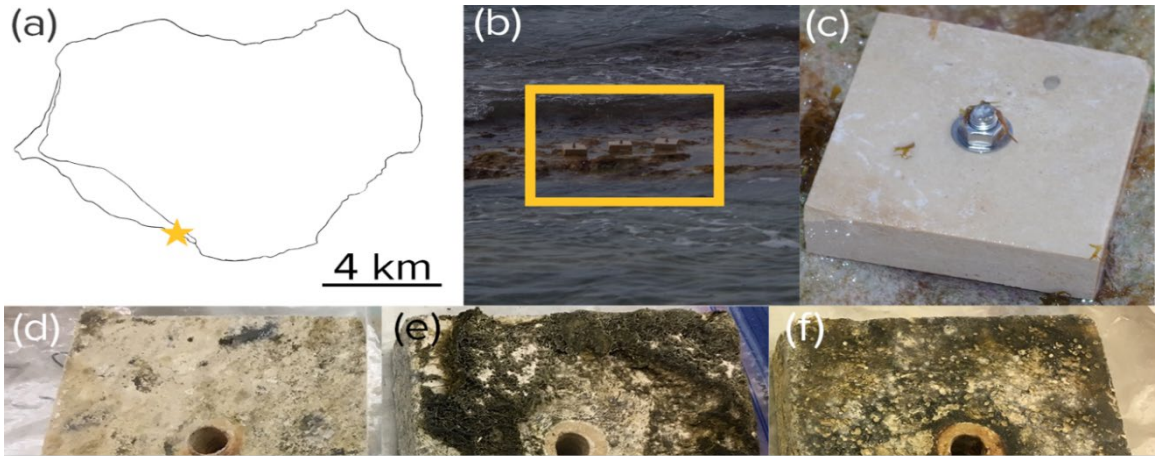

Figure 1. Experimental tile placement. (**a**) Location near Playa Uvero (yellow star) on Isla de Mona, Puerto Rico. (**b**) Anchoring on a stretch of intertidal beach rock (yellow box) as seen at low tide. (**c–f**) Aspect of virgin (**c**) and exposed tiles harvested after harvested after 3 (**d**), 6 (**e**), and 9 months (**f**). Part of the growth observable in the pictures was epilithic in nature.

2.2. Endolithic Community DNA Extraction

Tiles were vigorously brushed with sterile toothbrushes and sterilized seawater to remove epilithic biomass. To ensure a consistent sampling effort, each tile was sampled four times in 2 cm by 2 cm squares, 1 cm from the edge of the tile (sampling is shown in Figure 2d–f). Sampled fragments were ground in sterile mortars following the protocol described in Wade and Garcia-Pichel 2003 [26], and 0.5 g of powered rock was used as the input material for a MoBio PowerPlant Pro DNA extraction kit (Mo Bio Laboratories, Inc., Carlsbad, CA, USA) following the protocol provided, except that, before the first lysis step, the contents of the bead tubes were homogenized horizontally at 2200 rpm for 10 minutes, and, additionally, subjected to seven freeze–thaw cycles using liquid nitrogen to ensure full disruption of bacterial membranes.

2.3. Quantitative PCR of 16S rRNA Gene Content

In order to quantify the number of 16S rRNA gene copies in the extracts, quantitative real-time PCR was conducted using universal V3 16S rRNA gene primers 338F (5'- ACTCCTACGGGAGGCAGCAG-3') and 518R (5'-GTATTACCG CGGCTGCTGG-3'). PCRs were performed in triplicate using Sso Fast mix (Bio-Rad, Hercules, CA, USA) following Couradeau et al. [27]. Following quantification, triplicate 16S rRNA gene counts were averaged and then converted to counts per square meter using the surface area of the tile analyzed. The total counts per square meter were then multiplied by the associated proportional abundance of any clade of interest in order to obtain absolute population size for that clade. Separate biological replicates (i.e., tiles) were then averaged.

2.4. 16S rRNA Gene Library Preparation and Illumina Sequencing

Amplicon sequencing of the V3–V4 variable region of the 16S rRNA gene was performed using the universal bacterial PCR primers 341F (5'-CCTACGGGNGGCWGCAG) [28] and 806R (5'-GGACTACVSGGGTATCTAAT) [29]. PCR amplifications were done in triplicate, then pooled and quantified using Quant-iT™ PicoGreen® dsDNA Assay Kit (Invitrogen). Two hundred forty nanograms of DNA per sample were pooled and then cleaned using QIA quick PCR purification kit (QIAGEN). The PCR pool was quantified by Illumina library Quantification Kit ABI Prism® (Kapa Biosystems). DNA pool was determined and diluted to a final concentration of 4 nM then denatured diluted to a final concentration of 4 pM with a 30% of PhiX. Finally, the DNA library was loaded in the MiSeq Illumina sequencer using the chemistry version 3 (2 × 300 paired-end) and following the guidelines of the manufacturer.

2.5. Data Analysis Pipeline

Raw sequences were processed using the QIIME2 2018.2 analysis pipeline [30]. Demultiplexed sequences were imported into QIIME2 and processed using the DADA2 [31] denoised-paired plugin with the following parameters: trunc_len_f:280, trunc_len_r:235, trim_left_f:20, trim_left_r:25, and max_ee:8, so as to obtain amplicon sequence variants (ASVs). After resolving ASVs, any sequences found in the control tile extracts (uncolonized tiles) were filtered from the final feature table. Sequencing depth of the experimental tiles ranged from 21,897 to 180,168 (post filtering), and alpha-rarefaction analysis indicated that all samples had reached convergence (Figure S1). In order to conduct diversity analysis, representative sequences were aligned using MAFFT7 [32], and a phylogenetic tree was generated using FastTree

[33]. Diversity metrics were calculated using the core-metrics-phylogenetic plugin, including Weighted and Unweighted UniFrac metrics [34]. ASVs were initially classified using the classify-sklearn plugin, (Available online: https://github.com/qiime2/q2-feature-classifier) with a Green Genes 13_8 [35] based classifier. The feature table was then exported, and differential abundance analysis was conducted using the QIIME1 [36] plugin *differential_abundance.py* and the DESeq2 algorithm [37]. PCoAs were generated using the vegan package [38], and graphics were created using R [39] and the ggplot2 package [40]. Statistical analyses were conducted either using R (Student's *t*-test) or within Qiime2 (Kruskal–Wallis, PERMANOVA).

2.6. Cyanobacterial ASV Classification

To identify key euendolithic cyanobacterial clades, the representative sequence output from QIIME2 was filtered to only include cyanobacterial sequences (plastids were removed). These comprised at least 95% of the total number of reads within each sample. Next, the sequences were aligned to the Cydrasil reference alignment [41] using PaPaRa [42], and placed into the Cydrasil reference tree using the Evolutionary Placement Algorithm (based on the maximum-likelihood model) feature of RAxML8 [43]. The output was visualized using the ITOL3 website [44]. An ASV was considered a likely euendolith if it was placed on a branch containing only known euendolithic cyanobacteria with a >70% certainty. Biomass was calculated for each tile by multiplying the total relative abundance of the cluster by the total areal concentration of 16S rRNA genes in that sample. Then, biological replicates for each time point were averaged and graphed using R and ggplot2.

2.7. Steady-State Climax Community Comparisons

Three natural substrate samples from Couradeau et al. [14] and Roush et al. [15] (samples denoted as H001-H003 in SRA) were used as proxies for steady-state climax communities for comparison of colonization progress. The samples were chosen based upon their geographic proximity to the tile placement location and their similar geological composition (calcite). The raw sequencing data was processed using the same parameters and pipeline as described above. Pigment analysis was conducted in the same manner as for the tiles.

2.8. Unknown Boring Cluster (UBC) Phylogenetic Tree

In order to assess the nearest neighbors of the unknown boring cluster, a multiple sequence alignment (MSA) was generated using SSU-Align [45]. The MSA was comprised of the three most differentially abundant ASVs identified from DESeq2 and EPA placement analysis, the nearest sequences from Cydrasil, and the top seven most similar NCBI nr database sequences identified using BLAST [46]. The resulting alignment of 398 sequences was then used as input into RAxML8 [47] to generate a phylogenetic tree using the rapid-bootstrap algorithm with 1000 bootstraps and the GTR GAMMA model. The remaining ASVs were then checked using BLAST and the nr database for proximity to the resulting clade.

2.9. Pigment Extraction and Analysis

In order to extract lipid-soluble pigments, the remaining powdered sample (the same samples used for DNA extraction) was suspended in 7:2 acetone:methanol solvent and sonicated twice for 30 s in an ice bath in the dark. Extracts were centrifuged at 2100 × *g* for 10 min, decanted, and the supernatant filtered through a 0.22 µm nylon filter.

These steps were repeated and the supernatants were pooled until the extract was devoid of color. The resulting extract was then evaporated under a N_2 stream in the dark and resuspended in 200 µL of HPLC-grade acetone. HPLC analysis was conducted on a Waters Alliance e2695 HPLC with an inline Waters 2998 photodiode array detector, using a protocol adapted from Frigaard et al. [48] for use on a CORTECS C18 4.6 mm × 150 mm (90 Å pore size, 2.7 µm particles) column. Separation was performed as follows: the initial solvent gradient composed of ethyl-acetate:methanol:acetonitrile:water in a 21:23.9:47.6:7.5 ratio by volume and linearly changed to 30:20:50:0 ratio by volume in 13.43 minutes, held for 3.87 minutes, and then immediately returned to the initial ratio (21:23.9:47.6:7.5 by volume) and held for 5.7 minutes. Total runtime per sample was 23 minutes, at a flow rate of 2 mL min^{-1} and column temperature of 30 °C. Pigment identification was done by comparison of retention time and spectrum against standards of Chl *a* and BChl *a* obtained from Sigma Aldrich. All other pigments were identified from known spectra [49]. Injected pigment mass was calculated from the chromatogram using the equation $m = FA\,(e_m\,d)^{-1}$, where m is the mass of BChl or Chl in milligrams, *F* is the solvent flow rate (1 mL min^{-1}), A is the peak area (in Au), e_m is the extinction coefficient in L mg^{-1} cm^{-1}, and d is the path length of the PDA detector (1 cm). Extinction coefficients were taken from Ley et al. 2006 [50].

2.10. Data Availability

Isla de Mona steady-state climax community raw sequencing data is deposited under NCBI BioProject PRJNA603780. Raw sequencing data from the experimental tiles is deposited under NCBI BioProject PRJNA596277.

3. Results

3.1. Endolithic Bacterial and Phototrophic Growth

Visual inspection of the colonized tiles showed a marked increase in both pigmentation and erosion with time (Figure 2c–f). The tiles sustained both nonphototrophic and phototrophic bacterial growth over the 9-month exposure period. Bacterial biomass increased at an average rate of 3×10^{10} 16S rRNA gene copies per m^{-2} $month^{-1}$, reaching a mean value of 1.1×10^{11} 16S rRNA gene copies per m^{-2} at 9 months (Figure 3a). As expected, phototroph colonization followed a similar trend with photopigment content increasing at a rate of 2.3 mg m^{-2} $month^{-1}$, reaching an average value of 7.22 mg m^{-2} by 9 months. (Figure 3b), at which point 16S rRNA gene counts were not significantly different from those found in steady-state climax communities described by Couradeau et al. [51] and Roush et al. [15] (Student's *t*-test, $p < 0.05$). While total chlorophyll pigment concentrations were not significantly different between 9 months and steady-state climax communities either, cyanobacteria-specific counts were actually higher at 9 months than in steady-state climax communities.

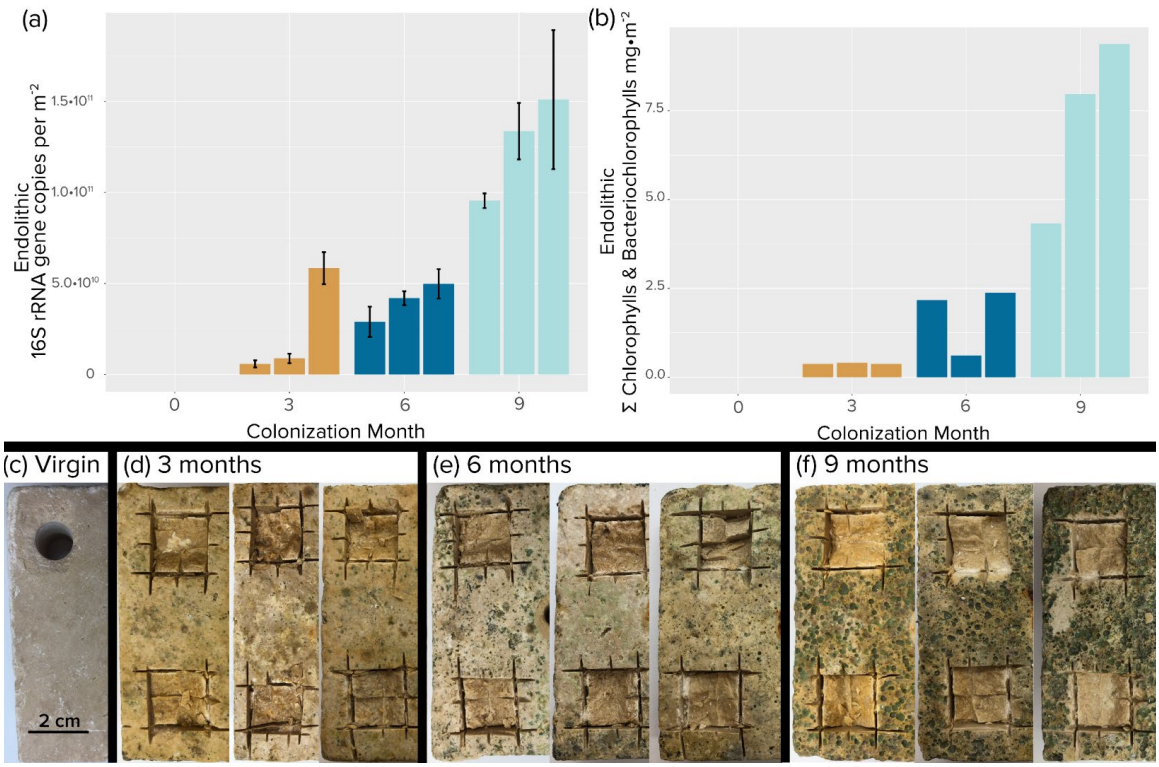

Figure 2. Endolithic colonization of travertine tiles. (**a**) Areal concentration of 16S rRNA gene copies. Each bar is an independent replicate. Error bars are from biological replicates. (**b**) Areal concentration of total photosynthetic chlorins (chlorophylls plus bacteriochlorophylls). Single determinations were carried out for each replicate tile. (**c–f**) Photographic evidence of colonization after removal of epilithic biomass by brushing. (**c**) Initial, virgin tile. Excising squares were samples used for analyses.

3.2. Incidence of Anoxygenic Phototrophs

APB abundance measured by bacteriochlorophylls increased with time but trailed in concentration by some two orders of magnitude to cyanobacterial abundance measured by chlorophylls during the colonization period. This situation obviously changed significantly later during succession, as bacteriochlorophylls were statistically as abundant as chlorophylls when compared to steady-state climax communities (Figure 3b). The magnitude of the difference between APB and cyanobacteria was less marked,

but still very significant, when measured by 16S rRNA gene abundance (Figure 3a). By using this metric, it was obvious that although APB trailed cyanobacteria during the colonization period, they eventually matched and even exceeded cyanobacteria in steady-state climax communities. We found very differing dynamics between populations of relevant APB groups: while Chloroflexales were only present in very small quantities during early phases (Figure 3c) and reached only 6.2×10^6 16s rRNA gene copies per m^{-2} after 9 months, *Erythrobacter* abundance was stable throughout the colonization, with an average of 1.2×10^9 16S rRNA gene copies per m^{-2} at 9 months. In comparison, the situation was reversed in steady-state mature communities, where *Erythrobacter* sp. decreased to some 1.6×10^8 16s rRNA gene copies per m^{-2} in steady-state climax communities, but Chloroflexales increased to populations close to those of cyanobacteria (Figure 3c). The apparent differences in trends between bacteriochlorophyll and 16S rRNA genes as proxies for population size can be explained by the relatively low bacteriochlorophyll content of *Erythrobacter spp.* compared to members of the Chloroflexales [52,53], which essentially made the total content of bacteriochlorophyll be very sensitive to the population size of the latter.

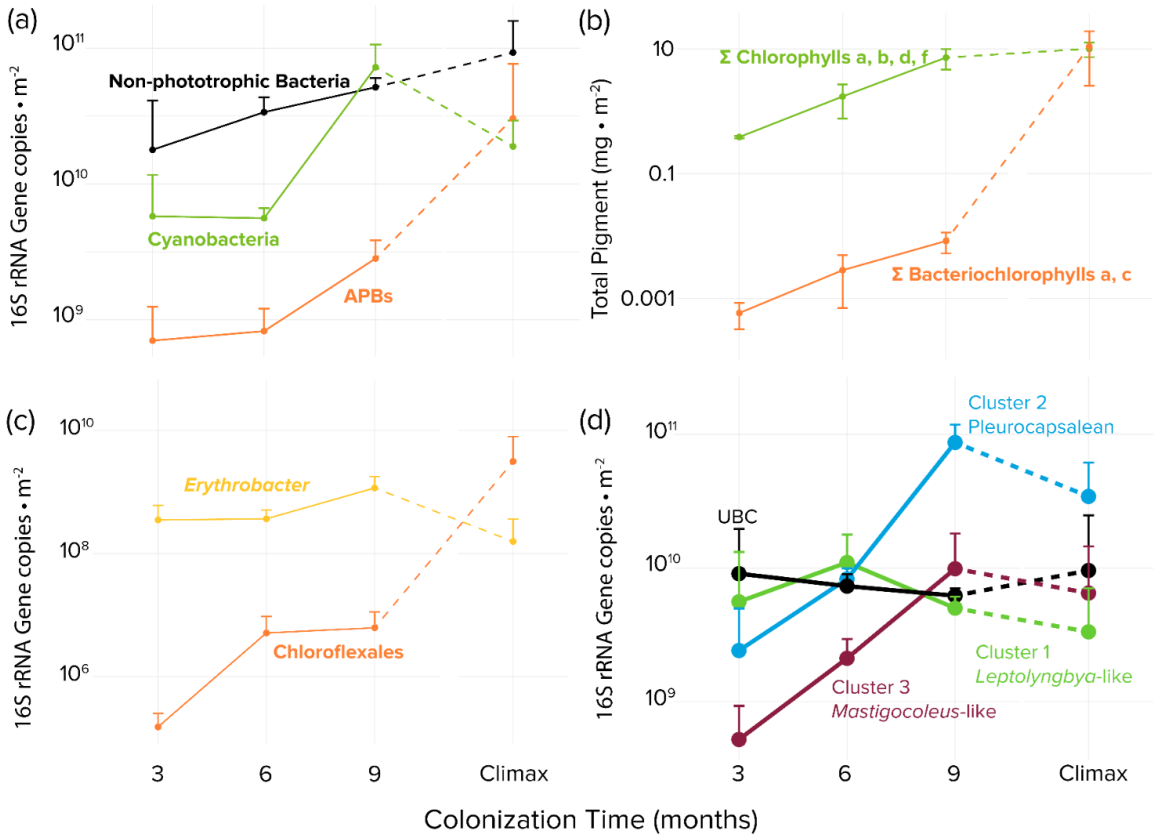

Figure 3. Time series of bacterial biomass proxies detected in colonized tiles and steady-state climax communities by guild or taxon. (**a**) Areal concentrations of 16S rRNA gene copies based on quantitative PCR and high-throughput sequencing phylogenetic assignments (**b**) Areal photosynthetic chlorins as biomarkers for oxygenic phototrophs (total chlorophylls) or APB (total bacteriochlorophylls) (**c**) areal population size of APB clades *Erythrobacter* spp. and Chloroflexales based on quantitative PCR and high-throughput sequencing phylogenetic assignments. (**d**) Endolithic colonization dynamics of specific microboring cyanobacterial clades, based on qPCR and bioinformatic placement of high-throughput environmental sequences using the Cydrasil cyanobacterial reference tree and database. Error bars are for biological sample triplicates.

