J. Phycol. **51,** 682–694 (2015)

DOI: 10.1111/jpy.12310

ANALYSIS OF *rbc*L SEQUENCES REVEALS THE GLOBAL BIODIVERSITY, COMMUNITY STRUCTURE, AND BIOGEOGRAPHICAL PATTERN OF THERMOACIDOPHILIC RED ALGAE (CYANIDIALES)[1]

Chia-Jung Hsieh

Department of Life Science, Tunghai University, Taichung 40704, Taiwan

Shing Hei Zhan

Department of Zoology, University of British Columbia, Vancouver, BC, Canada V6T 1Z4

Yiching Lin

Department of Life Science, Tunghai University, Taichung 40704, Taiwan

Sen-Lin Tang

Biodiversity Research Center, Academia Sinica, Taipei 11529, Taiwan

and Shao-Lun Liu[2]

Department of Life Science, Tunghai University, Taichung 40704, Taiwan

Thermoacidophilic cyanidia (Cyanidiales) are the primary photosynthetic eukaryotes in volcanic areas. These red algae also serve as important model organisms for studying life in extreme habitats. The global biodiversity and community structure of Cyanidiales remain unclear despite previous sampling efforts. Here, we surveyed the Cyanidiales biodiversity in the Tatun Volcano Group (TVG) area in Taiwan using environmental DNA sequencing. We generated 174 *rbc*L sequences from eight samples from four regions in the TVG area, and combined them with 239 publicly available *rbc*L sequences collected worldwide. Species delimitation using this large *rbc*L data set suggested at least 20 Cyanidiales OTUs (operational taxonomic units) worldwide, almost three times the presently recognized seven species. Results from environmental DNA showed that OTUs in the TVG area were divided into three groups: (i) dominant in hot springs with 92%–99% sequence identity to *Galdieria maxima*; (ii) largely distributed in drier and more acidic microhabitats with 99% identity to *G. partita*; and (iii) primarily distributed in cooler microhabitats and lacking identity to known cyanidia species (a novel Cyanidiales lineage). In both global and individual area analyses, we observed greater species diversity in non-aquatic than aquatic habitats. Community structure analysis showed high similarity between the TVG community and West Pacific-Iceland communities, reflecting their geographic proximity to each other. Our study is the first examination of the global species diversity and biogeographic affinity of cyanidia. Additionally, our data illuminate the influence of microhabitat type on Cyanidiales diversity and highlight intriguing questions for future ecological research.

Key index words: **biodiversity; biogeography; community; cyanidia; Cyanidiales; microhabitats; *rbc*L**

***Abbreviations*: ABGD, automated barcode gap discovery; Cc, *Cyanidium caldarium*; Cm, *Cyanidioschyzon merolae*; Csp, *Cyanidium* sp.; Gd, *Galdieria daedala*; Gm, *Galdieria maxima*; GMYC, generalized mixed Yule-coalescent model; Gp, *Galdieria partita*; Gph, *Galdieria phlegrea*; Gs, *Galdieria sulphuraria*; ML, maximum likelihood; PTP, Poisson Tree Process model; *rbc*L, large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase; SPN, statistical parsimony network**

Extremophiles are organisms that tolerate and grow optimally in environments deemed unsuitable for human habitation (MacElroy 1974). Most, if not all, extremophiles are bacteria and archaea (Brock 1967). Volcanic areas are common sites for discovering extremophiles. Active magma chambers found in volcanic areas exude intense heat and highly acidic sulfur fumes, creating habitats that only extremophilic (or more specifically, thermoacidophilic) organisms can withstand. Thermoacidophilic red algae belonging to Cyanidiales (also referred to as cyanidia; Lehr et al. 2007) are the primary photosynthetic eukaryotes found in these volcanic areas (e.g., Walker et al. 2005). Due to their

[1]Received 1 December 2014. Accepted 3 May 2015.
[2]Author for correspondence: e-mail shaolunliu@thu.edu.tw.
Editorial Responsibility: M. Vis (Associate Editor)

extremophilic lifestyle and the simplicity of their cellular structures, they are popular model organisms for understanding life in hostile habitats (Barbier et al. 2005) and for studying organelle biogenesis in eukaryotes (Misumi et al. 2005).