3.3. Cyanobacterial Succession: Diversity and Composition

Unexpectedly, cyanobacterial richness gauged by the number of observed amplicon sequence variants (ASVs) was not significantly different across time points and when compared to steady-state climax communities (Kruskal–Wallis, $p = 0.33$; Table 1),

whereas ASV evenness (measured as Pielou's Evenness) decreased significantly (Kruskal–Wallis, $p = 0.04$) with time. Pairwise Kruskal–Wallis comparisons indicated that the difference was driven by a drop in evenness between early (3 and 6 months) and late succession communities (9 month and steady-state climax) (adjusted $p = 0.07$ for all four comparisons). Shannon's diversity also followed the evenness trend, with significant differences with time (Kruskal–Wallis, $p = 0.02$) where late succession samples were less diverse than early succession samples (adjusted $p = 0.06$ for all four comparisons). Regarding cyanobacterial community composition (beta-diversity), all time points and steady-state climax communities were significantly different from each other (PERMANOVA, $p < 0.05$, pairwise Kruskal–Wallis $p < 0.05$), a result also supported statistically by a PCOA (principal coordinates ordination analysis; Weighted UniFrac metric; Figure S2).

Table 1. Alpha diversity metrics of cyanobacterial endolithic communities in tiles placed in the intertidal zone of Isla de Mona and metrics from geographically similar natural substrate communities on Isla de Mona described by Roush et al. [15].

Timepoint	*n*	Observed ASVs	Pielou's Evenness	Shannon's Diversity
3 months	3	78 ± 4	0.74 ± 0.06 [a]	4.55 ± 0.45 [a]
6 months	3	98 ± 6	0.79 ± 0.01 [a]	5.22 ± 0.07 [b]
9 months	3	73 ± 3	0.60 ± 0.06 [b]	3.67 ± 0.50 [c]
Climax[1]	3	69 ± 3	0.62 ± 0.02 [b]	3.67 ± 0.45 [c]

[1]Community composition of steady-state climax communities was taken from calcite samples published in Roush 2018. Lower-case letters denote samples not significantly different ($\alpha = 0.1$). ASVs, amplicon sequence variants.

3.4. Identification of Endolithic Cyanobacteria Clades

In nonporous virgin substrates, only euendolithic organisms can colonize and grow to large abundance. Since we removed all epilithic biomass before sequencing, those organisms found to be abundant early on can be deemed to be bona fide

euendoliths since they must have been able to excavate the substrate. Therefore, to identify pioneer euendolithic cyanobacterial clades, the most abundant cyanobacterial ASVs from the 3-month-old tiles were placed using the RAxML Evolutionary Placement Algorithm into the Cydrasil reference cyanobacterial 16S rRNA gene tree containing 980 curated cyanobacterial sequences, which includes all full-length 16S rRNA gene sequences traceable to known euendolithic cyanobacteria (Figures 3d and S3). Euendolithic sequences that were not full length were included in the query sequence list and checked for correlation with known clades. In order to pare down the dataset for placement, we ranked each sample's cyanobacterial ASVs in order of abundance until cumulative counts reached 95% of the total abundance in each sample, yielding 213 unique ASVs across all tile samples and steady-state climax communities. We then placed the resulting pared ASV dataset into the Cydrasil reference tree. Of the 213 initial ASVs, 139 were placed with high confidence and clustered onto four distinct tree nodes. Two of the nodes contained known euendolithic species: Cluster 2 (containing 37 unique ASVs) encompassed endolithic members in the Pleurocapsales, and Cluster 3 (27 ASVs) contained *Mastigocoleus testarum*. The other two did not align with known euendoliths: one contained *Leptolyngbya* species (Cluster 1; 60 ASVs) and the other was a novel clade that contained only environmental sequences lacking taxonomic assignment and only distantly related (<95.2% similarity) to *Stanieria cyanosphaera*. We named this clade UBC (15 ASVs), for "unknown boring cluster".

3.5. Colonization Dynamics of Euendolithic Cyanobacterial Clades

To quantify colonization dynamics, qPCR-normalized abundances of the euendolithic clusters were plotted over time (Figures 3d and S3). Members of the UBC were double to an order of magnitude more abundant than the other groups after 3

months of exposure, with an average biomass of 9.1×10^9 16S rRNA gene copies per m^{-2}. UBC abundance remained stable throughout the experiment and was not significantly different when compared to steady-state climax communities. Cluster 1 (*Leptolyngbya*-like) population size lagged that of UBC, reaching a maximum after 6 months (1.1×10^9 16s rRNA gene copies per m^{-2}). Clusters 2 (Pleurocapsalean) and 3 (*Mastigocoleus*-like) colonized substrate at the slowest rate, reaching maximum populations after 9 months (8.7×10^{10} and 9.9×10^9 16S rRNA gene copies per m^{-2}, respectively). Clusters 2 (Pleurocapsalean) and 3 (*Mastigocoleus*-like) also decreased in abundance in steady-state climax communities.

3.6. Differential Abundance Analysis

In order to identify which cyanobacterial colonizers were driving compositional differences between early (3-month) and late (9-month) tiles, we conducted a differential abundance analysis. The most abundant and significant ASVs at both time points were members of the four clades delineated above (Figure S4). At 3 months, representatives of the UBC were three of the four most abundant cyanobacteria ($p < 0.05$), both in total sequence count and in differential relative abundance (fold change) with respect to 9-month communities. The fourth ASV was a member of Cluster 1, allied to *Leptolyngbya*. At 9 months, Cluster 2 (Pleurocapsalean) and Cluster 3 (*Mastigocoleus*-like) sequences were found to be the most differentially abundant with respect to 3-month communities.

3.7. New Pioneer Euendolith Clade

Both qPCR-adjusted relative abundance and differential abundance analysis revealed that the previously unknown UBC clade played a significant role in early colonization of hard intertidal carbonates. In order to better constrain its identification,

we conducted a maximum-likelihood phylogenetic reconstruction of 395 sequences, largely from cyanobacterial isolates (Figure 4), but including those of the most differentially abundant UBC and the seven sequences most similar to UBC that we could find by BLAST analyses. As before (i.e., Figure 3d), UBC members were only distantly related (<5.2% similar) to cultured cyanobacteria, the nearest being *Stanieria cyanosphaera* (formerly *Chroococcidiopsis cyanosphaera*), an epilithic freshwater unicellular cyanobacterium [54]. UBC was distant from the canonical euendolithic groups, with the Cluster 2 (Pleurocapsalean) being the closest. However, UBC members were phylogenetically close to environmental sequences obtained from marine carbonate microbialites, a habitat not dissimilar from the interior of hard carbonates and containing known euendoliths [55].

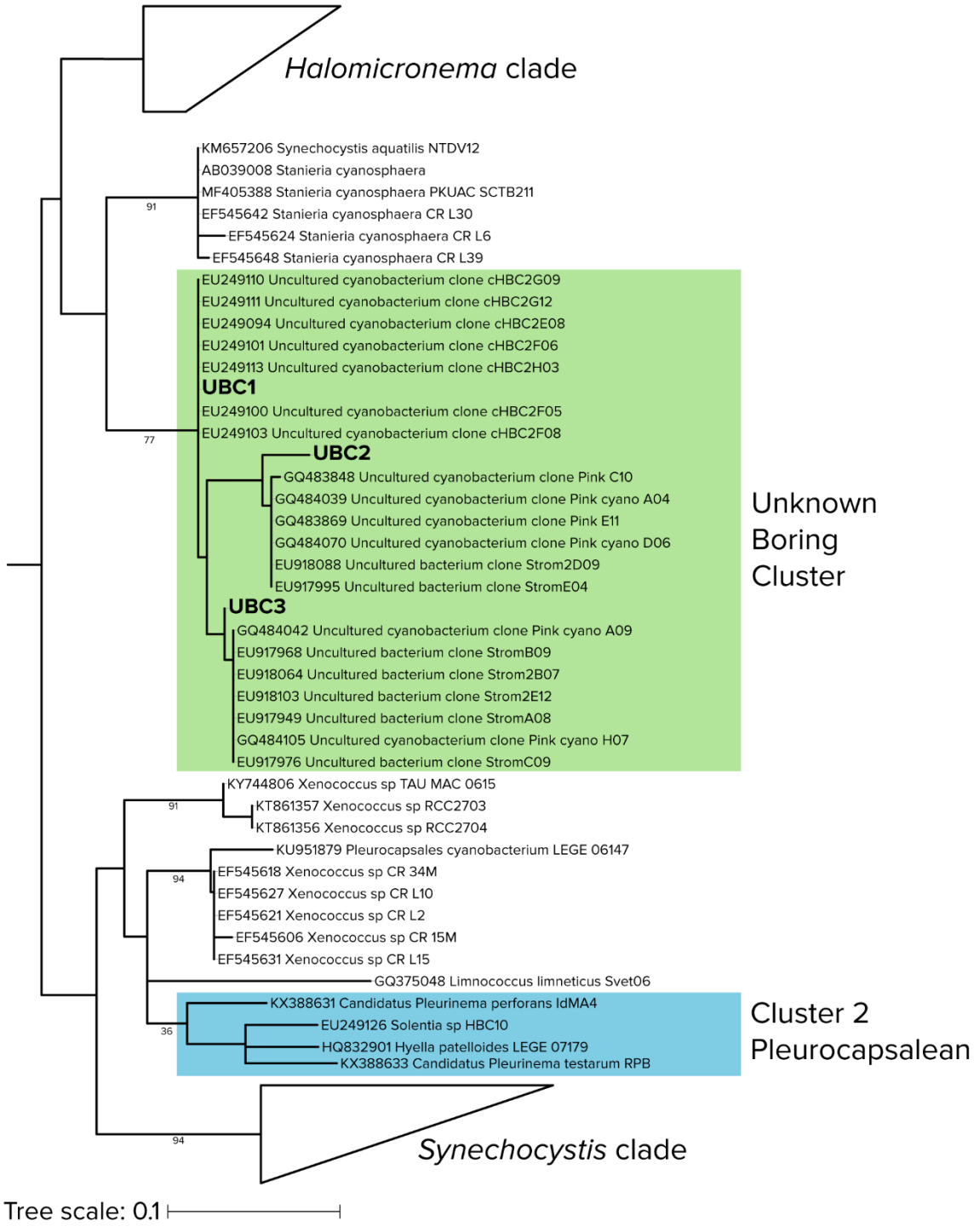

Figure 4. Detailed phylogenetic relationships of sequences in the "unknown boring cluster (UBC)", with environmental (uncultured) cyanobacterial sequences from stromatolites (shaded in green) and the closest known euendolith cluster (shaded in blue). Branch lengths are substitutions per site and node labels indicate bootstrap values.

4. Discussion

We recently reported that APBs can be major components of endolithic intertidal ecosystems and could potentially be euendolithic in nature [15], for which no precedent existed. Alternatively, these APBs may constitute secondary colonizers of opened pore space that rely on metabolic interactions with cyanobacteria, as they commonly do in other benthic environments like microbial mats or microbialites [50,55,56]. We hypothesized that examining colonization using molecular techniques and photopigment analysis specifically targeting APBs could help solve this question, in that early colonizers of bare substrates can be logically assumed to be active borers, while a dependency on cyanobacteria should result on delayed colonization by APBs. The temporal dynamics of endolithic population of Chloroflexales indeed suggest that this group of APBs are not euendoliths but instead act as secondary colonizers whose populations do not attain significance until communities of cyanobacteria are mature and the substrate has significantly eroded. The case of the proteobacterium *Erythrobacter* sp. was clearly different, since significant populations of *Erythrobacter* were present early in the colonization process and were sustained through the experimental period. *Erythrobacter* are aerobic anoxygenic phototrophs that conduct photoheterotrophy, have a low BChl *a* content, and require a source of organic carbon [24,57]. Our endolithic sequences were most similar to those in Group I *Erythrobacter* genomes [58]. Under our hypothesis, these organisms could still be euendoliths, even though their populations remained low throughout the experiment. Alternatively, since these small unicellular bacteria are abundant in coastal marine waters [24,57], they could have easily washed into fresh pits made by cyanobacteria in exposed tiles. Our current data cannot fully solve these alternatives. In fact, the metabolic action of

photoheterotrophs can increase pH levels around cells, leading the precipitation, not dissolution, of calcium carbonate [15], which would make a boring activity more difficult [59]. By contrast, the lack of the more complex photosynthetic Chloroflexales and low total bacteriochlorophylls suggests that, during colonization, euendolithic cyanobacteria dominate the photosynthetic niche due to their ability to excavate habitable space and utilize the mineral carbon for autotrophy [18,60]. Only once sufficient habitable space has been created by cyanobacteria can significant populations of APBs develop.

We found that the patterns of endolithic cyanobacterial succession within hard intertidal carbonates sustain three distinct phases (early, late succession, and steady-state climax). In our habitat, early colonization is predominantly conducted by a previously undescribed group of euendolithic cyanobacteria (UBC) that rapidly colonizes rock to maximal levels within 3 months. This clade could exceed 40% of endolithic cyanobacterial populations early on. Cluster 1 (*Leptolyngbya*-like) organisms also contribute to early colonization but only reach 60% of the biomass of UBC. By 9 months of incubation, the three canonical groups of euendolithic cyanobacteria, *Leptolyngbya (*which we tentatively equate to the *Plectonema terebrans* morphotype*),* boring members of the Pleurocapsales, and *Mastigocoleus testarum* gain a foothold. Finally, as the community reaches a steady-state climax composition, euendolithic cyanobacteria are displaced in relative importance by other cyanobacteria and by significant populations of Chloroflexalean APBs. The initial large abundance of the UBC could be explained by the presence of fast-growing propagules in natural seawater that quickly attach and bore into the substrate. Since boring microorganisms are fixed in place in their boreholes, competition for space, which can influence patterns of distribution in benthic cyanobacterial communities [61] is likely not a relevant factor until significant proportions of the rock surface become colonized. Hence, having an early foothold on the

substrate may have ensured their persistence through time, as we observed. However, UBC did not continue to increase in population size through the colonization, unlike the total cyanobacterial population, which did. The dynamics of the Cluster 1 (*Leptolyngbya*-like) members were not very different from those of UBC, although they seemed to sustain net population losses in late stages of colonization. The net gains in later stages can be attributed to Cluster 3 (*Mastigocoleus*) and, even more so, Cluster 2 (Pleurocapsalean) cyanobacteria (Figure 3d). As these slow colonizers begin to excavate more carbonate, they could reach a threshold where individual pore spaces become connected and pioneer organisms are no longer fully insulated from competition for space. Chlorophyll and qPCR data suggest that this carrying capacity is reached by 9 months of incubation. This density-dependent competition would also explain the overall decline in cyanobacterial evenness/Shannon diversity with successional progress. Finally, at maturity, as endolithic space has been colonized and the rock has become porous, nonboring endoliths can begin to colonize. One can imagine a scenario where nonboring endoliths, which need not spend energy for excavation, can outcompete borers in the outermost sections of the rock. Euendoliths would still have a competitive advantage deeper within the rock. This would be consistent with the relative decline of boring cyanobacterial ASVs in steady-state climax communities, as they are better adapted to diffusion-limited conditions. Interestingly, we did not see a difference in cyanobacterial pigment concentrations between the 9-month samples and steady-state climax communities, which suggests that nonboring phototrophs may colonize the upper interior of the rock, shading the deeper euendoliths and contributing to their decline.

A comparison of our results with prior colonization studies shows that there exist similarities, as well as marked differences, with the dynamics of porous, biogenic coral skeletons. For example, early work [4,62] demonstrated the divergence in euendolith

composition between shells and inorganic calcites. However, careful consideration must be taken as both substrate composition [9,10,12,14] and water depth [9,10] influence community structure, and, as discussed above, there are substantial differences in methodology. Even bearing those caveats in mind, the fact that all four major euendolithic clades are present after 3 months of colonization corroborates the prior conclusions that cyanobacterial colonization happens swiftly, in as little as 4 weeks, with *Plectonema, Mastigocoleus, Solentia,* and *Hyella* species all present [3,8–10]. Interestingly, there are no reports of any *Chroococcidiopsis*-like organism that could potentially represent our UBC. We also found that though *Mastigocoleus* does colonize quickly, it does not reach large abundances until the community approaches a steady-state climax composition, in contrast to the findings from corals where it is one of the first and most abundant pioneer organisms. Our observations on Cluster 1 *Leptolyngbya*-like euendoliths agree with the patterns of *P. terebrans* described by Grange et al. [11]. We find that this cluster peaks in abundance after 6 months, which was also found for coral systems. However, when comparing 9-month Cluster 1 *Leptolyngbya*-like populations to those of steady-state climax communities, we found that Cluster 1 *Leptolyngbya*-like populations were less than 10% of the 9-month totals, whereas in corals, *P. terebrans* remains very abundant through maturity [6]. Cluster 2 Pleurocapsalean euendoliths were not very abundant (sometimes <1%) in previous colonization experiments and surveys , which was attributed to their alleged susceptibility to grazing by fish and chitons due to their shallow mode of boring [11]. This was clearly not the case in our system, with Cluster 2 Pleurocapsalean organisms being the most abundant euendoliths after 9 months. Perhaps grazing pressure was unusually low in our setting, even though we did see abundant, actively grazing chitons during sampling. Though the abundance of eukaryotic euendoliths are widely reported in coral

systems [6,11], we did not find a significant contribution of plastid 16S rRNA genes in our samples, and those that were there were not phylogenetically related to known euendoliths.

In summary, by applying molecular approaches to euendolithic systems we were able to confirm that Chloroflexalean APBs act as secondary colonizers of marine carbonates, illustrate the complex dynamics of cyanobacterial colonization, and define a new clade of likely euendolithic cyanobacteria, highlighting the differences and similarities in succession dynamics between mineral and biogenic carbonates. Our work provides a first look at the complex colonization dynamics that drive bioerosion on these substrates.

Acknowledgments

We would like to thank A. Garrástazu, S. Velasco Ayuso, and B. Guida for field support. This research was partly funded by the National Science Foundation, grant number EAR 1224939 to F.G.-P.

Supplementary Materials

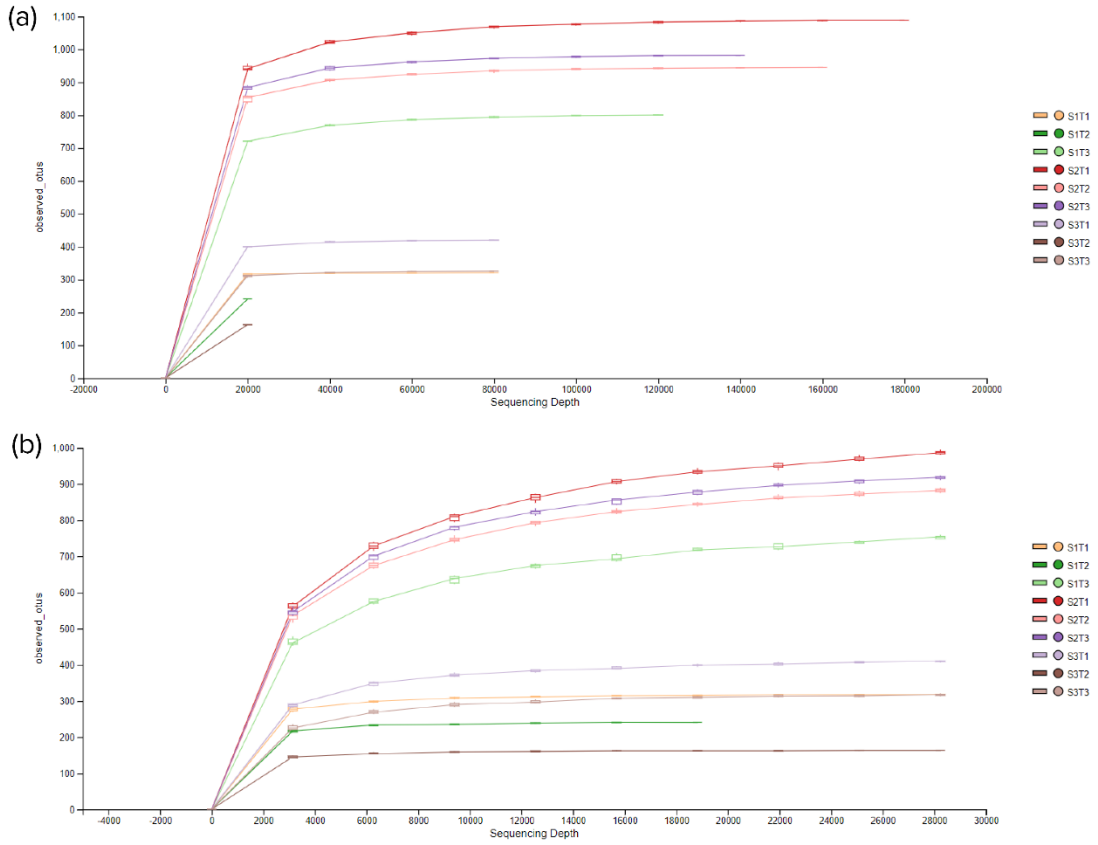

Figure S1. Alpha rarefaction curves of observed otus (Unique ASVs) at 180,000 (**a**) and 28,233 (**b**) sequencing depth. All samples reached convergence at their respective maximum sequencing depth.

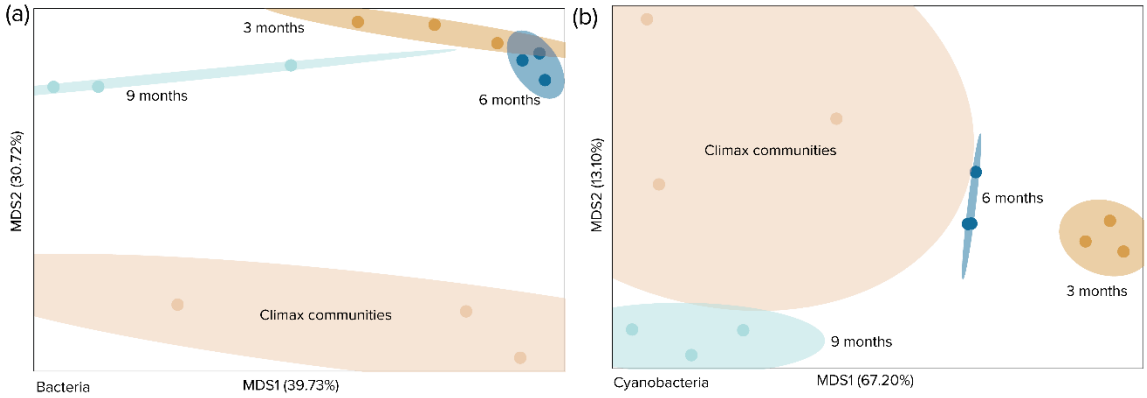

Figure S2. Principle coordinates analysis based on the weighted UniFrac metric of all bacteria (**a**) and all cyanobacteria (**b**) both showed that community composition was significantly different between all time points (PERMANOVA, $p < 0.05$).

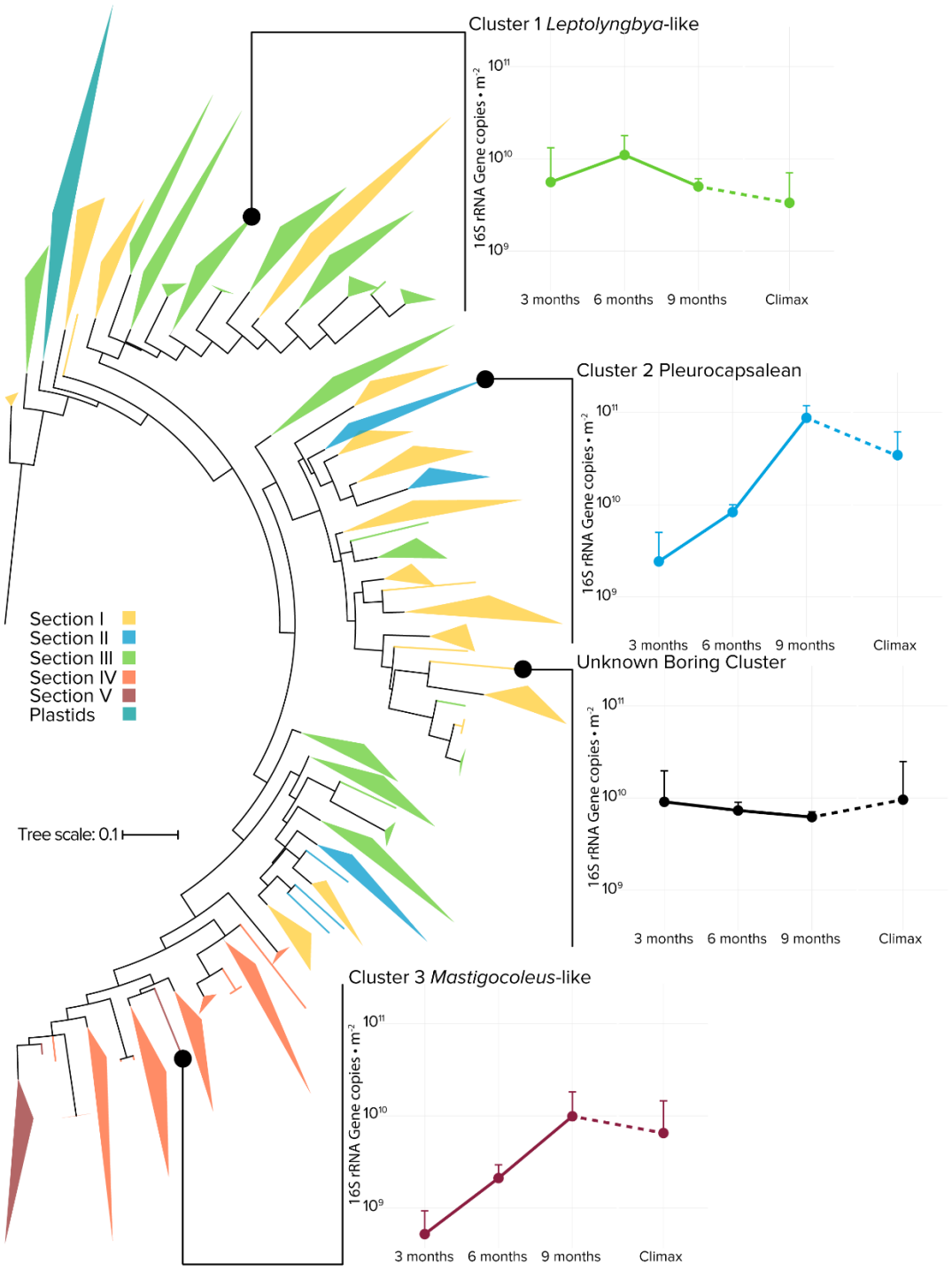

Figure S3. Endolithic colonization dynamics of specific microboring cyanobacterial clades, based on qPCR and bioinformatic placement of high-throughput environmental sequences using the Cydrasil cyanobacterial reference tree and database (left, colored by traditional morphotypical sections, *sensu* Rippka 1979 [1]). Error bars are for biological sample triplicates.

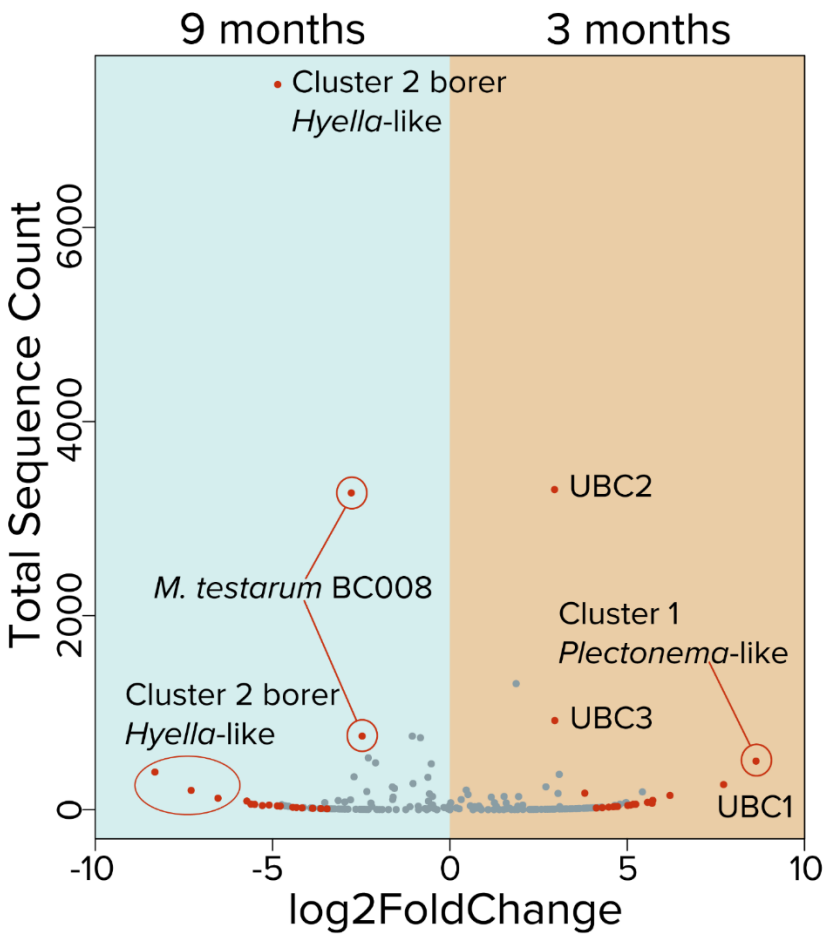

Figure S4. Volcano plot of cyanobacterial ASVs comparing differential abundance between tiles incubated for 3 months and 9 months. Red data points are significantly different ($p < 0.05$) between time points. DESeq2 analysis showed the presence of three previously unidentified ASVs not attributable to any known euendolithic genus (UBC).

References

1. Kölliker, A. On the frequent occurrence of vegetable parasites in the hard structures of animals. *Proc. R. Soc. London* **1859**, *10*, 95–99, DOI:10.5962/bhl.title.103118.

2. Duerden, J.E. Boring algae as agents in the disintegration of corals. *Bull. Am. Museum Nat. Hist.* **1902**, *889*, 323–332.

3. Le Campion-Alsumard, T. Étude Expérimentale De La Colonisation D'Éclats De Calcite Par Les Cyanophycées Endolithes Marines. *Cah. Biol. Mar.* **1975**, *16*, 177–185.

4. Le Campion-Alsumard, T. Les cyanophycées endolithes marines--Systématique, ultrastructure, écologie et biodestruction. Ocean. Acta **1979**, 2, 143–156.

5. Tribollet, A.; Golubic, S. Cross-shelf differences in the pattern and pace of bioerosion of experimental carbonate substrates exposed for 3 years on the northern Great Barrier Reef, Australia. *Coral Reefs* **2005**, *24*, 422–434, DOI:10.1007/s00338-005-0003-7.

6. Tribollet, A. Dissolution of dead corals by euendolithic microorganisms across the northern Great Barrier Reef (Australia). *Microb. Ecol.* **2008**, *55*, 569–580, DOI:10.1007/s00248-007-9302-6.

7. Chazottes, V.; Campion-Alsumard, T. Le; Peyrot-Clausade, M. Bioerosion rates on coral reefs: interactions between macroborers, microborers and grazers (Moorea, French Polynesia). *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **1995**, *113*, 189–198, DOI:10.1016/0031-0182(95)00043-L.

8. Vogel, K.; Gektidis, M.; Golubic, S.; Kiene, W.E.; Radtke, G. Experimental studies on microbial bioerosion at Lee Stocking Island, Bahamas and One Tree Island, Great Barrier Reef, Australia: Implications for paleoecological reconstructions. *Lethaia* **2000**, *33*, 190–204, DOI:10.1080/00241160025100053.

9. Gektidis, M. Development of microbial euendolithic communities : The influence of light and time. *Bull. Geol. Soc. Denmark.* **1999**, *45*, 147–150.

10. Kiene, W.; Radtke, G.; Gektidis, M.; Golubić, S.; Vogel, K. Factors controlling the distribution of microborers in Bahamian Reef environments. In *Facies*; Schuhmacher, H., Kiene, W., Dullo, W.-C., Eds.; **1995**; pp. 174–188.

11. Grange, J.S.; Rybarczyk, H.; Tribollet, A. The three steps of the carbonate biogenic dissolution process by microborers in coral reefs (New Caledonia). *Environ. Sci. Pollut. Res.* **2015**, *22*, 13625–13637, DOI:10.1007/s11356-014-4069-z.

12. Chacón, E.; Berrendero, E.; Garcia Pichel, F. Biogeological signatures of microboring cyanobacterial communities in marine carbonates from Cabo Rojo, Puerto Rico. *Sediment. Geol.* **2006**, *185*, 215–228, DOI:10.1016/j.sedgeo.2005.12.014.

13. Ramírez-Reinat, E.L.; Garcia-Pichel, F. Prevalence of Ca2+-ATPase-mediated carbonate dissolution among cyanobacterial euendoliths. *Appl. Environ. Microbiol.* **2012**, *78*, 7–13, DOI:10.1128/AEM.06633-11.

14. Couradeau, E.; Roush, D.; Guida, B.S.; Garcia-Pichel, F. Diversity and mineral substrate preference in endolithic microbial communities from marine intertidal outcrops (Isla de Mona, Puerto Rico). *Biogeosciences* **2017**, *14*, 311–324, DOI:10.5194/bg-14-311-2017.

15. Roush, D.; Couradeau, E.; Guida, B.; Neuer, S.; Garcia-Pichel, F. A new niche for anoxygenic phototrophs as endoliths. *Appl. Environ. Microbiol.* **2018**, *84*, AEM.02055-17, DOI:10.1128/AEM.02055-17.

16. Ramírez-Reinat, E.L.; Garcia-Pichel, F. Characterization of a marine cyanobacterium that bores into carbonates and the redescription of the genus *Mastigocoleus* . *J. Phycol.* **2012**, *48*, 740–749, DOI:10.1111/j.1529-8817.2012.01157.x.

17. Garcia-Pichel, F.; Ramirez-Reinat, E.; Gao, Q. Microbial excavation of solid carbonates powered by P-type ATPase-mediated transcellular Ca2+ transport. *Proc. Natl. Acad. Sci.* **2010**, *107*, 21749–21754, DOI:10.1073/pnas.1011884108.

18. Guida, B.S.; Garcia-Pichel, F. Extreme cellular adaptations and cell differentiation required by a cyanobacterium for carbonate excavation. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 5712–5717, DOI:10.1073/pnas.1524687113.

19. Guida, B.S.; Garcia-Pichel, F. Draft Genome Assembly of a Filamentous Euendolithic (True Boring) Cyanobacterium, Mastigocoleus testarum Strain BC008. *Genome Announc.* **2016**, *4*, 1–2, DOI:10.1128/genomeA.01574-15.

20. Tribollet, A.; Langdon, C.; Golubic, S.; Atkinson, M. Endolithic microflora are major primary producers in dead carbonate substrates of Hawaiian coral reefs. *J. Phycol.* **2006**, *42*, 292–303, DOI:10.1111/j.1529-8817.2006.00198.x.

21. Pierson, B.K.; Valdez, D.; Larsen, M.; Morgan, E.; Mack, E.E. Chloroflexus-like organisms from marine and hypersaline environments: Distribution and diversity. *Photosynth. Res.* **1994**, *41*, 35–52, DOI:10.1007/BF02184144.

22. Klappenbach, J.A.; Pierson, B.K. Phylogenetic and physiological characterization of a filamentous anoxygenic photoautotrophic bacterium “Candidatus Chlorothrix halophila” gen. nov., sp. nov., recovered from hypersaline microbial mats. *Arch. Microbiol.* **2004**, *181*, 17–25, DOI:10.1007/s00203-003-0615-7.

23. Hanada, S.; Takaichi, S.; Matsuura, K.; Nakamura, K. *Roseiflexus castenholzii* gen. nov., sp. nov., a thermophilic, filamentous, photosynthetic bacterium that lacks chlorosomes. *Int. J. Syst. Evol. Microbiol.* **2002**, *52*, 187–193, DOI:10.1099/00207713-52-1-187.

24. Koblížek, M.; Béjà, O.; Bidigare, R.R.; Christensen, S.; Benitez-Nelson, B.; Vetriani, C.; Kolber, M.K.; Falkowski, P.G.; Kolber, Z.S. Isolation and characterization of Erythrobacter sp. strains from the upper ocean. *Arch. Microbiol.* **2003**, *180*, 327–338, DOI:10.1007/s00203-003-0596-6.

25. Koblízek, M.; Janouškovec, J.; Oborník, M.; Johnson, J.H.; Ferriera, S.; Falkowski, P.G. Genome sequence of the marine photoheterotrophic bacterium *Erythrobacter* sp. Strain NAP1. *J. Bacteriol.* **2011**, *193*, 5881–5882, DOI:10.1128/JB.05845-11.

26. Wade, B.; Garcia-Pichel, F. Evaluation of DNA Extraction Methods for Molecular Analyses of Microbial Communities in Modern Calcareous Microbialites. *Geomicrobiol. J.* **2003**, *20*, 549–561, DOI:10.1080/01490450390249460.

27. Couradeau, E.; Karaoz, U.; Lim, H.C.; Nunes da Rocha, U.; Northen, T.; Brodie, E.; Garcia-Pichel, F. Bacteria increase arid-land soil surface temperature through the production of sunscreens. *Nat. Commun.* **2016**, *7*, 1–7, DOI:10.1038/ncomms10373.

28. Muyzer, G.; De Waal, E.; Uitterlinden, A. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain Reaction-Amplified Genes Coding for 16S rRNA. *Appl. Enviromental Microbiol.* **1993**, *59*, 695–700.

29. Caporaso, J.G.; Lauber, C.L.; Walters, W.A.; Berg-Lyons, D.; Lozupone, C.A.; Turnbaugh, P.J.; Fierer, N.; Knight, R. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci.* **2011**, *108*, 4516–4522, DOI:10.1073/pnas.1000080107.