To date, only about 20 species of microalgae from various major algal groups (e.g., Cyanobacteria, Chlorophyta, Chrysophyta, Euglenophyta, Heterokontophyta, and Rhodophyta) are known to inhabit highly acidic environments (reviewed in Gimmler 2001). Among them, cyanidia are the only algae that can thrive at high temperatures (e.g., Ciniglia et al. 2004, Pinto et al. 2007, Toplin et al. 2008). When the first cyanidia species were discovered, they were identified as green microalgae or cyanobacteria (*Coccochloris orsiniana* Meneghini, *Chroococcus varius* Tilden, *Protococcus botryoides* f. *caldarium* Tilden, *Palmellococcus thermalis* West, and *Pleurocapsa caldaria* Collins et al.; Guiry and Guiry 2014). Later, Geitler (1933) reclassified them as red algae and described the first species, *Cyanidium caldarium* (Cc; Tilden) Geitler. Based on morphological, physiological, and molecular evidence, there are three genera (*Cyanidium*, *Cyanidioschyzon*, and *Galdieria*) and seven species in Cyanidiales: Cc, *Cyanidioschyzon merolae* (Cm) De Luca et al., *Galdieria maxima* (Gm) Sentsova, *G. partita* (Gp) Sentsova, *G. daedala* (Gd) Sentsova, *G. phlegrea* (Gph) Pinto et al., and *G. sulphuraria* (Gs; Galdieri) Merola (see Pinto et al. 2007 for a taxonomic summary). Some species grow at lower temperatures and non-acidic environments; they are collectively referred to as mesophilic Cyanidiales (e.g., Ciniglia et al. 2004). Due to insufficient morphological and physiological data, all mesophilic cyanidia have been tentatively classified as members of the genus *Cyanidium* without further taxonomic treatment (e.g., Ciniglia et al. 2004, Toplin et al. 2008).

Cyanidia are mainly distributed in volcanic areas around the globe (e.g., Schwabe 1936, Gross and Oesterhelt 1999, Gross et al. 2001, Ciniglia et al. 2004, 2014, Kondo et al. 2004, Yoon et al. 2006, Toplin et al. 2008, Azúa-Bustos et al. 2009, Castenholz and McDermott 2010, Skorupa et al. 2013). Two sampling approaches have been used to survey the species diversity of Cyanidiales in volcanic areas. The first approach is to isolate algae in the laboratory and then determine species identity using conventional molecular biology techniques (i.e., culture-based methods; e.g., Gross et al. 2001, Kondo et al. 2004). The other approach is to assess Cyanidiales species diversity directly from environmental DNA (i.e., eDNA methods; e.g., Ciniglia et al. 2004). As of 2004, only a few large-scale surveys of Cyanidiales species diversity have been conducted. Ciniglia et al. (2004) and Yoon et al. (2006) examined Cyanidiales species diversity in Italy using the eDNA approach. These researchers discovered much greater Cyanidiales species richness than expected. In another study, Toplin et al. (2008) focused on samples from the Yellowstone National Park (YNP) in the United States, Japan, and New Zealand using a culture-based method and sequencing the *rbc*L and 18S rDNA loci. These authors uncovered high genetic diversity in *rbc*L but low genetic diversity in 18S rDNA, indicating that 18S rDNA is not suitable for surveying Cyanidiales species diversity. Recently, Skorupa et al. (2013) reexamined the species diversity of Cyanidiales in YNP using an eDNA method targeting *rbc*L, and showed that non-aquatic habitats harbored more cyanidia species than aquatic habitats. The other recent study investigating Cyanidiales species diversity was conducted in Iceland using a culture-based method (Ciniglia et al. 2014). These studies all suggest that the species diversity of Cyanidiales may be presently underappreciated. Indeed, the species diversity of these red algae in many volcanic areas around the globe—except in Italy and YNP—has only been skimmed using culture-based methods, and needs to be more thoroughly evaluated using the eDNA approach.

The Tatun Volcano Group (TVG) area is situated in the northern border of the Taipei basin in northern Taiwan. Active magma chambers in the TVG area emit sulfurous gases, making the entire area and its surroundings acidic (Liu et al. 2011). Therefore, the TVG area is an ideal spot for thermoacidophilic cyanidia to grow. Although two species, Cc and Gs, have been previously identified in the TVG area (Gross et al. 2001, Kondo et al. 2004), no systematic survey of Cyanidiales species diversity has yet been performed in this area. An environmental survey of cyanidia in the TVG area would provide new species diversity data for these algae in the West Pacific region, and thereby contribute significantly to the broader effort to document Cyanidiales biodiversity worldwide.

In this study, we characterized the global species diversity of Cyanidiales and examined the structure and biogeographical affinity of cyanidia communities from various volcanic areas around the world. We performed an eDNA survey of Cyanidiales species diversity in the TVG area in Taiwan, and combined our *rbc*L sequences with those gathered from various habitats worldwide in previous studies. Using this large comprehensive *rbc*L data set, we derived a novel estimate of the species richness of Cyanidiales, and inferred the phylogenetic relationships among known and potentially new species. In addition, we investigated whether non-aquatic and aquatic habitats affect the species abundance and composition of cyanidia communities. Finally, we compared the cyanidia community in the TVG area with communities from other volcanic areas around the world to better understand the biogeography and community similarities among locations. To the best of our knowledge, our study is the very first to explore the biodiversity, community structure, and biogeographical pattern of Cyanidiales on a global scale.