30. Bolyen, E.; Rideout, J.R.; Dillon, M.R.; Bokulich, N.A.; Abnet, C.C.; Al-Ghalith, G.A.; Alexander, H.; Alm, E.J.; Arumugam, M.; Asnicar, F.; et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **2019**, *37*, 852–857, DOI:10.1038/s41587-019-0209-9.

31. Callahan, B.J.; McMurdie, P.J.; Rosen, M.J.; Han, A.W.; Johnson, A.J.A.; Holmes, S.P. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **2016**, *13*, 581–583, DOI:10.1038/nmeth.3869.

32. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780, DOI:10.1093/molbev/mst010.

33. Price, M.N.; Dehal, P.S.; Arkin, A.P. FastTree 2 - Approximately maximum-likelihood trees for large alignments. *PLoS ONE* **2010**, *5*, DOI:10.1371/journal.pone.0009490.

34. Lozupone, C.; Knight, R. UniFrac : A New Phylogenetic Method for Comparing Microbial Communities UniFrac : a New Phylogenetic Method for Comparing Microbial Communities. *Appl. Environ. Microbiol.* **2005**, *71*, 8228–8235, DOI:10.1128/AEM.71.12.8228.

35. DeSantis, T.Z.; Hugenholtz, P.; Larsen, N.; Rojas, M.; Brodie, E.L.; Keller, K.; Huber, T.; Dalevi, D.; Hu, P.; Andersen, G.L. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **2006**, *72*, 5069–5072, DOI:10.1128/AEM.03006-05.

36. Caporaso, J.G.; Kuczynski, J.; Stombaugh, J.; Bittinger, K.; Bushman, F.D.; Costello, E.K.; Fierer, N.; Peña, A.G.; Goodrich, J.K.; Gordon, J.I.; et al. QIIME allows analysis of high-throughput community sequencing data. *Nat. Publ. Gr.* **2010**, *7*, 335–336, DOI:10.1038/nmeth0510-335.

37. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 1–21, DOI:10.1186/s13059-014-0550-8.

38. Oksanen, A.J.; Blanchet, F.G.; Friendly, M.; Kindt, R.; Legendre, P.; Mcglinn, D.; Minchin, P.R.; Hara, R.B.O.; Simpson, G.L.; Solymos, P.; et al. Vegan: Community ecology package. *https://github.com/vegandevs/vegan* **2018**, DOI:10.4135/9781412971874.n145.

39. R Development Core Team R: A Language and Environment for Statistical Computing. *R Found. Stat. Comput.* **2011**, DOI:10.1007/978-3-540-74686-7.
40. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*; Springer-Verlag: New York, NY, USA, 2016; ISBN 978-0-387-98140-6.

41. Roush, D.; Giraldo-Silva, A.; Fernandes, V.M.C.; Maria Machado de Lima, N.; McClintock, S.; Velasco Ayuso, S.; Klicki, K.; Dirks, B.; Arantes Gama, W.; Sorochkina, K.; et al. Cydrasil: A comprehensive phylogenetic tree of cyanobacterial 16s rRNA gene sequences Available online: https://github.com/FGPLab/cydrasil (accessed on 08, August, 2018).

42. Berger, S.A.; Stamatakis, A. Aligning short reads to reference alignments and trees. *Bioinformatics* **2011**, *27*, 2068–2075, DOI:10.1093/bioinformatics/btr320.

43. Berger, S.A.; Krompass, D.; Stamatakis, A. Performance, accuracy, and web server for evolutionary placement of short sequence reads under maximum likelihood. *Syst. Biol.* **2011**, *60*, 291–302, DOI:10.1093/sysbio/syr010.

44. Letunic, I.; Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **2016**, *44*, W242–W245, DOI:10.1093/nar/gkw290.

45. Madden, T.L.; Camacho, C.; Ma, N.; Coulouris, G.; Avagyan, V.; Bealer, K.; Papadopoulos, J. BLAST+: architecture and applications. *BMC Bioinformatics* **2009**, *10*, 421, DOI:10.1186/1471-2105-10-421.

46. Nawrocki, E. Structural RNA Homology Search and Alignment Using Covariance Models, Ph. D. Thesis, Washington University School of Medicine, Saint Louis, MS, USA, December 2009.

47. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313, DOI:10.1093/bioinformatics/btu033.

48. Frigaard, N.U.; Takaichi, S.; Hirota, M.; Shimada, K.; Matsuura, K. Quinones in chlorosomes of green sulfur bacteria and their role in the redox-dependent fluorescence studied in chlorosome-like bacteriochlorophyll *c* aggregates. *Arch. Microbiol.* **1997**, *167*, 343–349, DOI:10.1007/s002030050453.

49. Frigaard, N.U.; Larsen, K.L.; Cox, R.P. Spectrochromatography of photosynthetic pigments as a fingerprinting technique for microbial phototrophs. *FEMS Microbiol. Ecol.* **1996**, *20*, 69–77, DOI:10.1016/0168-6496(96)00005-0.

50. Ley, R.E.; Harris, J.K.; Wilcox, J.; Spear, J.R.; Miller, S.R.; Bebout, B.M.; Maresca, J.A.; Bryant, D.A.; Sogin, M.L.; Pace, N.R. Unexpected Diversity and Complexity of the Guerrero Negro Hypersaline Microbial Mat Unexpected Diversity and Complexity of the Guerrero Negro Hypersaline Microbial Mat. *Appl. Environ. Microbiol.* **2006**, *72*, 3685–3695, DOI:10.1128/AEM.72.5.3685.

51. Koblížek, M. Ecology of aerobic anoxygenic phototrophs in aquatic environments. *FEMS Microbiol. Rev.* **2015**, *39*, 854–870, DOI:10.1093/femsre/fuv032.

52. Golecki, J.R.; Oelze, J. Quantitative relationship between bacteriochlorophyll content, cytoplasmic membrane structure and chlorosome size in *Chloroflexus aurantiacus*. *Arch. Microbiol.* **1987**, *148*, pp 236-241.

53. Komarek, J.; Hindak, F. Taxonomy of the new isolated strains of Chroococcidiopsis (Cyanophyceae). *Arch. fur Hydrobiol. Suppl. 46, Algol. Stud.* **1975**, *13*, 311–329.

54. Couradeau, E.; Benzerara, K.; Moreira, D.; Gérard, E.; Kaźmierczak, J.; Tavera, R.; López-García, P. Prokaryotic and eukaryotic community structure in field and cultured microbialites from the alkaline Lake Alchichica (Mexico). *PLoS One* **2011**, *6*, DOI:10.1371/journal.pone.0028767.

55. Lee, J.Z.; Burow, L.C.; Woebken, D.; Craig Everroad, R.; Kubo, M.D.; Spormann, A.M.; Weber, P.K.; Pett-Ridge, J.; Bebout, B.M.; Hoehler, T.M. Fermentation couples Chloroflexi and sulfate-reducing bacteria to Cyanobacteria in hypersaline microbial mats. *Front. Microbiol.* **2014**, *5*, 1–17, DOI:10.3389/fmicb.2014.00061.

56. Shiba, T.; Simidu, U. Erythrobacter longus gen. nov., sp. nov., an aerobic bacterium which contains bacteriochlorophyll a. *Int. J. Syst. Bacteriol.* **1982**, *32*, 211–217, DOI:10.1099/00207713-32-2-211.

57. Zheng, Q.; Lin, W.; Liu, Y.; Chen, C.; Jiao, N. A comparison of 14 Erythrobacter genomes provides insights into the genomic divergence and scattered distribution of phototrophs. *Front. Microbiol.* **2016**, *7*, DOI:10.3389/fmicb.2016.00984.

58. Garcia-Pichel, F. Plausible mechanisms for the boring on carbonates by microbial phototrophs. *Sediment. Geol.* **2006**, *185*, 205–213, DOI:10.1016/j.sedgeo.2005.12.013.

59. Guida, B.S.; Bose, M.; Garcia-Pichel, F. Carbon fixation from mineral carbonates. *Nat. Commun.* **2017**, *8*, 1–6, DOI:10.1038/s41467-017-00703-4.

60. Nübel, U.; Garcia-Pichel, F.; Kühl, M.; Muyzer, G. Spatial scale and the diversity of benthic cyanobacteria and diatoms in a salina. In *Molecular Ecology of Aquatic Communities*; Zehr, J.P., Voytek, M.A., Eds.; Springer Netherlands: Dordrecht, The Netherlands, 1999; pp. 199–206 ISBN 978-94-011-4201-4.

61. Perkins, R.D.; Tsentas, C.I. Microbial infestation of carbonate substrates planted on the St. Croix shelf, West Indies. *Bull. Geol. Soc. Am.* **1976**, *87*, 1615–1628, DO I:10.1130/0016-7606(1976)87<1615:MIOCSP>2.0.CO;2.

CHAPTER 4

CYDRASIL 2, A CURATED 16S RRNA GENE REFERENCE PACKAGE AND WEB APP FOR CYANOBACTERIAL SEQUENCE PLACEMENT

Manuscript in preparation for *Scientific Data*, **2020**

Reference package available at https://github.com/FGPLab/cydrasil **Launch**: **2018**

Web application available at https://cydrasil.org **Launch: 2020**

Coauthors have acknowledged the use of this manuscript in my dissertation

Authors:

Daniel Roush, Ana Giraldo-Silva, and Ferran Garcia-Pichel

Abstract: Cyanobacteria are one of the most widespread and important bacterial groups on the planet, providing essential ecosystem services like carbon and nitrogen fixation. However, reliable and accurate taxonomic classification of cyanobacterial 16S rRNA gene sequences is muddled due to mixed nomenclature rules and conflicting definitions of the phylum Cyanobacteria. To address this problem, we present Cydrasil 2 (https://www.cydrasil.org), a curated 16S rRNA gene reference package, database, and web application designed to provide a full phylogenetic perspective for cyanobacterial systematics and quick identification. The database, containing over 1400 manually curated sequences longer than 1100 base pairs, can be used with sequence placement algorithms, or as a reference sequence set for *de novo* phylogenetic reconstructions. The web application (utilizing PaPaRA and EPA-ng) can place tens of thousands of sequences into our reference tree and has detailed instructions on how to analyze the results. While

it makes no taxonomic assignments, it provides the framework to do so. Cydrasil 2 includes a searchable database with relevant metadata, notes and curation notes, and a mechanism for community feedback.

---

1. Introduction

Cyanobacteria are photosynthetic microorganisms that are widespread in both terrestrial and marine ecosystems [1,2] and are responsible for many ecosystem services including carbon [3] and nitrogen fixation [4]. Their detection in an environment informs researchers about the primary productivity potential of an ecosystem [1,3]. The rise of inexpensive high-throughput DNA sequencing, the development of easy to use analysis pipelines like Qiime2 [5] and mothur [6], and the availability of comprehensive taxonomic databases has made it easy for microbial ecologists to conduct preliminary microbial surveys of an ecosystem, including the abundance and diversity of cyanobacteria. Unfortunately, assigning taxonomy to organisms in a survey has inherent uncertainty due to both the computational methods commonly employed and the accuracy of reference databases. Indeed, errors in the two most used databases, Silva [7] and Greengenes [8], have been well documented [9–11], leading to spurious taxonomic assignments and error amplification. Though this uncertainty may be accepted as a trade-off for efficiency, it becomes amplified when examining cyanobacteria, due to the complex history of cyanobacterial systematics.

Unlike all other prokaryotes, cyanobacterial taxonomy is governed by both the International Code of Nomenclature of Prokaryotes [12] (ICNP) and the International Code of Nomenclature for algae, fungi, and plants [13] (ICN). Historically, cyanobacteria were originally classified as blue green algae, and much of the early work on cyanobacterial classification utilized botanical principles, including identifying new

isolates based upon their morphology. The application of molecular phylogeny techniques has since been beneficial in that it not only sped up the identification process, but also the detection and discovery of new organisms. However, the current use of molecular phylogeny techniques has at the same time resulted in the discovery of many "bad" genera, remnants of the early taxonomic system. Frequently, researchers find two organisms, sharing the same genus and at times the same species, are in fact separated by vast phylogenetic distances (Two examples of which are *Leptolyngbya* [14] and *Microcoleus* [15]). This stems from the practice of delineating new taxa by comparing sequence-based phylogenies from just a few sequences of known organisms most closely related to the one being studied, given the efforts associated to place them into comprehensive phylogenies that would provide the big picture. More uncertainty arises from the recent push for the inclusion of non-phototrophic organisms in the Cyanobacterial phylum. Soo et al. [16,17] described two sibling clades, *Melainabacteria* and *Sericytochromatia*, that are the phylogenetically closest non-phototrophic prokaryotes to the previously defined phylum of Cyanobacteria. Though this classification has been contentious within the cyanobacteria research community (see Garcia-Pichel et al. 2019 [18]), their classification as Cyanobacteria has been propagated through the major taxonomic databases, leading to confusion by researchers.

To overcome these sources of uncertainty, researchers should optimally move away from a laissez faire approach based on sequence similarity algorithms and databases with cosmopolitan criteria for sequence inclusion, to either manually curate the resulting amplicon sequences after traditional taxonomic assignment, or better, to use a complete phylogenetic perspective based upon curated, organism specific databases. Bioinformaticians have already taken steps to alleviate these issues by developing new algorithms that use the principle of phylogenetic placement.

Phylogenetic placement algorithms represent a phylogenetic-accurate and efficient way to perform classification if done on trusted databases. These algorithms were developed to place query sequences (like those from an amplicon survey) onto a precalculated reference phylogenetic tree, inferred from a curated set of reference sequences [19]. Maximum-likelihood based programs such as PPLACER [20], RAxML-EPA [21], and EPA-ng [22] take any query sequence (or a set) supplied by the user, along with a reference package (precalculated reference phylogenetic tree and alignment of all the curated sequences included in the tree), and generate a placement file (JPLACE) that contains the query sequences placed onto leaves of the reference tree with confidence values. Although phylogenetic placement is a robust method of classification, it has not been widely adopted perhaps due to the time investment required for the creation of reference packages, or the complexity in initial data analysis.

Here we present such a reference package, Cydrasil 2, intended for Cyanobacteria and its sibling bacterial clades. It offers a framework to simplify cyanobacterial classification, by providing a comprehensive and curated alignment, phylogeny, database, and web application (available at https://www.cydrasil.org) that a researcher with moderate experience can use to conduct a broad examination of the phylogeny of their sequence(s) of interest. For example, researchers could gain clarity as to which *Leptolyngbya* clade a new isolate belongs or if a set of environmental sequences classified by automated classifiers as Cyanobacteria are truly photosynthetic. In addition, we have also developed a JSON-based database file that includes notes and warnings about inconsistencies for every sequence in the reference package. We envision three common use cases for the Cydrasil reference package: provide a “first look” at the phylogenetic location of a given 16S rRNA cyanobacterial sequence within the context of a full phylogenetic reconstruction, alleviate researchers need to spend time collecting

sequences for *de novo* phylogenetic analysis, and act as a reference package for sequence placement algorithms.

2. Materials and Methods

2.1. Cydrasil Database Construction and Sequence Inclusion Criteria

From the onset, we implemented strict criteria for sequence inclusion and curation procedures of the Cydrasil database (Figure 1). Inclusion criteria were as follows:

1. The sequence must come from an isolated strain or a single-cell genome. Exceptions were made for metagenome-assembled genomes on a case by case basis after manual review of the genome (needed for most sibling clade sequences).
2. Each reference sequence must be 1100 base pairs or longer. The minimum length was chosen as a compromise between strain coverage and sequence information for phylogenetic reconstruction. Of note, this excludes all cyanobacteria sequenced using the Nübel et al. [23] cyanobacteria-specific primers.

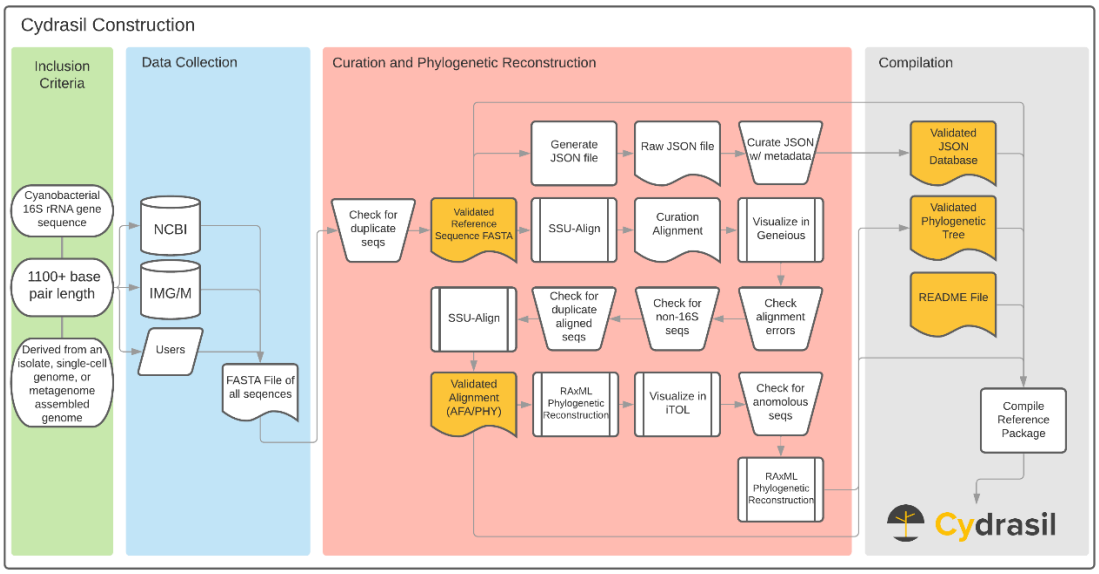

Figure 1. Process diagram describing the Cydrasil database construction workflow for each release. Yellow shapes indicate final reference package files. Standard flowchart diagram shapes were used.

Initial data collection for the Cydrasil version 1 (v1) included all available cyanobacterial 16S rRNA gene sequences that fit the above criteria and were available through the NCBI taxonomic browser up until June 2017 (Table 1), along with an outgroup comprising the closest known non-phototrophic organisms at the time, *Vampirovibrio spp*. (sister clade Melainabacteria).

Cydrasil 1.5 (v1.5) then expanded upon v1, by including all available cyanobacterial 16S rRNA sequences from genomes available on the JGI IMG/M database up until June 2019 and incorporated the first set of researcher-submitted sequences (Table 1). In this release, we also included six plastid sequences to help resolve newly discovered Greengenes errors. A new outgroup was also added that did not include any members of the sibling clades, as we were planning on doing a large sibling clade sequence addition in the 2.0 release.

Cydrasil 2, the current version, is an expansion of the v1.5 release, it includes additional curation measures reducing the total number of Cyanobacteria sequences from the previous release (described below, Table 1). For the reference package (sequence list, alignment, phylogenetic tree), we also incorporated all available 16S rRNA sequences from the sibling clades (Gastranaerophilales, Obscuribacterales, Vampirovibrionales, Margulisbacteria Saganbacteria, and Sericytochromatia) that were available through the NCBI taxonomy browser in August 2019. Since this release was also going to coincide with the release of the Cydrasil website and web application, we also created a JSON file (searchable on the Cydrasil website; included in the GitHub data deposit) that included metadata for every sequence.

Table 1. Summary statistics for major releases of Cydrasil.

Cydrasil Version	Number of Sequences				
	Total	Cyanobacteria	Outgroup	Plastids	Sibling clades
1 (rc1)	982	980	0	0	2 (root)
1.5	1494	1481	3	6	4
2	**1482**[1]	**1405**	**3**	**6**	**68**

[1] The source distribution of sequences are as follows: 1003 NCBI, 440 IMG/JGI, and 36 researcher submissions.

2.2. Data curation and phylogenetic reconstruction

The post-data collection curation procedure for each release began with a global check of the reference sequence file for header or sequence duplication. Duplicate sequences were first removed. Then, if sequence headers were found to be identical but the respective sequences were found to be unique, the headers were then modified (appending _1, _2, etc. to the header). Next, to identify duplicate aligned sequences, a curation alignment was generated using SSU-Align [24] with default parameters and masked using the ssu-mask feature of SSU-Align with per-alignment calculated masks, which removes alignment insert columns and those columns that aligned with low confidence (posterior probability < 0.95). The curation alignment was then manually examined in Geneious version 8 [25] for duplicate aligned sequences, non-16S rRNA sequences, and alignment errors. Sequences that were found to be identical post-masking were combined (one sequence was kept, and the headers combined) to reduce database redundancy and computational overhead. Sequences that were poorly aligned (typically due to the unlabeled inclusion of the ITS and/or 23S regions) were trimmed and removed if their length fell below the 1100 base pair threshold. Alignment-based curation for all Cydrasil versions 1 (rc1) and 1.5 ended here with a validated reference sequence file and the generation of a final "validated alignment."

For Cydrasil 2, an additional alignment curation step was added. Due to the introduction of 440 JGI IMG/M genome 16S rRNA sequences in v1.5, some organisms had both an NCBI sequence and an IMG/M sequence, or even multiple IMG/M sequences. These sequences were kept due to the previous curation protocol indicating that the sequences were unique. However, upon closer examination, some entries were found to be the exact same sequence, with the only difference being sequence overhang on the 5' and/or 3' end. The inclusion of both the parent and child sequences had little effect on the final reference alignment and phylogenetic tree, but for reduction of computational overhead and user readability, the longest sequence was kept, and the children sequences were removed. Then, a final "validated alignment" was generated.

After the alignment curation step, each release then underwent the same phylogenetic tree-based curation procedure. In the case of Cydrasil 2, the validated alignment was used as the input for a full maximum likelihood phylogenetic reconstruction using RAxML 8.2.12 [26] in the RAxML-HPC2 Workflow on XSEDE, part of the CIPRES [27] science gateway. The run included a maximum likelihood tree search and thorough bootstrap workflow that ran for 1000 bootstrap iterations using the GTR model with GAMMA distribution. The output curation phylogenetic tree was examined using iTOL v4 [28] for inconsistencies in taxonomic groupings, anomalous phenotypic clustering, lone wolf sequences, and sequences that directly contradict widely accepted theories regarding the evolution of Cyanobacteria. If a sequence fit any of these criteria, a literature search was conducted to identify possible causes for the abnormality. Typically, the erroneous sequences were removed, but in special cases where the organism was the type species for a genus or in common databases for taxonomic assignment, the sequence was kept with a clear warning in the header. This warning was also included in the Cydrasil 2 JSON database metadata file. Once all anomalous

sequences were removed, the tree was re-run and marked as the validated phylogenetic tree. A collapsed version of the Cydrasil 2 validated tree is shown in Figure 2, with the full tree available online (APPENDIX A).

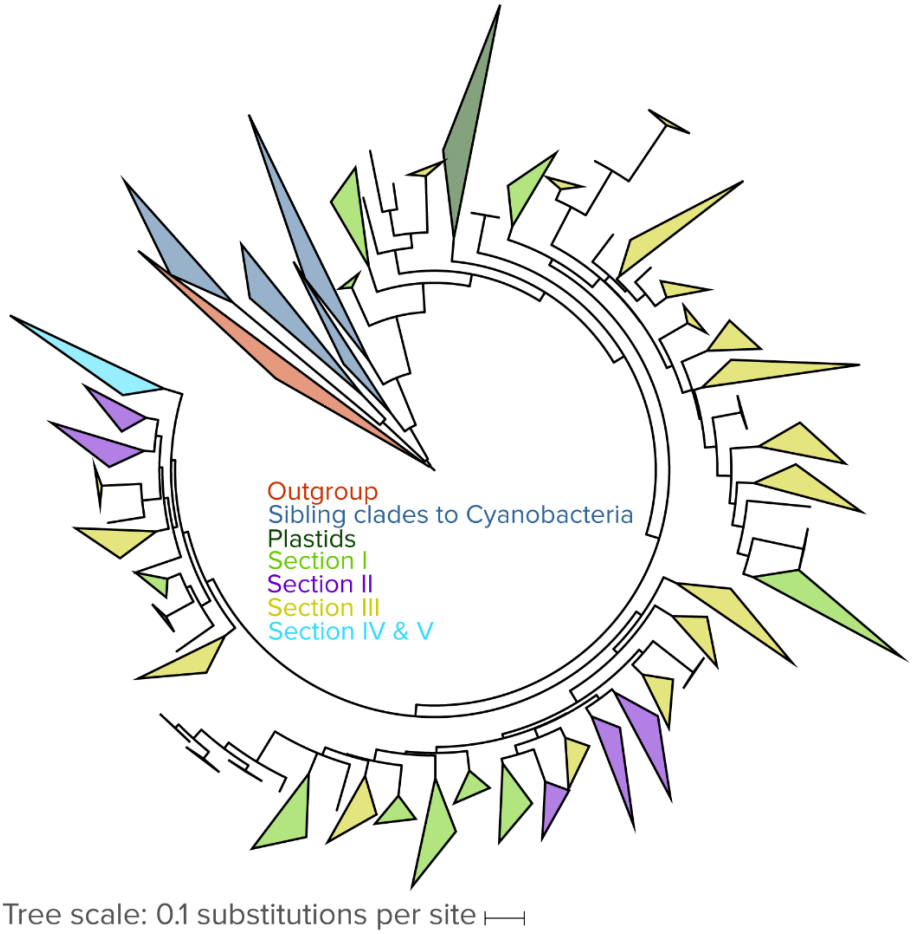

Figure 2. The Cydrasil 2 validated reference 16S rRNA cyanobacterial tree collapsed and colored based upon sections identified in Rippka et al. [29].

3. Discussion

3.1. Data Records

The Cydrasil reference dataset contains a reference sequence list (FASTA), an alignment (FASTA and PHYLIP), a SSU-Align masking file (.mask), a model file (.bestModel), a phylogenetic tree (NEWICK), a database file (JSON), and a README file

that includes instructions for using the reference database for sequence placement and data analysis using iTOL. All files are deposited in GitHub (https://github.com/FGPLab/cydrasil) and available in APPENDIX B.

The reference sequence list is a standard FASTA file with a header name that either includes the IMG id (IMG_XXXXXXXXXX) or the NCBI accession number. The header name has been formatted to be compatible with various alignment and phylogenetic programs, and as such, most non-alphanumeric characters have been converted to underscores to improve program compatibility. If a sequence is shared by multiple organisms after alignment masking with SSU-Align, the entries were combined, and the header formatted to include all names.

The Cydrasil alignment is provided in both FASTA and relaxed PHYLIP formats. This allows for the user to use various popular algorithms for aligning query sequences to the Cydrasil reference alignment. We have also included the SSU-Align mask file that was generated during reference alignment construction, in the case the user desires to use SSU-Align or Infernal for query sequence alignment.

An unrooted tree file is included in the dataset for use with sequence placement algorithms. A tree model parameter file has also been included for use with epa-ng for sequence placement.

The JSON formatted database file contains the sequence and metadata fields to provide the user with basic information about the organism and a link to the data in its respective database. An overview of these fields is described in Table 2.

Table 2. JSON keys, data type, and description for cydrasil-v2.json.

JSON Key	**Type**	**Description**
name	string	The name of the sequence from the FASTA file with underscores removed.
id	string	The id number of the sequence, initially assigned alphabetically.
sequence	string	The DNA sequence corresponding to the 16S rRNA gene with no masking.
dataSource	string	Database, publication, or lab where the sequence was retrieved.
dataSourceLink	string	A link to the entry in the corresponding dataSource or contact information for the submitting lab.
notes	string	Contains notes about sequences including other names if strains are identical, or if the organism is part of an outgroup.
warnings	string	This is reserved for warnings regarding sequence quality, taxonomic naming errors, or other oddities.

Each release also contains a README file that includes instructions and tips for using the Cydrasil reference package. The file contains step by step instructions on using Cydrasil on a local computer, links to the Cydrasil web application and a visualization of the phylogenetic tree, instructions on how to interpret results, and contact information.

3.2. Technical Validation

Database construction was entirely based upon a manual search and download of sequences from either NCBI or JGI IMG/M. Each entry was manually verified to fit the inclusion criteria before going through the extensive curation process described in the Methods. In the case of researcher submitted sequences, each sequence was manually checked for fidelity and then the submitter contacted for verification of the inclusion criteria. All sequences contain their original header names. Every sequence in the

database has dataSource and dataSourceLink information to allow for the end user to verify the original source of the sequence.

3.3. Usage Notes

The Cydrasil reference database and web application (available at https://www.cydrasil.org) is provided as a free public resource for researchers conducting taxonomic and phylogenetic analyses of cyanobacteria. We encourage any researcher looking to identify a new isolate or those conducting amplicon surveys to examine their data in the context of the full Cydrasil 2 phylogenetic reconstruction.

The construction of the Cydrasil 2 reference dataset allows for multiple use cases. The two most common (a quick check of a single sequence or an analysis of a full amplicon survey) are both based on the same sequence placement bioinformatic pipeline designed to place sequences (originally short reads from amplicon surveys, but long sequences work as well) onto the branches of a reference tree without modifying the topology. The sequence placement algorithms at the heart of this workflow, like EPA-ng and PPLACER, require a comprehensive reference package that is typically time consuming to create and curate. Cydrasil solves this problem for cyanobacterial research. We designed a web application to simplify the workflow so a user could obtain a “first look” at the phylogenetic location of a given 16S rRNA cyanobacterial sequence or analyze a full amplicon survey without the need to install any programs locally. Importantly, the web app does not provide any taxonomic assignments, but rather provides a framework for the user to those assignments on their own. The Cydrasil web application has a free, user-friendly sequence placement pipeline based on PaPaRa[30] and EPA-ng with in-depth instructions on how to analyze the output using iTOL. In the case of a full amplicon survey, the app can scale to tens of thousands of 16S rRNA

sequences and includes instructions on how to prepare the output of Qiime2 for use with Cydrasil. If a user wants to use the Cydrasil 2 database for sequence placement locally, an in-depth README file containing detailed instructions is available on the Cydrasil website and the GitHub repository.

A third use case is to use the reference dataset as a framework for *de novo* phylogenetic reconstruction of novel long to full length sequences. A user would first conduct an exploratory analysis using the sequence placement pipeline. Then, they would retrieve sequences from the database where the query sequence was placed, along with sequences belonging to the nearest neighbors and a small collection of phylogenetically close, but unrelated sequences to act as an outgroup. With the addition of other high similarity sequences from NCBI (found using a simple BLAST [31] search), a user could then generate and alignment and conduct a full *de novo* phylogenetic analysis.

Important to the underlying design of Cydrasil is the use of community feedback for future updates. Cydrasil is intended as a "living" reference package, that grows and expands with researchers' needs. We invite all users of Cydrasil to suggest possible new clades for inclusion, and additionally, submit new sequences to be incorporated in the next release. Cydrasil is under continuous development and we intend on Cydrasil being a mainstay in cyanobacterial systematics moving forward.

3.4. Code Availability

The source code for the web app is in APPENDIX C.

Acknowledgements

The authors would like to thank V. M. C. Fernandes, N. M. Machado de Lima, C. Nelson, S. McClintock, S. Velasco Ayuso, K.Klicki, B. Dirks, W. Arantes, and K. Sorochkina for their contributions to the early versions of the Cydrasil database. We also would like to thank P. Mateo, B. Roncero-Ramos and N. M. Machado de Lima for sequence contributions.

References

1. Whitton, B.A.; Potts, M. Introduction to the cyanobacteria. In *Ecology of Cyanobacteria II: Their Diversity in Space and Time*; Whitton, B.A., Ed.; Springer Netherlands, 2012; pp. 1–13 ISBN 9789400738553.

2. Soule, T.; Garcia-Pichel, F. Cyanobacteria. In *Encyclopedia of Microbiology*; Schmidt, T.M., Ed.; Elsevier Inc., 2019; pp. 799–817 ISBN 9780128117378.

3. Garcia-Pichel, F.; Belnap, J.; Neuer, S.; Schanz, F. Estimates of global cyanobacterial biomass and its distribution. *Arch. Hydrobiol. Suppl. Algol. Stud.* **2003**, *109*, 213–227, DOI:10.1127/1864-1318/2003/0109-0213.

4. Karl, D.; Michaels, A.; Bergman, B.; Capone, D.; Carpenter, E.; Letelier, R.; Lipschultz, F.; Paerl, H.; Sigman, D.; Stal, L. Dinitrogen fixation in the world's oceans. In *The Nitrogen Cycle at Regional to Global Scales*; Boyer, E.W., Howarth, R.W., Eds.; Springer, Dordrecht, 2002; pp. 47–98.

5. Bolyen, E.; Rideout, J.R.; Dillon, M.R.; Bokulich, N.A.; Abnet, C.C.; Al-Ghalith, G.A.; Alexander, H.; Alm, E.J.; Arumugam, M.; Asnicar, F.; et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **2019**, *37*, 852–857, DOI:10.1038/s41587-019-0209-9.

6. Schloss, P.D.; Westcott, S.L.; Ryabin, T.; Hall, J.R.; Hartmann, M.; Hollister, E.B.; Lesniewski, R.A.; Oakley, B.B.; Parks, D.H.; Robinson, C.J.; et al. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* **2009**, *75*, 7537–7541, DOI:10.1128/AEM.01541-09.

7. Quast, C.; Pruesse, E.; Yilmaz, P.; Gerken, J.; Schweer, T.; Yarza, P.; Peplies, J.; Glöckner, F.O. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **2013**, *41*, 590–596, DOI:10.1093/nar/gks1219.

8. DeSantis, T.Z.; Hugenholtz, P.; Larsen, N.; Rojas, M.; Brodie, E.L.; Keller, K.; Huber, T.; Dalevi, D.; Hu, P.; Andersen, G.L. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **2006**, *72*, 5069–5072, DOI:10.1128/AEM.03006-05.

9. Edgar, R. Taxonomy annotation and guide tree errors in 16S rRNA databases. *PeerJ* **2018**, *2018*, DOI:10.7717/peerj.5030.

10. Park, S.-C.; Won, S. Evaluation of 16S rRNA Databases for Taxonomic Assignments Using a Mock Community. *Genomics Inform.* **2018**, *16*, e24, DOI:10.5808/gi.2018.16.4.e24.

11. Lydon, K.A.; Lipp, E.K. Taxonomic annotation errors incorrectly assign the family Pseudoalteromonadaceae to the order Vibrionales in Greengenes: Implications for microbial community assessments. *PeerJ* **2018**, *2018*, DOI:10.7717/peerj.5248.

12. Parker, C.T.; Tindall, B.J.; Garrity, G.M. International code of nomenclature of Prokaryotes. *Int. J. Syst. Evol. Microbiol.* **2019**, DOI:10.1099/ijsem.0.000778.

13. Turland, N.J.; Wiersema, J.H.; Barrie, F.R.; Greuter, W.; Hawksworth, D.L.; Herendeen, P.S.; Knapp, S.; Kusber, W.-H.; Li, D.-Z.; Marhold, K.; et al. *International Code of Nomenclature for algae, fungi, and plants (Shenzhen Code) adopted by the Nineteenth International Botanical Congress Shenzhen, China, July 2017*; 2018; ISBN 9783946583165.

14. Stoyanov, P.; Moten, D.; Mladenov, R.; Dzhambazov, B.; Teneva, I. Phylogenetic relationships of some filamentous cyanoprokaryotic species. *Evol. Bioinforma.* **2014**, *10*, 39–49, DOI:10.4137/EBo.s13748.

15. Boyer, S.L.; Johansen, J.R.; Flechtner, V.R.; Howard, G.L. Phylogeny and genetic variance in terrestrial Microcoleus (Cyanophyceae) species based on sequence analysis of the 16S rRNA gene and associated 16S-23S its region. *J. Phycol.* **2002**, *38*, 1222–1235, DOI:10.1046/j.1529-8817.2002.01168.x.

16. Soo, R.M.; Skennerton, C.T.; Sekiguchi, Y.; Imelfort, M.; Paech, S.J.; Dennis, P.G.; Steen, J.A.; Parks, D.H.; Tyson, G.W.; Hugenholtz, P. An expanded genomic representation of the phylum cyanobacteria. *Genome Biol. Evol.* **2014**, *6*, 1031–1045, DOI:10.1093/gbe/evu073.

17. Soo, R.M.; Hemp, J.; Hugenholtz, P. Evolution of photosynthesis and aerobic respiration in the cyanobacteria. *Free Radic. Biol. Med.* **2019**, *140*, 200–205, DOI:10.1016/j.freeradbiomed.2019.03.029.

18. Garcia-Pichel, F.; Zehr, J.P.; Bhattacharya, D.; Pakrasi, H.B. What's in a name? The case of cyanobacteria. *J. Phycol.* **2020**, *56*, 1–5, DOI:10.1111/jpy.12934.

19. Felsenstein, J. Evolutionary trees from DNA sequences: A maximum likelihood approach. *J. Mol. Evol.* **1981**, DOI:10.1007/BF01734359.

20. Matsen, F.A.; Kodner, R.B.; Armbrust, E.V. pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics* **2010**, DOI:10.1186/1471-2105-11-538.

21. Berger, S.A.; Krompass, D.; Stamatakis, A. Performance, accuracy, and web server for evolutionary placement of short sequence reads under maximum likelihood. *Syst. Biol.* **2011**, *60*, 291–302, DOI:10.1093/sysbio/syr010.

22. Barbera, P.; Kozlov, A.M.; Czech, L.; Morel, B.; Darriba, D.; Flouri, T.; Stamatakis, A. EPA-ng: Massively Parallel Evolutionary Placement of Genetic Sequences. *Syst. Biol.* **2019**, *68*, 365–369, DOI:10.1093/sysbio/syy054.

23. Nübel, U.; Garcia-Pichel, F.; Muyzer, G. PCR primers to amplify 16S rRNA genes from cyanobacteria. *Appl. Environ. Microbiol.* **1997**, *63*, 3327–3332, DOI:10.1128/aem.63.8.3327-3332.1997.

24. Nawrocki, E. Structural RNA Homology Search and Alignment Using Covariance Models, Washington University School of Medicine, 2009.

25. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647–1649, DOI:10.1093/bioinformatics/bts199.

26. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313, DOI:10.1093/bioinformatics/btu033.

27. Miller, M.A.; Pfeiffer, W.; Schwartz, T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *2010 Gatew. Comput. Environ. Work. GCE 2010* **2010**, DOI:10.1109/GCE.2010.5676129.

28. Letunic, I.; Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **2016**, *44*, W242–W245, DOI:10.1093/nar/gkw290.

29. Rippka, R.; Deruelles, J.; Waterbury, J.B. Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *J. Gen. Microbiol.* **1979**, *111*, 1–61, DOI:10.1099/00221287-111-1-1.

30. Berger, S.A.; Stamatakis, A. PaPaRa 2.0: a vectorized algorithm for probabilistic phylogeny-aware alignment extension. *... Inst. Theor. Stud. http//sco. h ...* **2012**.