MATERIALS AND METHODS

Sample collection, measurements of environmental variables, and cultivation. Eight algal samples were collected from four different volcanic regions (GengZiPeng [25°11′ 16″ N, 121°36′ 50″ E], DaYouKeng [25°10′ 21″ N, 121°34′ 48″ E], MaChao [25°10′ 41″ N, 121°33′ 52″ E], and DiReGu [25°08′ 17″ N, 121°30′ 41″ E]) in the Tatun Volcanic Group (TVG) area, YangMingShan National Park, Taiwan. The samples were collected from non-aquatic and aquatic habitats. The non-aquatic samples—from endolithic and epilithic microhabitats, sulfur fumes, and soils (see Fig. S1 in the Supporting Information)—were collected using a hammer with a chisel (endolithic or epilithic samples) or a shovel (sulfur fumes or soil samples). These ~10 g samples were stored in 50 mL centrifuge tubes. The aquatic samples—from pools and streams (see Fig. S1)—were collected using a toothbrush, and then stored in 50 mL centrifuge tubes. During transportation, the samples were preserved in a thermostat bag. Once in the laboratory, a portion of each sample (~5 g) was treated with liquid nitrogen and preserved in a −80°C freezer for molecular analysis, and the remaining sample was used for subsequent isolation and cultivation. Temperature and pH were measured in situ using a digital microsensor thermometer and a pH meter, respectively. After collection, algal samples were immediately grown on 1.5% agar plates using modified Allen's medium (Minoda et al. 2004). The pH of the medium was adjusted to 2.5 using 1 N H_2SO_4. Pure isolates were subsequently cultivated in 50 mL of modified Allen's medium solution in the flask tube. The culture was maintained on a shaker at 45°C with a photoperiod of 12 h of light and 12 h of darkness.

DNA extraction, gene amplification, and sequencing. Total DNA was extracted using the PowerSoil™ DNA Isolation Kit (Mo Bio, Watson Biotechnology Co. Ltd., New Taipei City, Taiwan) following the manufacturer's instructions. After extraction, the total DNA pellet was dissolved in 50 μL ddH_2O. The quantity and quality of total DNA were examined using a Maestro Nanodrop Spectrophotometer (Green BioResearch Corp., Hsinchu, Taiwan), and were also checked by running the samples on a 1.5% agarose gel using gel electrophoresis. The resulting DNA sample was stored at −20°C. We used the same *rbc*L-specific primers as in Toplin et al. (2008): RbcL1F, 5′-AACCT TTCATGCGTTGGAGAGA-3′ (forward primer) and RbcL1R, 5′-CCTGCATGAATACCACCAGAAGC-3′ (reverse primer). Polymerase chain reaction (PCR) was conducted using a Bio-Rad C1000 Touch™ PCR thermal cycler (Bio-Rad Laboratories, Inc., Taipei, Taiwan). Briefly, PCR was performed on a mixture of 10 μL chemical reaction consisting of 0.25 μL of 5U Super-Run *Taq* (Protech Technology, Taipei, Taiwan), which is a high-fidelity thermostable DNA polymerase, 5 μL of 10× Super-Run *Taq* buffer, 4 μL of 2.5 mM dNTP mixture, 0.1 μL of each 10 mM primer, and 1 μL of DNA (100 ng · μL^{-1}). The PCR conditions were 94°C for 4 min, 34 cycles of 94°C for 30 s, 54°C for 30 s, and 72°C for 1 min, and 72°C for 6 min. The PCR products were visualized on a 1.5% agarose gel using gel electrophoresis. For sequencing, the PCR products were first cloned to TA plasmid vectors using the RBC-TA cloning kit (RBC Bioscience Corp., Taipei, Taiwan) following the manufacturer's recommendations. Then, the TA clone plasmids were transformed into DH5α competent cells (Protech Technology) according to the manufacturer's instructions. Successfully transformed clones were sent to Mission Biotech (Taipei, Taiwan) for sequencing.