31. Madden, T.L.; Camacho, C.; Ma, N.; Coulouris, G.; Avagyan, V.; Bealer, K.; Papadopoulos, J. BLAST+: architecture and applications. *BMC Bioinformatics* **2009**, *10*, 421, DOI:10.1186/1471-2105-10-421.

CHAPTER 5

CONCLUSIONS

1. General Conclusions

Through the use of modern molecular methodologies for microbial ecology and phylogeny, the work presented in this dissertation contributes a phylogenetically sound, functionally more defined understanding of the carbonate endolithic microbiome as it pertains to its component microorganisms and their dynamics. As an additional contribution, I demonstrated that the use of a bioinformatic pipeline based on curated, comprehensive databases can effectively assist in the notoriously difficult task of cyanobacterial identification and systematics.

I demonstrated that anoxygenic phototropic bacteria (APBs) are numerically important endoliths of marine carbonates in a variety of rock types and geographical locations. These APBs belong to the Chloroflexi (genera *Roseiflexus* and *Chlorothrix*), the proteobacterial *Erythrobacter* sp. (*affiliated to strain* NAP1) and purple non-sulfur alphaproteobacterial. Furthermore, APBs were differentially distributed with geography, with Chloroflexi dominant on Isla de Mona (Caribbean Sea) and *Erythrobacter* on Menorca (Mediterranean Sea), indicating that endolithic APB composition may be dependent upon water temperature. Given their ubiquity and numbers, endolithic APBs may have been misidentified as morphologically similar cyanobacteria and could potentially play important metabolic roles by exerting geomicrobial effects on coastal carbonates.

I could also establish through experiments using de novo colonization of artificial carbonates, that filamentous Chloroflexalean APBs were not euendoliths but secondary colonizers that invade the substrate only after it has been excavated and made porous by cyanobacterial euendoliths. In contrast, small unicellular *Erythrobacter* were present

throughout the colonization period making it unclear if they were euendolithic, or opportunistic endolithic colonizers. In these experiments, the endolithic photosynthetic microbiome reached steady state maturity within nine months and showed clear temporal dynamics. I proposed a three-phase endolithic succession pattern (early, late succession, and steady-state climax) to describe it. The early phase was populated by a novel clade of cyanobacteria (UBC), accompanied by a lesser number of *Leptolyngbya*-like organisms. Late successional phases were characterized by the advent of traditional cyanobacterial euendoliths: *Leptolyngbya* (equated to the *Plectonema terebrans* morphotype*),* boring members of the Pleurocapsales, and *Mastigocoleus testarum*. The steady-state climax phase exhibited density-dependent competition where non-boring phototrophic endoliths partly displaced euendolithic cyanobacteria, including significant populations of Chloroflexalean APBs and cyanobacteria. This molecular-based colonization model differs significantly from that found in the traditional literature, illustrating that euendolith ecology is more complex than previously thought.

Lastly, I developed and presented Cydrasil, a curated 16S rRNA gene reference package, database, and web application designed to provide a complete phylogenetic perspective for cyanobacterial systematics and molecular ecology. While fully functional, Cydrasil was designed as an evolving tool, with inherent capacity for continual expansion and updating by users. It should find application well beyond the field of endolithic microbiomes.

2. Future Perspectives

While I set out to establish a baseline understanding of phototropic endolith ecology in carbonates, many questions remained unanswered. The most interesting of which concern astrobiology and planetary evolution. On young planets (including Earth)

with no atmospheric oxygen, the endolithic habitat provides a ready-made microbial bunker to protect from UV-C DNA damage. Additionally, anoxygenic photosynthesis is thought to have evolved early in Earth's history. Could endolithic anoxygenic phototrophs have evolved as some of the first land inhabiting organisms, supporting the first non-aquatic biospheres? Earth's early chemistry would support anoxygenic photosynthesis, and combined with an abundance of exposed rocky substrates, one could imagine a possibility where APBs act as the primary producers in the some of the first land-based biospheres. There are examples of endolithic and euendolithic microbes in the fossil record, but they have been attributed to cyanobacteria. The fossils may need to be reexamined, taking into consideration the existence of endolithic APBs. These fossils may in fact support early land-based APB biospheres. Extending this hypothesis to exoplanet exploration and astrobiology, researchers should now include APB photopigments as potential biomarkers when looking for life on other worlds.

In modern endolithic habitats, I hypothesized that endolithic anoxygenic phototrophs were most likely photoheterotrophs due to their phylogenetic proximity to known photoheterotrophic APBs. Understanding the geomicrobial effects of both APBs (through cultures) and whole prokaryotic community metabolism will be essential to properly model and understand bioerosion in this era of rapid climate change. Unpublished data from both high-throughput carbonate surveys [1,2] revealed a large non-phototrophic prokaryote community comprised of fermentative organisms, suggesting that carbonates submerged at high tide and/or at night rapidly become anoxic, supporting diverse fermentative metabolisms. Many fermentation products are acidic and would have the ability to dissolve carbonate, potentially providing another bioerosive vector in these ecosystems. This question should be answered using combined

isolation-, metagenomics-, and metabolomics-based approaches to ascertain the metabolic potential of these organisms.

While I saw clear geographic differences in the survey, the small number of climatically different locations, made generalization impossible. Mapping the global distribution of endolithic APBs thus remains an open question. Preliminary data (APPENDIX V) obtained from Churchill, Manitoba Canada, Tonto Natural Bridge, Arizona USA, and Isla de Mona hint that APBs may be widespread within carbonate endolithic communities, and not just limited to carbonate platform islands. Churchill sits on the Hudson Bay at the edge of the Arctic, and dolomite endolithic microbiomes from the intertidal zone there showed the presence of the same APB groups as both Isla de Mona and Menorca. Similar findings came from freshwater and terrestrial carbonates from Tonto and cliffs in the rainforest interior of Isla de Mona. If endolithic APBs are indeed a global phenomenon, then the potential biomass reservoir must be considered when looking at phototrophic potential in the environment. The geomicrobial contributions from APBs on a global scale would need to be reevaluated, including their role in the carbon cycle, and other potential knock-on effects in various nutrient cycles.

The colonization experiments I conducted also provided a few unexpected discoveries that have opened questions about the nature of cyanobacterial euendoliths. The novel clade, UBC, is phylogenetically distant from all known euendoliths and most cultured cyanobacteria. It likely represents a new genus. UBC's nearest cultured relative is *Stanieria cyanosphaera* which may indicate that UBC may be the elusive unicellular euendolith that has a passing mention in the literature [3]. Only microscopic identification, or a cultured isolate will be able to resolve this question. Finally, the phylogenetic location of *Leptolyngbya*-like organisms identified in my colonization experiment may provide a clue to resolving the *Plectonema* mystery and provide

coherence between ages old morphological descriptions and modern phylogenetics. Efforts to cultivate such novel euendoliths should be pursued as a means to advance the impasse.

The intention of my dissertation was to lay a modern microbial ecology foundation and provide a framework that other researchers would benefit from in the field of endolith ecology. The application of modern methods led to novel findings and suggests that further investigation of endolithic microbiomes in this manner will undoubtedly reveal more complex microbial and geomicrobial interactions.

References

1. Couradeau, E.; Roush, D.; Guida, B.S.; Garcia-Pichel, F. Diversity and mineral substrate preference in endolithic microbial communities from marine intertidal outcrops (Isla de Mona, Puerto Rico). *Biogeosciences* **2017**, *14*, 311–324, DOI:10.5194/bg-14-311-2017.

2. Roush, D.; Couradeau, E.; Guida, B.; Neuer, S.; Garcia-Pichel, F. A new niche for anoxygenic phototrophs as endoliths. *Appl. Environ. Microbiol.* **2018**, *84*, AEM.02055-17, DOI:10.1128/AEM.02055-17.

3. Ercegović, A. La végétation lithophytes sur les calcaires et les dolomites en Croatie. *Acta Bot. Croat.* **1925**.

REFERENCES

1. Edwards, K.J.; Bach, W.; Rogers, D.R. Geomicrobiology of the ocean crust: A role for chemoautotrophic Fe-bacteria. *Biol. Bull.* **2003**, *204*, 180–185, DOI:10.2307/1543555.

2. Singer, E.; Chong, L.S.; Heidelberg, J.F.; Edwards, K.J. Similar microbial communities found on two distant seafloor basalts. *Front. Microbiol.* **2015**, *6*, 1–11, DOI:10.3389/fmicb.2015.01409.

3. Russell, N.C.; Edwards, H.G.M.; Survey, B.A.; Environment, N.; Cross, H.; Road, M.; Oet, C.C.B.; Wynn-Williams, D.D. FT-Raman spectroscopic analysis of endolithic microbial communities from Beacon sandstone in Victoria Land, Antarctica. *Antart. Sci.* **1998**, *10*, 63–74.

4. Budel, B.; Weber, B.; Kuhl, M.; Pfanz, H.; Sultemeyer, D.; Wessels, D. Reshaping of sandstone surfaces by cryptoendolithic cyanobacteria: bioalkalization causes chemical weathering in arid landscapes. *Geobiology* **2004**, *2*, 261–268, DOI:10.1111/j.1472-4677.2004.00040.x.

5. Horath, T.; Bachofen, R. Molecular characterization of an endolithic microbial community in dolomite rock in the central Alps (switzerland). *Microb. Ecol.* **2009**, *58*, 290–306, DOI:10.1007/s00248-008-9483-7.

6. Chacón, E.; Berrendero, E.; Garcia Pichel, F. Biogeological signatures of microboring cyanobacterial communities in marine carbonates from Cabo Rojo, Puerto Rico. *Sediment. Geol.* **2006**, *185*, 215–228, DOI:10.1016/j.sedgeo.2005.12.014.

7. Ramírez-Reinat, E.L.; Garcia-Pichel, F. Prevalence of Ca2+-ATPase-mediated carbonate dissolution among cyanobacterial euendoliths. *Appl. Environ. Microbiol.* **2012**, *78*, 7–13, DOI:10.1128/AEM.06633-11.

8. Friedmann, E.I.; Ocampo, R. Endolithic blue-green algae in the dry valleys: primary producers in the antartic desert ecosystem. *Science (80-.).* **1976**, *193*, 1247–1249.

9. Friedmann, E.I. Endolithic microorganisms in the antarctic cold desert. *Science* **1982**, *215*, 1045–1053, DOI:10.1126/science.215.4536.1045.

10. Stjepko Golubic, Imre Friedmann, Ju, S.; Friedmann, E.I.; Schneider, J. The lithobiontic ecological niche, with special reference to microorganisms. *J Sediment Res* **1981**, DOI:10.1306/212F7CB6-2B24-11D7-8648000102C1865D.

11. Wierzchos, J.; Ascaso, C.; McKay, C.P. Endolithic cyanobacteria in halite rocks from the hyperarid core of the Atacama Desert. *Astrobiology* **2006**, *6*, 415–422, DOI:10.1089/ast.2006.6.415.

12. Dong, H.; Rech, J.A.; Jiang, H.; Sun, H.; Buck, B.J. Endolithnic cyanobacteria in soil gypsum: Occurences in Atacama (Chile), Mojave (United States), and Al-Jafr Basin (Jordan) Deserts. *J. Geophys. Res. Biogeosciences* **2007**, *112*, 1–11, DOI:10.1029/2006JG000385.

13. Ziolkowski, L. a; Wierzchos, J.; Davila, A.F.; Slater, G.F. Radiocarbon evidence of active endolithic microbial communities in the hyperarid core of the Atacama Desert. *Astrobiology* **2013**, *13*, 607–616, DOI:10.1089/ast.2012.0854.

14. De La Torre, J.R.; Goebel, B.M.; Friedmann, E.I.; Pace, N.R. Microbial diversity of cryptoendolithic communities from the McMurdo Dry Valleys, Antarctica. *Appl Env. Microbiol* **2003**, *69*, 3858–3867, DOI:10.1128/AEM.69.7.3858.

15. Palmer, R.J.; Friedmann, E.I. Water relations and photosynthesis in the cryptoendolithic microbial habitat of hot and cold deserts. *Microb. Ecol.* **1990**, *19*, 111–118, DOI:10.1007/BF02015057.

16. Couradeau, E.; Roush, D.; Guida, B.S.; Garcia-Pichel, F. Diversity and mineral substrate preference in endolithic microbial communities from marine intertidal outcrops (Isla de Mona, Puerto Rico). *Biogeosciences* **2017**, *14*, 311–324, DOI:10.5194/bg-14-311-2017.

17. Golubić, S.; Pietrini, A.M.; Ricci, S. Euendolithic activity of the cyanobacterium Chroococcus lithophilus Erc. In biodeterioration of the Pyramid of Caius Cestius, Rome, Italy. *Int. Biodeterior. Biodegrad.* **2015**, *100*, 7–16, DOI:10.1016/j.ibiod.2015.01.019.

18. Kobluk, D.R.; Risk, M.J. Rate and nature of infestation of a carbonate substratum by a boring alga. *J. Exp. Mar. Bio. Ecol.* **1977**, *27*, 107–115, DOI:10.1016/0022-0981(77)90131-9.

19. Ramírez-Reinat, E.L.; Garcia-Pichel, F. Characterization of a marine cyanobacterium that bores into carbonates and the redescription of the genus *Mastigocoleus* . *J. Phycol.* **2012**, *48*, 740–749, DOI:10.1111/j.1529-8817.2012.01157.x.

20. Roush, D.; Couradeau, E.; Guida, B.; Neuer, S.; Garcia-Pichel, F. A new niche for anoxygenic phototrophs as endoliths. *Appl. Environ. Microbiol.* **2018**, *84*, AEM.02055-17, DOI:10.1128/AEM.02055-17.

21. Tribollet, A. Dissolution of dead corals by euendolithic microorganisms across the northern Great Barrier Reef (Australia). *Microb. Ecol.* **2008**, *55*, 569–580, DOI:10.1007/s00248-007-9302-6.

22. Grange, J.S.; Rybarczyk, H.; Tribollet, A. The three steps of the carbonate biogenic dissolution process by microborers in coral reefs (New Caledonia). *Environ. Sci. Pollut. Res.* **2015**, *22*, 13625–13637, DOI:10.1007/s11356-014-4069-z.

23. Enochs, I.C.; Manzello, D.P.; Tribollet, A.; Valentino, L.; Kolodziej, G.; Donham, E.M.; Fitchett, M.D.; Carlton, R.; Price, N.N. Elevated Colonization of Microborers at a Volcanically Acidified Coral Reef. *PLoS One* **2016**, *11*, 1–16, DOI:10.1371/journal.pone.0159818.

24. Tribollet, A.; Langdon, C.; Golubic, S.; Atkinson, M. Endolithic microflora are major primary producers in dead carbonate substrates of Hawaiian coral reefs. *J. Phycol.* **2006**, *42*, 292–303, DOI:10.1111/j.1529-8817.2006.00198.x.

25. Guida, B.S.; Bose, M.; Garcia-Pichel, F. Carbon fixation from mineral carbonates. *Nat. Commun.* **2017**, *8*, 1–6, DOI:10.1038/s41467-017-00703-4.

26. Shachak, M.; Jones, C.G.; Granot, Y. Herbivory in rocks and the weathering of a desert. *Science (80-.).* **1987**, *236*, 1098–1099, DOI:10.1126/science.236.4805.1098.

27. Vogel, K.; Gektidis, M.; Golubic, S.; Kiene, W.E.; Radtke, G. Experimental studies on microbial bioerosion at Lee Stocking Island, Bahamas and One Tree Island, Great Barrier Reef, Australia: Implications for paleoecological reconstructions. *Lethaia* **2000**, *33*, 190–204, DOI:10.1080/00241160025100053.

28. Tribollet, A.; Payri, C. Bioerosion of the coralline alga Hydrolithon onkodes by microborers in the coral reefs of Moorea, French Polynesia. *Oceanol. Acta* **2001**, *24*, 329–342, DOI:10.1016/S0399-1784(01)01150-1.

29. Ford, D.; Williams, P. The Global Distribution Of Karst. In *Karst Hydrogeology and Geomorphology*; 2013; pp. 1–8 ISBN 9781118684986.

30. Ćurin, M.; Peharda, M.; Calcinai, B.; Golubić, S. Incidence of damaging endolith infestation of the edible mytilid bivalve Modiolus barbatus. *Mar. Biol. Res.* **2014**, *10*, 179–189, DOI:10.1080/17451000.2013.814793.

31. Pfister, C.A.; Meyer, F.; Antonopoulos, D.A. Metagenomic profiling of a microbial assemblage associated with the california mussel: A node in networks of carbon and nitrogen cycling. *PLoS One* **2010**, *5*, 1–10, DOI:10.1371/journal.pone.0010518.

32. Kaehler, S. Incidence and distribution of phototrophic shell-degrading endoliths of the brown mussel *Perna perna* . *Mar. Biol.* **1999**, *135*, 505–514, DOI:10.1007/s002270050651.

33. Raghukumar, C.; Sharma, S.; Lande, V. Distribution and biomass estimation of shell-boring algae in the intertidal at Goa, India. *Phycologia* **1991**, *30*, 303–309, DOI:10.2216/i0031-8884-30-3-303.1.

34. Kiene, W.; Radtke, G.; Gektidis, M.; Golubić, S.; Vogel, K. Factors controlling the distribution of microborers in Bahamian Reef environments. In *Facies*; Schuhmacher, H., Kiene, W., Dullo, W.-C., Eds.; 1995; pp. 174–188.

35. Gektidis, M. Development of microbial euendolithic communities : The influence of light and time. *Bull. Geol. Soc. Denmark*. **1999**, *45*, 147–150.

36. Guida, B.S.; Garcia-Pichel, F. Extreme cellular adaptations and cell differentiation required by a cyanobacterium for carbonate excavation. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 5712–5717, DOI:10.1073/pnas.1524687113.

37. Knoll, A.H.; Golubic, S.; Green, J.; Swett, K. Organically preserved microbial endoliths from the late Proterozoic of East Greenland. *Nature* **1986**, *321*, 856–857, DOI:10.1038/321856a0.

38. Zhang, X. guang; Pratt, B.R. Microborings in Early Cambrian phosphatic and phosphatized fossils. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **2008**, *267*, 185–195, DOI:10.1016/j.palaeo.2008.06.015.

39. Radtke, G.; Glaub, I.; Vogel, K.; Golubic, S. isp. nov., Distribution, Variability and Biological Origin. *Ichnos* **2010**, *17*, 25–33, DOI:10.1080/10420940903358628.

40. Glaub, I. Paleobathymetric reconstructions and fossil microborings. *Bull. Geol. Soc. Denmark* **1999**, *45*, 143–146.

41. Elias, R.J.; Lee, D. Microborings and Growth in Late Ordovician Halysitids and Other Corals. *J. Paleontol.* **1993**, *67*, 922–934.

42. Kölliker, A. On the frequent occurrence of vegetable parasites in the hard structures of animals. *Proc. R. Soc. London* **1859**, *10*, 95–99, DOI:10.5962/bhl.title.103118.

43. Bornet, E.; Flahault, C. Note sur deux nouveaux genres d'algues perforantes. *J. Bot.* **1888**.

44. Duerden, J.E. Boring algae as agents in the disintegration of corals. *Bull. Am. Museum Nat. Hist.* **1902**, *889*, 323–332.

45. Ercegović, A. La végétation lithophytes sur les calcaires et les dolomites en Croatie. *Acta Bot. Croat.* **1925**.

46. Goiubic, S. Distribution, taxonomy, and boring patterns of marine endolithic algae. *Integr. Comp. Biol.* **1969**, DOI:10.1093/icb/9.3.747.