Phylogenetic analysis. A total of 174 novel *rbc*L sequences were produced for this study. These sequences were deposited in NCBI GenBank (www.ncbi.nlm.nih.gov; accessions provided in Tables S1 and S2 in the Supporting Information). We integrated our new data with 239 available *rbc*L sequences from GenBank (accessions listed in Tables S1 and S2) for a final data set of 413 *rbc*L sequences. The *rbc*L sequences were aligned using MUSCLE (Edgar 2004) along with 11 *rbc*L sequence accessions from non-Cyanidiales red algae as the outgroup taxa (Table S3 in the Supporting Information). Two phylogenetic analyses were performed to obtain (i) a non-ultrametric maximum likelihood (ML) tree and (ii) a maximum clade credibility (MCC) tree in a set of ultrametric Bayesian trees. The ML tree was constructed using MEGA v.6 (Tamura et al. 2013). The GTR+G+I model was identified as the best-fitting nucleotide substitution model. Statistical support was obtained from 1,000 bootstrap replicates (Felsenstein 1985). To sample clock-like Bayesian phylogenies, we employed MrBayes v3.2.2 (Huelsenbeck and Ronquist 2001). Nucleotides were assumed to evolve under the GTR+G+I model, and branch lengths under the Independent Gamma Rate (IGR) relaxed clock model. Two runs, with four chains each, were set off for 40 million generations with a sampling frequency of once per 2,000 steps; 90% of the samples were discarded as burn-in. The MCC tree was identified among the MrBayes trees using TreeAnnotator v.1.7.5 (Drummond and Rambaut 2007). The ML tree and MCC tree were used in subsequent analyses.

Species delimitation. To identify operational taxonomic units (OTUs, which are interpreted as hypothetical species in this study), two coalescent-based methods were applied to the *rbc*L phylogenies: (i) the generalized mixed Yule-coalescent (GMYC) method (Pons et al. 2006) and (ii) the Poisson Tree Processes (PTP) method (Zhang et al. 2013). Both methods determine species boundaries (threshold point) by inferring shifts in branching rates in a given phylogenetic tree from the interspecific rate variation (speciation-level model) to intraspecific (population-level model) rate variation. The GMYC model describes branching events using information from time-scaled branch lengths (as in our MCC tree). We analyzed the MCC tree using the GMYC method implemented in the SPLITS v1.0-18 software using the "single threshold" model (Pons et al. 2006). The PTP model utilizes either branch length scaled to time or to number of substitutions (as in our ML tree) using the PTP v1.2 software (Zhang et al. 2013).

GMYC and PTP are fundamentally similar because both these models share an underlying evolutionary model (i.e., the coalescent). Hence, we confirmed the OTU number estimates using orthogonal techniques. We applied two additional methods: (i) a statistical parsimony network (SPN) approach and (ii) an automated barcode gap discovery (ABGD) analysis (Puillandre et al. 2012). SPN defines the number of OTUs based on a haplotype network (Morrison 2005) derived using a maximum parsimony method (Templeton et al. 1992). Using the SPN method implemented in TCS v1.21 (Clement et al. 2000), we estimated the OTU number and reported it with a 95% confidence interval (see example in Hart and Sunday 2007). The ABGD analysis was performed on the abgd website (wwwabi.snv.jussieu.fr/public/abgd/abgdweb.html, date accessed: Nov. 26, 2014) with the relative gap width (the "X" parameter) equal to 1 and the number of steps equal to 100. The pair-wise distance matrix was estimated under a T92+G model in MEGA v.6 (Tamura et al. 2013). ABGD detects putative species based on a series of prior thresholds (ranging from 0.001 to 0.1) for a gap in the pair-wise distribution of genetic distance. The gap could be considered as the threshold of the upper limit of intraspecific distances and the lower limit of interspecific distances. We only considered the outcome of the ABGD initial partition in this study (Puillandre et al. 2012). We compared the estimates of OTU number obtained from the four methods, and took the most conservative one (i.e., the lowest count).

Diversification analysis. To test whether the rate of diversification altered in the evolutionary history of Cyanidiales, we

conducted a model-based analysis using Parametric Rates Comparison (PRC; Shah et al. 2013). Here, we used only the MCC tree, and assumed that the branch lengths were distributed exponentially (the default setting). Before the analysis, we randomly selected one representative per OTU group, and then pruned out all the other members. Also, the MCC tree was pruned four separate times, once based on the OTU membership assignments by each of the four species delimitation methods. Consequently, PRC was conducted for each of the resulting pruned phylogenies (containing 20, 21, 22, or 29 OTUs). Applying a *P*-value threshold of 0.05, we identified clades that might have diversified at significantly higher (or lower) rates than the remainder of the phylogeny. The PRC test was performed using the package iterates v3.1 (Fordyce et al. 2014) in the R environment (R Core Team 2014; www.r-project.org/).

Diversity analysis among microhabitats or habitats. To assess Cyanidiales biodiversity among microhabitats or habitats, we calculated the Shannon-Wiener (H) index and the phylogenetic diversity (PD) index using functions implemented in the R packages vegan v2.0-10 (Oksanen 2009) and picante v1.6-2 (Kembel et al. 2010), respectively. The H index describes both the richness and evenness of OTUs that are present in a community (Hill 1973); a higher H index indicates greater species diversity and/or more uniform species relative abundance. The PD index considers not only OTU richness, but also the phylogenetic relatedness among OTUs within a community (Faith 1992). Communities in which taxa are more closely clustered on the phylogeny have lower PD indices (i.e., phylogenetic clustering) and vice versa (i.e., phylogenetic evenness) (e.g., Horner-Devine and Bohannan 2006).

Community similarity analysis among geothermal areas. To reveal similarity among cyanidia communities from various geographic locations, a cluster analysis was performed using picante. The picante method accounts for the phylogenetic relatedness among OTUs within each community or between two communities by measuring the mean pair-wise phylogenetic distance (MPD; Webb et al. 2008). The pair-wise phylogenetic relatedness among OTUs within each community (i.e., pair-wise branch length) was first calculated using the function *cophenetic*. Then, MPD was calculated for every pair of communities using the function *comdist*. Here, two data sets were used, one containing data only from the eDNA method (namely, the eDNA data set) and the other having combined data from both the eDNA and culture-based methods (namely, the combined data set). The eDNA data set contains relative abundance data, whereas the combined data set have only presence or absence data. To characterize the relative contribution of cyanidia OTUs to community similarity among geothermal areas, a principle component analysis (PCA) was conducted using the function *princomp* in vegan.

RESULTS

Species delimitation. Using the eDNA approach, we obtained 166 clones from eight samples from four geothermal areas: (i) GengZiPing (GZP; $n = 71$), (ii) DaYouKeng (DYK; $n = 57$), (iii) DiReGu (DRG; $n = 19$), and (iv) MaChao (MC; $n = 19$) in the TVG area of Taiwan (Table S1). Eight additional sequences were obtained from GZP and DRG using a culture-based method (Table S2), thereby resulting in a total of 174 *rbc*L sequences from Taiwan. Another 239 *rbc*L GenBank sequences from geothermal areas worldwide (Indonesia, Japan, New Zealand, Iceland, USA, Mexico and Italy) were combined with our data. The MUSCLE alignment of 424 *rbc*L sequences (including 11 outgroup sequences; 480 aligned sites with no gaps) was used to determine species diversity using four different statistical methods. Two coalescent-based methods, GMYC and PTP, produced similar OTU number estimates—22 and 20, respectively (Table 1; Fig. S2 in the Supporting Information). In the GMYC analysis, the null model (a pure coalescent process) was rejected in favor of the GMYC model ($\ln L_{GMYC} = 4317.477 > \ln L_0 = 4308.642$, $P < 0.001$; Table 1). Similarly, in the PTP analysis, the PTP model was supported ($\ln L_{PTP} = 711.635 > \ln L_0 = 543.469$, $P < 0.001$). The OTU number estimates were further corroborated by a complementary analysis using the SPN method (21 OTUs) and the ABGD method (29 OTUs; Fig. S2).

Using the most liberal estimate of global OTUs of Cyanidiales, we chose one representative sequence from each of 29 OTUs, and reconstructed phylogenetic trees using ML and Bayesian analyses. A total of seven well-separated lineages were inferred from our phylogenetic analysis (Figs. 1 and S2), consisting of five thermophilic lineages (viz., the "*Galdieria maxima* (Gm)" assemblage, *Cyanidium caldarium* (Cc), *Cyanidioschyzon merolae* (Cm), an uncharacterized *Cyanidium* species that is phylogenetically close to *C. merolae*, and the "*G. sulphuraria* (Gs)" assemblage), and two mesophilic *Cyanidium* lineages that are not genetically similar to any recognized species. In Ciniglia et al. (2014), they defined cyanidia from non-thermal and non-acidic sites as mesophilic *Cyanidium* (CspOTU3-6; as non-acidic mesophilic *Cyanidium* hereafter). In this study, we additionally identified a novel mesophilic *Cyanidium* clade from non-thermal, but acidic sites (CspOTU2; as acidic mesophilic *Cyanidium* hereafter). Among them, the two phylogenetically distant *Galdieria* lineages (the "Gm" and "Gs" assemblages) indicated that *Galdieria*

TABLE 1. Result of species delimitation of Cyanidiales.

Methods	Input tree	T	L_{null}	$L_{alternative}$	*P*	*N*	CI
GMYC	MCC tree	−0.014	4,308.642	4,317.477	1.455e-4	22	13–27
PTP	ML tree	N.A.	543.469	711.635	<0.001	20	N.D.

GMYC, general mixed Yule-coalescent; PTP, Poisson Tree Processes; T, relative threshold time; L_{null}, log likelihood of the null model; $L_{alternative}$, log likelihood of the alternative model; *P*, *P*-value of a likelihood ratio test; MCC tree, ultrametric maximum clade credibility tree; ML tree, non-ultrametric maximum likelihood tree; N, number of species; CI, confidence interval; N.A., not applicable; N.D., not determined.