47. Budd, D.A.; Perkins, R.D. Bathymetric zonation and paleoecological significance of algal microborings in Puerto Rican shelf and slope sediments. *J. Sediment. Petrol.* **1980**, DOI:10.1306/212F7B17-2B24-11D7-8648000102C1865D.

48. Chazottes, V.; Campion-Alsumard, T. Le; Peyrot-Clausade, M. Bioerosion rates on coral reefs: interactions between macroborers, microborers and grazers (Moorea, French Polynesia). *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **1995**, *113*, 189–198, DOI:10.1016/0031-0182(95)00043-L.

49. Pantazidou, A.; Louvrou, I.; Economou-Amilli, A. Euendolithic shell-boring cyanobacteria and chlorophytes from the saline lagoon Ahivadolimni on Milos Island, Greece. *Eur. J. Phycol.* **2006**, DOI:10.1080/09670260600649420.

50. Wisshak, M.; Tribollet, A.; Golubic, S.; Jakobsen, J.; Freiwald, A. Temperate bioerosion: Ichnodiversity and biodiversity from intertidal to bathyal depths (Azores). *Geobiology* **2011**, DOI:10.1111/j.1472-4669.2011.00299.x.

51. Campbell, S.E. The modern distribution and geological history of calcium carbonate boring microorganisms. *Biominer. Biol. Met. Accumul.* **1983**, 99–104, DOI:10.1007/978-94-009-7944-4.

52. Perkins, R.D.; Tsentas, C.I. Microbial infestation of carbonate substrates planted on the St. Croix shelf, West Indies. *Bull. Geol. Soc. Am.* **1976**, *87*, 1615–1628, DOI:10.1130/0016-7606(1976)87<1615:MIOCSP>2.0.CO;2.

53. Campbell, S.E.; Cole, K. Developmental studies on cultured endolithic conchocelis (Rhodophyta). *Hydrobiologia* **1984**, *116/117*, 201–208, DOI:10.1007/BF00027666.

54. Brito, Â.; Ramos, V.; Seabra, R.; Santos, A.; Santos, C.L.; Lopo, M.; Ferreira, S.; Martins, A.; Mota, R.; Frazão, B.; et al. Culture-dependent characterization of cyanobacterial diversity in the intertidal zones of the Portuguese coast: A polyphasic study. *Syst. Appl. Microbiol.* **2012**, *35*, 110–119, DOI:10.1016/j.syapm.2011.07.003.

55. Foster, J.S.; Green, S.J.; Ahrendt, S.R.; Golubic, S.; Reid, R.P.; Hetherington, K.L.; Bebout, L. Molecular and morphological characterization of cyanobacterial diversity in the stromatolites of Highborne Cay, Bahamas. *ISME J.* **2009**, *3*, 573–587, DOI:10.1038/ismej.2008.129.

56. Walker, J.J.; Spear, J.R.; Pace, N.R. Geobiology of a microbial endolithic community in the Yellowstone geothermal environment. *Nature* **2005**, *434*, 1011–1014, DOI:10.1038/nature03447.

57. Walker, J.J.; Pace, N.R. Endolithic Microbial Ecosystems. *Annu. Rev. Microbiol.* **2007**, *61*, 331–347, DOI:10.1146/annurev.micro.61.080706.093302.

58. Crits-Christoph, A.; Gelsinger, D.R.; Ma, B.; Wierzchos, J.; Ravel, J.; Davila, A.; Casero, M.C.; DiRuggiero, J. Functional interactions of archaea, bacteria and viruses in a hypersaline endolithic community. *Environ. Microbiol.* **2016**, *18*, 2064–2077, DOI:10.1111/1462-2920.13259.

59. Couradeau, E.; Roush, D.; Guida, B.S.; Garcia-Pichel, F. Diversity and mineral substrate preference in endolithic microbial communities from marine intertidal outcrops (Isla de Mona, Puerto Rico). *Biogeosciences* **2017**, *14*, 311–324, DOI:10.5194/bg-14-311-2017.

60. Belnap, J.; Büdel, B.; Lange, O.L. Biological Soil Crusts: Characteristics and Distribution. In *Biological Soil Crusts: Structure, Function, and Management*; 2003; Vol. 150, pp. 3–30 ISBN 978-3-540-43757-4.

61. Couradeau, E.; Benzerara, K.; Moreira, D.; Gérard, E.; Kaźmierczak, J.; Tavera, R.; López-García, P. Prokaryotic and eukaryotic community structure in field and cultured microbialites from the alkaline Lake Alchichica (Mexico). *PLoS One* **2011**, *6*, DOI:10.1371/journal.pone.0028767.

62. Magnusson, S.H.; Fine, M.; Kühl, M. Light microclimate of endolithic phototrophs in the scleractinian corals *Montipora monasteriata* and *Porites cylindrica* . *Mar. Ecol. Prog. Ser.* **2007**, *332*, 119–128, DOI:10.3354/meps332119.

63. Yang, S.H.; Lee, S.T.M.; Huang, C.R.; Tseng, C.H.; Chiang, P.W.; Chen, C.P.; Chen, H.J.; Tang, S.L. Prevalence of potential nitrogen-fixing, green sulfur bacteria in the skeleton of reef-building coral *Isopora palifera* . *Limnol. Oceanogr.* **2016**, *61*, 1078–1086, DOI:10.1002/lno.10277.

64. Overmann, J.; Garcia-Pichel, F. The phototrophic way of life. In *The Prokaryotes: Prokaryotic Communities and Ecophysiology*; 2013; Vol. 9783642301, pp. 203–257 ISBN 9783642301230.

65. Koblížek, M. Ecology of aerobic anoxygenic phototrophs in aquatic environments. *FEMS Microbiol. Rev.* **2015**, *39*, 854–870, DOI:10.1093/femsre/fuv032.

66. Bryant, D.A.; Costas, A.M.G.; Maresca, J.A.; Chew, A.G.M.; Klatt, C.G.; Bateson, M.M.; Tallon, L.J.; Hostetler, J.; Nelson, W.C.; Heidelberg, J.F.; et al. *Candidatus* Chloracidobacterium thermophilum: An Aerobic Phototrophic Acidobacterium. *Science (80-.).* **2007**, *317*, 523–526, DOI:10.1126/science.1143236.

67. Ward, L.M.; McGlynn, S.E.; Fischera, W.W. Draft genome sequence of chloracidobacterium sp. CP2_5A, a phototrophic member of the phylum acidobacteria recovered from a Japanese hot spring. *Genome Announc.* **2017**, *5*, DOI:10.1128/genomeA.00821-17.

68. Zeng, Y.; Feng, F.; Medova, H.; Dean, J.; Kobli ek, M. Functional type 2 photosynthetic reaction centers found in the rare bacterial phylum Gemmatimonadetes. *Proc. Natl. Acad. Sci.* **2014**, *111*, 7795–7800, DOI:10.1073/pnas.1400295111.

69. Imhoff, J.F. The family Chlorobiaceae. In *The Prokaryotes: Other Major Lineages of Bacteria and The Archaea*; 2014; Vol. 9783642389, pp. 501–514 ISBN 9783642301230.

70. Imhoff, J.F. The family Chromatiaceae. In *The Prokaryotes: Gammaproteobacteria*; 2014; Vol. 9783642389, pp. 151–178 ISBN 9783642389221.

71. Oren, A. The family ectothiorhodospiraceae. *The Prokaryotes: Gammaproteobacteria* **2014**, *9783642389*, 199–222, DOI:10.1007/978-3-642-38922-1_248.

72. Nicholson, J.A.; Stolz, J.F.; Pierson, B.K. Structure of a microbial mat at Great Sippewisset Marsh, Cape Cod, Masschusetts. *Fems Microbiol. Ecol.* **1987**, *45*, 343–364.

73. Manske, A.K.; Glaeser, J.; Kuypers, M.M.M.; Overmann, J. Physiology and phylogeny of green sulfur bacteria forming a monospecific phototrophic assemblage at a depth of 100 meters in the Black Sea. *Appl. Environ. Microbiol.* **2005**, *71*, 8049–8060, DOI:10.1128/AEM.71.12.8049-8060.2005.

74. Findlay, A.J.; Bennett, A.J.; Hanson, T.E.; Luther, G.W. Light-dependent sulfide oxidation in the anoxic zone of the chesapeake bay can be explained by small populations of phototrophic bacteria. *Appl. Environ. Microbiol.* **2015**, *81*, 7560–7569, DOI:10.1128/AEM.02062-15.

75. Lauro, F.M.; DeMaere, M.Z.; Yau, S.; Brown, M. V; Ng, C.; Wilkins, D.; Raftery, M.J.; Gibson, J.A.; Andrews-Pfannkoch, C.; Lewis, M.; et al. An integrative study of a meromictic lake ecosystem in Antarctica. *ISME J.* **2011**, *5*, 879–895, DOI:10.1038/ismej.2010.185.

76. Hanada, S. The phylum Chloroflexi, the family Chloroflexaceae, and the related phototrophic families Oscillochloridaceae and Roseiflexaceae. In *The Prokaryotes: Other Major Lineages of Bacteria and The Archaea*; 2014; Vol. 9783642389, pp. 515–532 ISBN 9783642301230.

77. Madigan, M.T.; Jung, D.O. An Overview of Purple Bacteria: Systematics, Physiology, and Habitats. *Purple Phototrophic Bact.* **2009**, *28*, 1–15, DOI:10.1007/978-1-4020-8815-5.

78. Castanier, S.; Le Métayer-Levrel, G.; Perthuisot, J.P. Ca-carbonates precipitation and limestone genesis - the microbiogeologist point of view. *Sediment. Geol.* **1999**, *126*, 9–23, DOI:10.1016/S0037-0738(99)00028-7.

79. Bosak, T.; Greene, S.E.; Newman, D.K. A likely role for anoxygenic photosynthetic microbes in the formation of ancient stromatolites. *Geobiology* **2007**, *5*, 119–126, DOI:10.1111/j.1472-4669.2007.00104.x.

80. Bundeleva, I.A.; Shirokova, L.S.; Bénézeth, P.; Pokrovsky, O.S.; Kompantseva, E.I.; Balor, S. Calcium carbonate precipitation by anoxygenic phototrophic bacteria. *Chem. Geol.* **2012**, *291*, 116–131, DOI:10.1016/j.chemgeo.2011.10.003.

81. Li, R.Y.; Fang, H.H.P. Hydrogen production characteristics of photoheterotrophic *Rubrivivax gelatinosus* L31. *Int. J. Hydrogen Energy* **2008**, *33*, 974–980, DOI:10.1016/j.ijhydene.2007.12.001.

82. de Wit, R.; van Gemerden, H. Chemolithotrophic growth of the phototrophic sulfur bacterium Thiocapsa roseopersicina. *FEMS Microbiol. Lett.* **1987**, *45*, 117–162, DOI:10.1016/0378-1097(87)90033-4.

83. Le Campion-Alsumard, T. Étude Expérimentale De La Colonisation D'Éclats De Calcite Par Les Cyanophycées Endolithes Marines. *Cah. Biol. Mar.* **1975**, *16*, 177–185.

84. Le Campion-Alsumard, T. Les cyanophycées endolithes marines--Systématique, ultrastructure, écologie et biodestruction. *Ocean. Acta* **1979**, *2*, 143–156.

85. Tribollet, A.; Golubic, S. Cross-shelf differences in the pattern and pace of bioerosion of experimental carbonate substrates exposed for 3 years on the northern Great Barrier Reef, Australia. *Coral Reefs* **2005**, *24*, 422–434, DOI:10.1007/s00338-005-0003-7.

86. Vogel, K.; Gektidis, M.; Golubic, S.; Kiene, W.E.; Radtke, G. Experimental studies on microbial bioerosion at Lee Stocking Island, Bahamas and One Tree Island, Great Barrier Reef, Australia: Implications for paleoecological reconstructions. *Lethaia* **2000**, *33*, 190–204, DOI:10.1080/00241160025100053.

87. Garcia-Pichel, F.; Ramirez-Reinat, E.; Gao, Q. Microbial excavation of solid carbonates powered by P-type ATPase-mediated transcellular Ca2+ transport. *Proc. Natl. Acad. Sci.* **2010**, *107*, 21749–21754, DOI:10.1073/pnas.1011884108.

88. Guida, B.S.; Garcia-Pichel, F. Draft Genome Assembly of a Filamentous Euendolithic (True Boring) Cyanobacterium, Mastigocoleus testarum Strain BC008. *Genome Announc.* **2016**, *4*, 1–2, DOI:10.1128/genomeA.01574-15.Copyright.

89. Parker, C.T.; Tindall, B.J.; Garrity, G.M. International code of nomenclature of Prokaryotes. *Int. J. Syst. Evol. Microbiol.* **2019**, DOI:10.1099/ijsem.0.000778.

90. Turland, N.J.; Wiersema, J.H.; Barrie, F.R.; Greuter, W.; Hawksworth, D.L.; Herendeen, P.S.; Knapp, S.; Kusber, W.-H.; Li, D.-Z.; Marhold, K.; et al. *International Code of Nomenclature for algae, fungi, and plants (Shenzhen Code) adopted by the Nineteenth International Botanical Congress Shenzhen, China, July 2017*; 2018; ISBN 9783946583165.

91. Stoyanov, P.; Moten, D.; Mladenov, R.; Dzhambazov, B.; Teneva, I. Phylogenetic relationships of some filamentous cyanoprokaryotic species. *Evol. Bioinforma.* **2014**, *10*, 39–49, DOI:10.4137/EBo.s13748.

92. Boyer, S.L.; Johansen, J.R.; Flechtner, V.R.; Howard, G.L. Phylogeny and genetic variance in terrestrial Microcoleus (Cyanophyceae) species based on sequence analysis of the 16S rRNA gene and associated 16S-23S its region. *J. Phycol.* **2002**, *38*, 1222–1235, DOI:10.1046/j.1529-8817.2002.01168.x.

93. Soo, R.M.; Skennerton, C.T.; Sekiguchi, Y.; Imelfort, M.; Paech, S.J.; Dennis, P.G.; Steen, J.A.; Parks, D.H.; Tyson, G.W.; Hugenholtz, P. An expanded genomic representation of the phylum cyanobacteria. *Genome Biol. Evol.* **2014**, *6*, 1031–1045, DOI:10.1093/gbe/evu073.

94. Soo, R.M.; Hemp, J.; Hugenholtz, P. Evolution of photosynthesis and aerobic respiration in the cyanobacteria. *Free Radic. Biol. Med.* **2019**, *140*, 200–205, DOI:10.1016/j.freeradbiomed.2019.03.029.

95. Garcia-Pichel, F.; Zehr, J.P.; Bhattacharya, D.; Pakrasi, H.B. What's in a name? The case of cyanobacteria. *J. Phycol.* **2020**, *56*, 1–5, DOI:10.1111/jpy.12934.

96. Quast, C.; Pruesse, E.; Yilmaz, P.; Gerken, J.; Schweer, T.; Yarza, P.; Peplies, J.; Glöckner, F.O. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **2013**, *41*, 590–596, DOI:10.1093/nar/gks1219.

97. DeSantis, T.Z.; Hugenholtz, P.; Larsen, N.; Rojas, M.; Brodie, E.L.; Keller, K.; Huber, T.; Dalevi, D.; Hu, P.; Andersen, G.L. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **2006**, *72*, 5069–5072, DOI:10.1128/AEM.03006-05.

98. Edgar, R. Taxonomy annotation and guide tree errors in 16S rRNA databases. *PeerJ* **2018**, *2018*, DOI:10.7717/peerj.5030.

99. Park, S.-C.; Won, S. Evaluation of 16S rRNA Databases for Taxonomic Assignments Using a Mock Community. *Genomics Inform.* **2018**, *16*, e24, DOI:10.5808/gi.2018.16.4.e24.

100. Lydon, K.A.; Lipp, E.K. Taxonomic annotation errors incorrectly assign the family Pseudoalteromonadaceae to the order Vibrionales in Greengenes: Implications for microbial community assessments. *PeerJ* **2018**, *2018*, DOI:10.7717/peerj.5248.

101. Felsenstein, J. Evolutionary trees from DNA sequences: A maximum likelihood approach. *J. Mol. Evol.* **1981**, DOI:10.1007/BF01734359.

102. Matsen, F.A.; Kodner, R.B.; Armbrust, E.V. pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics* **2010**, DOI:10.1186/1471-2105-11-538.

103. Berger, S.A.; Krompass, D.; Stamatakis, A. Performance, accuracy, and web server for evolutionary placement of short sequence reads under maximum likelihood. *Syst. Biol.* **2011**, *60*, 291–302, DOI:10.1093/sysbio/syr010.

104. Barbera, P.; Kozlov, A.M.; Czech, L.; Morel, B.; Darriba, D.; Flouri, T.; Stamatakis, A. EPA-ng: Massively Parallel Evolutionary Placement of Genetic Sequences. *Syst. Biol.* **2019**, *68*, 365–369, DOI:10.1093/sysbio/syy054.

105. Golubić, S.; Le Campion-Alsumard, T. Boring behavior of marine blue-green algae *Mastigocoleus testarum* Lagerheim and *Kyrtuthrix dalmatica* Ercegović, as a taxonomic character. *Aquat. Sci.* **1973**, *35*, 157–161, DOI:10.1007/BF02502070.

106. Carreiro-Silva, M.; Kiene, W.E.; Golubic, S.; McClanahan, T.R. Phosphorus and nitrogen effects on microbial euendolithic communities and their bioerosion rates. *Mar. Pollut. Bull.* **2012**, *64*, 602–613, DOI:10.1016/j.marpolbul.2011.12.013.

107. Macedo, M.F.; Miller, A.Z.; Dionísio, A.; Saiz-Jimenez, C. Biodiversity of cyanobacteria and green algae on monuments in the Mediterranean Basin: An overview. *Microbiology* **2009**, *155*, 3476–3490, DOI:10.1099/mic.0.032508-0.

108. Godinot, C.; Tribollet, A.; Grover, R.; Ferrier-Pagès, C. Bioerosion by euendoliths decreases in phosphate-enriched skeletons of living corals. *Biogeosciences* **2012**, *9*, 2377–2384, DOI:10.5194/bg-9-2377-2012.

109. Reid, R.P.; Foster, J.S.; Radtke, G.; Golubic, S. Modern marine stromatolites of Little Darby Island, exuma archipelago, Bahamas: Environmental setting, accretion mechanisms and role of euendoliths. *Lect. Notes Earth Sci.* **2011**, *131*, 77–89, DOI:10.1007/978-3-642-10415-2.

110. Hutchings, P.A. Biological destruction of coral reefs. *Coral Reefs* 1986, *4*, 239–252.

111. Tribollet, A.; Veinott, G.; Golubic, S.; Dart, R. Infestation of the North American freshwater mussel Elliptio complanata (Head Lake, Canada) by the euendolithic cyanobacterium *Plectonema terebrans* Bornet et Flahault. *Arch. Hydrobiol. Suppl. Algol. Stud.* **2008**, *128*, 65–77, DOI:10.1127/1864-1318/2008/0128-0065.

112. Campbell, S.E. Precambrian endoliths discovered. *Nature* **1982**, *299*, 429–431, DOI:10.1038/299429a0.

113. Walker, J.J.; Pace, N.R. Endolithic Microbial Ecosystems. *Annu. Rev. Microbiol.* **2007**, *61*, 331–347, DOI:10.1146/annurev.micro.61.080706.093302.