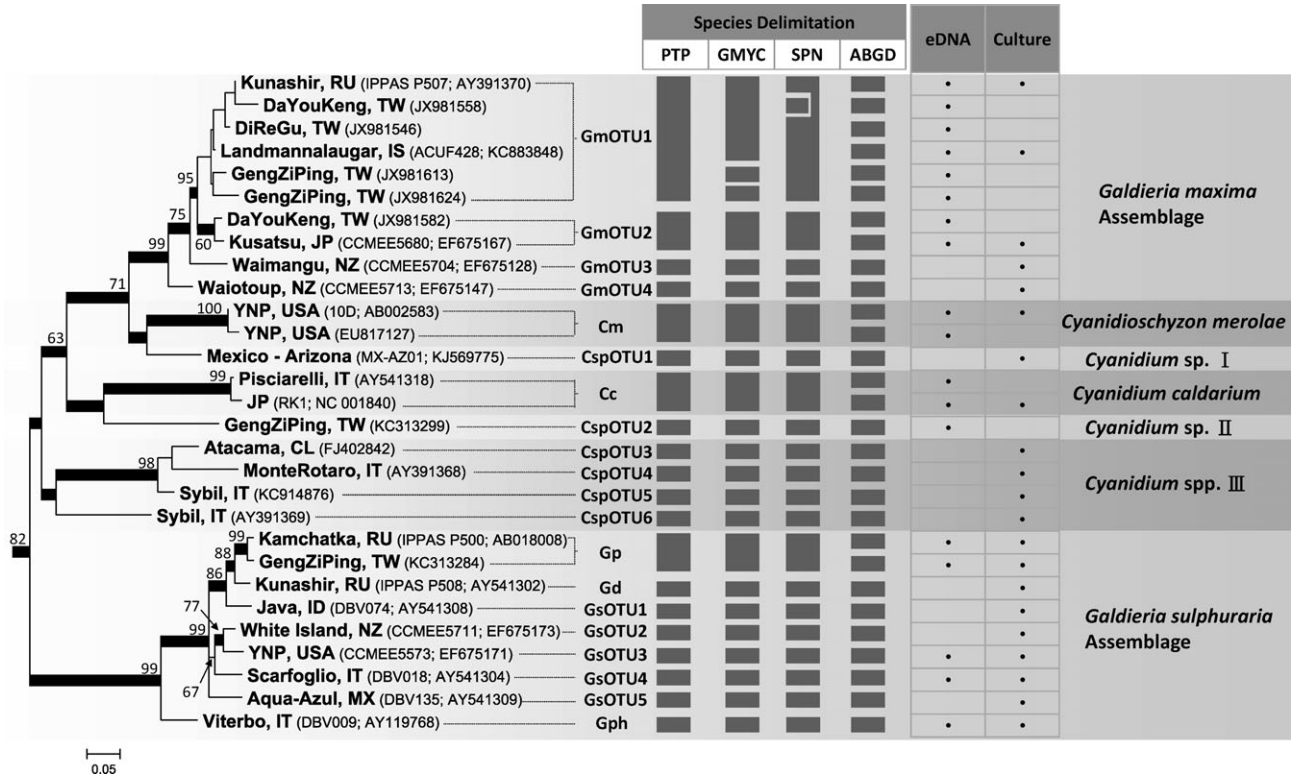

FIG. 1. Maximum likelihood *rbc*L phylogeny showing 29 ABGD operational taxonomic units (OTUs, or hypothetical species) worldwide. There are seven well-separated, supported clades: the "*Galdieria maxima*" assemblage (GmOTU1, GmOTU2, GmOTU3, and GmOTU4), *Cyanidioschyzon merolae* (Cm), *Cyanidium caldarium* (Cc), the "*Galdieria sulphuraria*" assemblage (GsOTU1-GsOTU5, Gd, Gp, and Gph) and three uncharacterized lineages (viz, *Cyanidium* sp. clade I [CspOTU1], *Cyanidium* sp. clade II [CspOTU2], and *Cyanidium* sp. clade III [CspOTU3-6]). Results of four different statistical species delimitation (PTP, GMYC, SPN, and ABGD) are summarized as vertical bars next to the OTU clusters (see Fig. S2 for details), as well as the OTUs discovered by the eDNA and culture-based methods. Statistical support based on 1,000 bootstrapping replicates is shown on branches (>60%), whereas the statistical support from the Bayesian posterior probability is presented by bold lines (>0.85). The strain identifier and GenBank accession number are indicated in parentheses. Abbreviation: Gd, *Galdieria daedala*; Cm, *Cyanidioschyzon merolae*; Csp, *Cyanidium* sp.; Gm, *Galdieria maxima*; Gp, *Galdieria partita*; Gph, *Galdieria phlegrea*; Gs, *Galdieria sulphuraria*. Scale bar is substitution per site.

is polyphyletic, consistent with previous observations based on multilocus analyses in Ciniglia et al. (2004) and Yoon et al. (2006). In addition, both lineages clearly showed higher OTU diversity than other lineages even when the most conservative OTU estimate was considered (Fig. 1). In the "Gm" assemblage, at least four OTUs (designated from GmOTU1 to GmOTU4) were discovered, but their taxonomic status requires further study (Fig. 1). In the "Gs" assemblage, at least eight OTUs were observed, including five GsOTUs (designated from GsOTU1 to GsOTU5) that also require further taxonomic investigation, *G. partita* (Gp), *G. daedala* (Gd), and *G. phlegrea* (Gph). In stark contrast to the two *Galdieria* lineages, there is only one OTU for Cc and one for Cm (Fig. 1). This trend would be more pronounced if the most liberal OTU estimates were considered (Figs. 1 and S2). For instance, six and one additional ABGD lineages were proposed in GmOTU1 and Gp, whereas only one additional ABGD lineages was proposed in Cm and Cc (Figs. 1 and S2). In this study, we identified a new thermophilic lineage from Mexico that is phylogenetically close to Cm, referred to as CspOTU1 (Figs. 1 and S2). At present, only one OTU could be recognized from this novel thermophilic clade due to limited sampling effort. Of the two mesophilic cyanidia lineages, the non-acidic clade consists of four species (CspOTU3-6), whereas the acidic clade comprises only one (CspOTU2).

Diversification dynamics in Cyanidiales. In the phylogenies, we observed greater OTU diversity in the two *Galdieria* assemblages compared to other lineages, and so we speculated that these assemblages might have undergone increased diversification that made them more OTU-rich than other lineages. To test this hypothesis, we applied the model-based PRC test on the four pruned MCC trees. We found no significant signals of elevated rates of cladogenesis in the 20-, 21-, and 22-tip trees (Fig. S3, A–C in the Supporting Information). We did, however, detect some significant evidence of rate increase in the 29-tip tree, in support of our hypothesis (Fig. S3D). The PRC result indicated that

diversification might have increased along the branch leading to the "Gm" assemblage. Nevertheless, we did not find robust evidence corroborating our hypothesis overall, but we believe that additional data need to be collected and then the hypothesis retested.

Niche differentiation in some cyanidia species. After determining OTUs, we investigated their distribution among microhabitats in the TVG area of Taiwan. Using the conservative OTU estimates, a total of four OTUs were present in the TVG area based on the eDNA data. These four OTUs were in the "Gm" assemblage (GmOTU1 and GmOTU2), Gp, and one unknown mesophilic cyanidia species (CspOTU2; Fig. 2). For the two OTUs in the "Gm" assemblage, GmOTU1 was widely distributed across many different microhabitats—including pools, streams, sulfur fumes, and endolithic environments—except for soil at low temperature (Fig. 2). In contrast, GmOTU2 was only found in pools and endolithic microhabitats at DYK, but less frequently or not at all at lower temperature environments (Fig. 2). Unlike GmOTU1 and GmOTU2, Gp primarily occurs in non-aquatic microhabitats, such as sulfur fumes, epilithic and endolithic microhabitats (Fig. 2), implying that Gp might prefer drier microhabitats. Expectedly, the acidic mesophilic cyanidia (CspOTU2) was predominantly found at lower temperature microhabitats (Fig. 2), suggesting that this alga might better adapt to low temperatures than other thermoacidophilic cyanidia. Overall, our data demonstrated that the type of microhabitat might influence the abundance and composition of different cyanidia OTUs in Taiwan.

To improve our understanding of the impact of microhabitat on cyanidia biodiversity, we compared our results with published results from two other geothermal areas, central Italy volcanic area (CIV) and YNP in the USA using the most conservative OTU estimates. Since OTU diversity surveys of mesophilic cyanidia are largely incomplete in most studies (e.g., Ciniglia et al. 2004, Yoon et al. 2006, Toplin et al. 2008, Skorupa et al. 2013, Ciniglia et al. 2014, this study), we did not include mesophilic cyanidia in the subsequent analyses. We also did not consider culture-based data because they would be only presence or absence data. From previous studies, moisture appears to correlate with cyanidia diversity. Therefore, we categorized the microhabitats into two groups: aquatic (A) or non-aquatic (NA). The aquatic microhabitats include pools and streams, whereas the non-aquatic microhabitats include soils, endolithics, epilithics, and sulfur fumes. Globally, Cyanidiales OTU diversity in the non-aquatic habitat was shown to be greater than that in the aquatic habitats ($Global_{NA}$: H: 1.85, PD: 1.38 > $Global_{A}$: H: 0.86, PD: 0.81; Fig. 3A). Four OTUs were present in the aquatic habitats, whereas seven in the non-aquatic habitats (Fig. 3A). To determine if this global trend is shown in each geothermal area, we repeated the analyses for each geothermal area. Both OTU and PD were shown to be