114. Wade, B.; Garcia-Pichel, F. Evaluation of DNA Extraction Methods for Molecular Analyses of Microbial Communities in Modern Calcareous Microbialites. *Geomicrobiol. J.* **2003**, *20*, 549–561, DOI:10.1080/01490450390249460.

115. Schindelin, J.; Arganda-Carreras, I.; Frise, E.; Kaynig, V.; Longair, M.; Pietzsch, T.; Preibisch, S.; Rueden, C.; Saalfeld, S.; Schmid, B.; et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **2012**, *9*, 676–682, DOI:10.1038/nmeth.2019.

116. Muyzer, G.; De Waal, E.; Uitterlinden, A. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain Reaction-Amplified Genes Coding for 16S rRNA. *Appl. Enviromental Microbiol.* **1993**, *59*, 695–700.

117. Caporaso, J.G.; Lauber, C.L.; Walters, W.A.; Berg-Lyons, D.; Lozupone, C.A.; Turnbaugh, P.J.; Fierer, N.; Knight, R. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci.* **2011**, *108*, 4516–4522, DOI:10.1073/pnas.1000080107.

118. Caporaso, J.G.; Kuczynski, J.; Stombaugh, J.; Bittinger, K.; Bushman, F.D.; Costello, E.K.; Fierer, N.; Peña, A.G.; Goodrich, J.K.; Gordon, J.I.; et al. QIIME allows analysis of high- throughput community sequencing data. *Nat. Publ. Gr.* **2010**, *7*, 335–336, DOI:10.1038/nmeth0510-335.

119. Rognes, T.; Flouri, T.; Nichols, B.; Quince, C.; Mahé, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **2016**, *4*, e2584, DOI:10.7717/peerj.2584.

120. Kopylova, E.; Noé, L.; Touzet, H. SortMeRNA: Fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **2012**, *28*, 3211–3217, DOI:10.1093/bioinformatics/bts611.

121. Mercier, C.; Boyer, F.; Bonin, A.; Coissac, E. SUMATRA and SUMACLUST: fast and exact comparison and clustering of sequences Available online: http://metabarcoding.org/sumatra.

122. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780, DOI:10.1093/molbev/mst010.

123. Sela, I.; Ashkenazy, H.; Katoh, K.; Pupko, T. GUIDANCE2: Accurate detection of unreliable alignment regions accounting for the uncertainty of multiple parameters. *Nucleic Acids Res.* **2015**, *43*, W7–W14, DOI:10.1093/nar/gkv318.

124. Miller, M.A.; Pfeiffer, W.; Schwartz, T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *2010 Gatew. Comput. Environ. Work. GCE 2010* **2010**, DOI:10.1109/GCE.2010.5676129.

125. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313, DOI:10.1093/bioinformatics/btu033.

126. Berger, S.A.; Stamatakis, A. Aligning short reads to reference alignments and trees. *Bioinformatics* **2011**, *27*, 2068–2075, DOI:10.1093/bioinformatics/btr320.

127. Letunic, I.; Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **2016**, *44*, W242–W245, DOI:10.1093/nar/gkw290.

128. Frigaard, N.U.; Takaichi, S.; Hirota, M.; Shimada, K.; Matsuura, K. Quinones in chlorosomes of green sulfur bacteria and their role in the redox-dependent fluorescence studied in chlorosome-like bacteriochlorophyll *c* aggregates. *Arch. Microbiol.* **1997**, *167*, 343–349, DOI:10.1007/s002030050453.

129. Frigaard, N.U.; Larsen, K.L.; Cox, R.P. Spectrochromatography of photosynthetic pigments as a fingerprinting technique for microbial phototrophs. *FEMS Microbiol. Ecol.* **1996**, *20*, 69–77, DOI:10.1016/0168-6496(96)00005-0.

130. Hanada, S.; Takaichi, S.; Matsuura, K.; Nakamura, K. *Roseiflexus castenholzii* gen. nov., sp. nov., a thermophilic, filamentous, photosynthetic bacterium that lacks chlorosomes. *Int. J. Syst. Evol. Microbiol.* **2002**, *52*, 187–193, DOI:10.1099/00207713-52-1-187.

131. Koblízek, M.; Janouškovec, J.; Oborník, M.; Johnson, J.H.; Ferriera, S.; Falkowski, P.G. Genome sequence of the marine photoheterotrophic bacterium *Erythrobacter* sp. Strain NAP1. *J. Bacteriol.* **2011**, *193*, 5881–5882, DOI:10.1128/JB.05845-11.

132. Caple, M.B.; Chow, H.; Strouse, C.E. Photosynthetic pigments of green sulfur bacteria. The esterifying alcohols of bacteriochlorophylls *c* from *Chlorobium limicola* . *J. Biol. Chem.* **1978**, *253*, 6730–6737.

133. Abed, R.M.M.; Garcia-Pichel, F.; Hernández-Mariné, M. Polyphasic characterization of benthic, moderately halophilic, moderately thermophilic cyanobacteria with very thin trichomes and the proposal of *Halomicronema excentricum* gen. nov., sp. nov. *Arch. Microbiol.* **2002**, *177*, 361–370, DOI:10.1007/s00203-001-0390-2.

134. Pierson, B.K.; Valdez, D.; Larsen, M.; Morgan, E.; Mack, E.E. Chloroflexus-like organisms from marine and hypersaline environments: Distribution and diversity. *Photosynth. Res.* **1994**, *41*, 35–52, DOI:10.1007/BF02184144.

135. Blankenship, R.E. Reaction centers and electron transport pathways in anoxygenic phototrophs. In *Molecular Mechanisms of Photosynthesis*; 2008; pp. 89–110 ISBN 9780470758472.

136. Tonon, L.A.C.; Moreira, A.P.B.; Thompson, F. The family Erythrobacteraceae. In *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria*; 2014; Vol. 9783642301, pp. 213–235 ISBN 9783642301971.

137. Dubinina, G.A.; Gorlenko, V.M. [New filamentous photosynthesizing green bacteria with gas vacuoles]. *Mikrobiologiia* **1975**, *44*, 511–517.

138. Gorlenko, V.M. A new phototrophic green sulphur bacterium–*Prosthecochloris aestuarii* nov. gen. nov. spec. *Z. Allg. Mikrobiol.* **1970**, *10*, 147–149, DOI:10.1002/jobm.19700100207.

139. Garcia-Pichel, F.; Loza, V.; Marusenko, Y.; Mateo, P.; Potrafka, R.M. Temperature Drives the Continental-Scale Distribution of Key Microbes in Topsoil Communities. *Science (80-.).* **2013**, *340*, 1574–1577, DOI:10.1126/science.1236404.

140. Reynolds, R.W.; Smith, T.M.; Liu, C.; Chelton, D.B.; Casey, K.S.; Schlax, M.G. Daily high-resolution-blended analyses for sea surface temperature. *J. Clim.* **2007**, *20*, 5473–5496, DOI:10.1175/2007JCLI1824.1.

141. Karr, E.A.; Sattley, W.M.; Jung, D.O.; Madigan, M.T.; Achenbach, L.A. Remarkable diversity of phototrophic purple bacteria in a permanently frozen Antarctic lake. *Appl. Environ. Microbiol.* **2003**, *69*, 4910–4914, DOI:10.1128/AEM.69.8.4910-4914.2003.

142. Madigan, M.T.; Jung, D.O.; Karr, E. a; Sattley, W.M.; Achenbach, L. a.; van der Meer, M.T.J. Diversity of Anoxygenic Phototrophs in Contrasting Extreme Environments. *Geotherm. Biol. Geochemistry YNP* **2005**, *1*, 203–220.

143. De Philippis, R.; Vincenzini, M. Exocellular polysccharides from cyanobacteria and their possible applications. *FEMS Microbiol. Rev.* **1998**, *22*, 151–175, DOI:10.1111/j.1574-6976.1998.tb00365.x.

144. Gich, F.; Garcia-Gil, J.; Overmann, J. Previously unknown and phylogenetically diverse members of the green nonsulfur bacteria are indigenous to freshwater lakes. *Arch. Microbiol.* **2002**, *177*, 1–10, DOI:10.1007/s00203-001-0354-6.

145. Lami, R.; Cottrell, M.T.; Ras, J.; Ulloa, O.; Obernosterer, I.; Claustre, H.; Kirchman, D.L.; Lebaron, P. High abundances of aerobic anoxygenic photosynthetic bacteria in the South Pacific Ocean. *Appl. Environ. Microbiol.* **2007**, *73*, 4198–4205, DOI:10.1128/AEM.02652-06.

146. Garcia-Pichel, F.; Belnap, J.; Neuer, S.; Schanz, F. Estimates of global cyanobacterial biomass and its distribution. *Arch. Hydrobiol. Suppl. Algol. Stud.* **2003**, *109*, 213–227, DOI:10.1127/1864-1318/2003/0109-0213.

147. Klappenbach, J.A.; Pierson, B.K. Phylogenetic and physiological characterization of a filamentous anoxygenic photoautotrophic bacterium “Candidatus Chlorothrix halophila” gen. nov., sp. nov., recovered from hypersaline microbial mats. *Arch. Microbiol.* **2004**, *181*, 17–25, DOI:10.1007/s00203-003-0615-7.

148. Koblížek, M.; Béjà, O.; Bidigare, R.R.; Christensen, S.; Benitez-Nelson, B.; Vetriani, C.; Kolber, M.K.; Falkowski, P.G.; Kolber, Z.S. Isolation and characterization of Erythrobacter sp. strains from the upper ocean. *Arch. Microbiol.* **2003**, *180*, 327–338, DOI:10.1007/s00203-003-0596-6.

149. Couradeau, E.; Karaoz, U.; Lim, H.C.; Nunes da Rocha, U.; Northen, T.; Brodie, E.; Garcia-Pichel, F. Bacteria increase arid-land soil surface temperature through the production of sunscreens. *Nat. Commun.* **2016**, *7*, 1–7, DOI:10.1038/ncomms10373.

150. Bolyen, E.; Rideout, J.R.; Dillon, M.R.; Bokulich, N.A.; Abnet, C.C.; Al-Ghalith, G.A.; Alexander, H.; Alm, E.J.; Arumugam, M.; Asnicar, F.; et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **2019**, *37*, 852–857, DOI:10.1038/s41587-019-0209-9.

151. Callahan, B.J.; McMurdie, P.J.; Rosen, M.J.; Han, A.W.; Johnson, A.J.A.; Holmes, S.P. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **2016**, *13*, 581–583, DOI:10.1038/nmeth.3869.

152. Price, M.N.; Dehal, P.S.; Arkin, A.P. FastTree 2 - Approximately maximum-likelihood trees for large alignments. *PLoS One* **2010**, *5*, DOI:10.1371/journal.pone.0009490.

153. Lozupone, C.; Knight, R. UniFrac : a New Phylogenetic Method for Comparing Microbial Communities UniFrac : a New Phylogenetic Method for Comparing Microbial Communities. *Appl. Environ. Microbiol.* **2005**, *71*, 8228–8235, DOI:10.1128/AEM.71.12.8228.

154. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 1–21, DOI:10.1186/s13059-014-0550-8.

155. Oksanen, A.J.; Blanchet, F.G.; Friendly, M.; Kindt, R.; Legendre, P.; Mcglinn, D.; Minchin, P.R.; Hara, R.B.O.; Simpson, G.L.; Solymos, P.; et al. Vegan: Community ecology package. *https://github.com/vegandevs/vegan* **2018**, DOI:10.4135/9781412971874.n145.

156. R Development Core Team R: A Language and Environment for Statistical Computing. *R Found. Stat. Comput.* **2011**, DOI:10.1007/978-3-540-74686-7.

157. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*; Springer-Verlag New York, 2016; ISBN 978-0-387-98140-6.

158. Roush, D.; Giraldo-Silva, A.; Fernandes, V.M.C.; Maria Machado de Lima, N.; McClintock, S.; Velasco Ayuso, S.; Klicki, K.; Dirks, B.; Arantes Gama, W.; Sorochkina, K.; et al. Cydrasil: A comprehensive phylogenetic tree of cyanobacterial 16s rRNA gene sequences Available online: https://github.com/FGPLab/cydrasil.

159. Nawrocki, E. Structural RNA Homology Search and Alignment Using Covariance Models, Washington University School of Medicine, 2009.

160. Madden, T.L.; Camacho, C.; Ma, N.; Coulouris, G.; Avagyan, V.; Bealer, K.; Papadopoulos, J. BLAST+: architecture and applications. *BMC Bioinformatics* **2009**, *10*, 421, DOI:10.1186/1471-2105-10-421.

161. Ley, R.E.; Harris, J.K.; Wilcox, J.; Spear, J.R.; Miller, S.R.; Bebout, B.M.; Maresca, J.A.; Bryant, D.A.; Sogin, M.L.; Pace, N.R. Unexpected Diversity and Complexity of the Guerrero Negro Hypersaline Microbial Mat Unexpected Diversity and Complexity of the Guerrero Negro Hypersaline Microbial Mat. *Appl. Environ. Microbiol.* **2006**, *72*, 3685–3695, DOI:10.1128/AEM.72.5.3685.

162. Golecki, J.R.; Oelze, J. Quantitative relationship between bacteriochlorophyll content, cytoplasmic membrane structure and chlorosome size in *Chloroflexus aurantiacus*. *Arch. Microbiol.* **1987**, *148*, pp 236-241.

163. Komarek, J.; Hindak, F. Taxonomy of the new isolated strains of Chroococcidiopsis (Cyanophyceae). *Arch. fur Hydrobiol. (Suppl. 46, Algol. Stud.* **1975**, *13*, 311–329.

164. Lee, J.Z.; Burow, L.C.; Woebken, D.; Craig Everroad, R.; Kubo, M.D.; Spormann, A.M.; Weber, P.K.; Pett-Ridge, J.; Bebout, B.M.; Hoehler, T.M. Fermentation couples Chloroflexi and sulfate-reducing bacteria to Cyanobacteria in hypersaline microbial mats. *Front. Microbiol.* **2014**, *5*, 1–17, DOI:10.3389/fmicb.2014.00061.

165. Shiba, T.; Simidu, U. Erythrobacter longus gen. nov., sp. nov., an aerobic bacterium which contains bacteriochlorophyll a. *Int. J. Syst. Bacteriol.* **1982**, *32*, 211–217, DOI:10.1099/00207713-32-2-211.

166. Zheng, Q.; Lin, W.; Liu, Y.; Chen, C.; Jiao, N. A comparison of 14 Erythrobacter genomes provides insights into the genomic divergence and scattered distribution of phototrophs. *Front. Microbiol.* **2016**, *7*, DOI:10.3389/fmicb.2016.00984.

167. Garcia-Pichel, F. Plausible mechanisms for the boring on carbonates by microbial phototrophs. *Sediment. Geol.* **2006**, *185*, 205–213, DOI:10.1016/j.sedgeo.2005.12.013.

168. Nübel, U.; Garcia-Pichel, F.; Kühl, M.; Muyzer, G. Spatial scale and the diversity of benthic cyanobacteria and diatoms in a salina. In *Molecular Ecology of Aquatic Communities*; Zehr, J.P., Voytek, M.A., Eds.; Springer Netherlands: Dordrecht, 1999; pp. 199–206 ISBN 978-94-011-4201-4.

169. Whitton, B.A.; Potts, M. Introduction to the cyanobacteria. In *Ecology of Cyanobacteria II: Their Diversity in Space and Time*; Whitton, B.A., Ed.; Springer Netherlands, 2012; pp. 1–13 ISBN 9789400738553.

170. Soule, T.; Garcia-Pichel, F. Cyanobacteria. In *Encyclopedia of Microbiology*; Schmidt, T.M., Ed.; Elsevier Inc., 2019; pp. 799–817 ISBN 9780128117378.

171. Karl, D.; Michaels, A.; Bergman, B.; Capone, D.; Carpenter, E.; Letelier, R.; Lipschultz, F.; Paerl, H.; Sigman, D.; Stal, L. Dinitrogen fixation in the world's oceans. In *The Nitrogen Cycle at Regional to Global Scales*; Boyer, E.W., Howarth, R.W., Eds.; Springer, Dordrecht, 2002; pp. 47–98.

172. Schloss, P.D.; Westcott, S.L.; Ryabin, T.; Hall, J.R.; Hartmann, M.; Hollister, E.B.; Lesniewski, R.A.; Oakley, B.B.; Parks, D.H.; Robinson, C.J.; et al. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* **2009**, *75*, 7537–7541, DOI:10.1128/AEM.01541-09.

173. Nübel, U.; Garcia-Pichel, F.; Muyzer, G. PCR primers to amplify 16S rRNA genes from cyanobacteria. *Appl. Environ. Microbiol.* **1997**, *63*, 3327–3332, DOI:10.1128/aem.63.8.3327-3332.1997.

174. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647–1649, DOI:10.1093/bioinformatics/bts199.

175. Rippka, R.; Deruelles, J.; Waterbury, J.B. Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *J. Gen. Microbiol.* **1979**, *111*, 1–61, DOI:10.1099/00221287-111-1-1.

176. Berger, S.A.; Stamatakis, A. PaPaRa 2.0: a vectorized algorithm for probabilistic phylogeny-aware alignment extension. *... Inst. Theor. Stud. http//sco. h ...* **2012**.

APPENDIX A

EXPANDED CYDRASIL PHYLOGENETIC RECONSTRUCTION

[Consult Attached Files]

APPENDIX B

CYDRASIL REFERENCE PACKAGE

[Consult Attached Files]

[Also available at https://github.com/FGPLab/cydrasil]

APPENDIX C

CYDRASIL WEB APPLICATION SOURCE CODE

[Consult Attached Files]

APPENDIX D

POLAR AND TERRESTRIAL ENDOLITHIC PHOTOTROPH DIVERSITY

Figure 1. Phototrophic composition of dolomite endoliths based on 16S rRNA gene sequence reads from Churchill, Manitoba Canada. Sample identifiers are on the very left. (a) Distribution of major phototrophic groups showing that all samples contained both oxygenic and anoxygenic phototrophs. (b) Ratio of phototrophic prokaryotes to non-phototrophic prokaryotes. Churchill dolomites tended to have very low phototrophic sequence abundance compared to temperate and tropical endolithic communities. (c) Phylogenetic distribution of 16S rRNA gene sequences reads attributed to cyanobacteria.

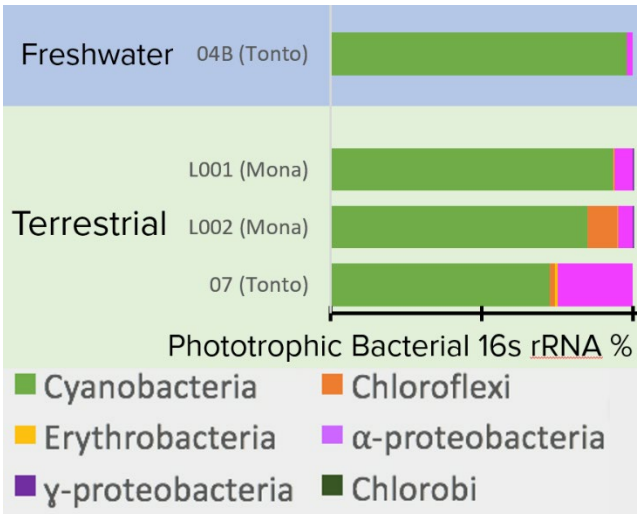

Figure 2. Phototrophic composition of dolomite endoliths based on 16S rRNA gene sequence reads from freshwater and terrestrial environments.