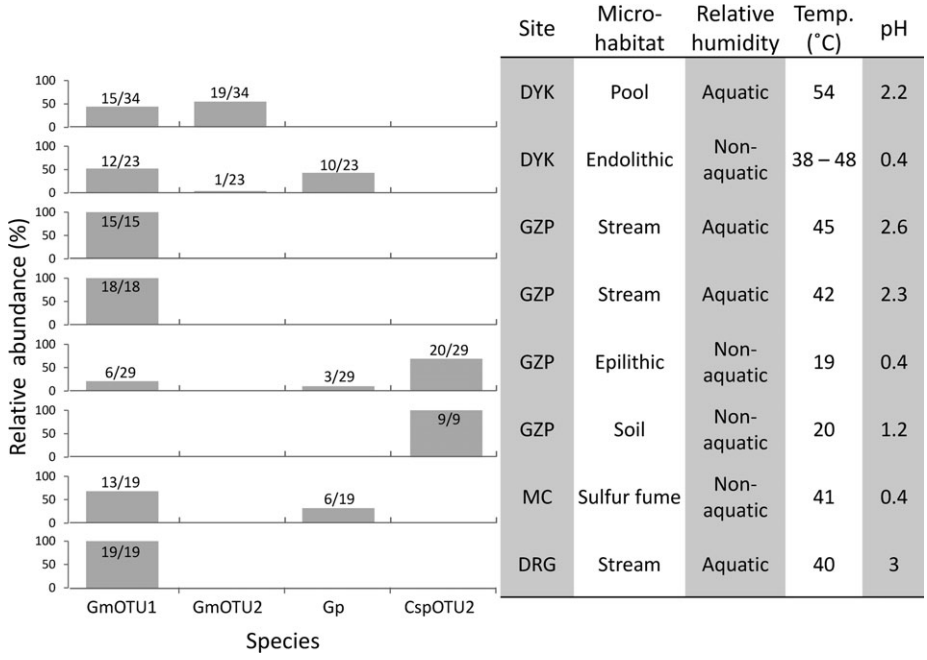

FIG. 2. Relative abundance of Cyanidiales species in the Tatun Volcano Group (TVG) area, Taiwan. A total of four Cyanidiales species were found in different microhabitats at four study sites, GengZiPeng (GZP), DaYouKeng (DYK), MaChao (MC), and DiReGu (DRG) based on the eDNA method. Abbreviation: Csp, *Cyanidium* sp.; Gm, *Galdieria maxima*; Gp, *Galdieria partita*; OTU, operational taxonomic unit.

greater in the non-aquatic habitats than in the aquatic habitats across the three geothermal areas (TVG_{NA}: H: 0.75, PD: 0.73 > TVG_{A}: H: 0.53, PD: 0.38; YNP_{NA}: H: 0.68, PD: 0.59 > YNP_{A}: H: 0.00; PD: 0.20; CIV_{NA}: H: 1.14, PD: 1.06 > CIV_{A}: H: 0.47, PD: 0.55; Fig. 3B). In the TVG area, three OTUs (GmOTU1, GmOTU2, and Gp) were in non-aquatic habitats, whereas only two OTUs (GmOTU1 and GmOTU2) in aquatic habitats. In YNP, two OTUs (Cm and GsOTU3) were in non-aquatic environments, but only one OTU (Cm) in aquatic environments. In the CIV area, four OTUs (Cc, Cm, GsOTU4, and Gph) were observed in non-aquatic environments, but only two OTUs in aquatic environments. The CIV area in Italy possesses both the highest OTU and PD, followed by the TVG area in Taiwan, with the lowest diversity in YNP, USA (Fig. 3B). According to our analyses, OTUs in the "Gs" assemblage (Gp, Gph, GsOTU3, and GsOTU4) prefer non-aquatic environments. On the other hand, the other two lineages (the "Gm" assemblage and Cm), which are distantly related to the "Gs" assemblage, but more closely related to each other, appear to prefer aquatic environments. Hence, the habitat preference in Cyanidiales might be associated with phylogenetic relatedness. This result indicated a much higher number of cyanidia OTU that is phylogenetically distant in the CIV area in Italy (Fig. 3B). To better compare the community structure between aquatic and non-aquatic habitats, we

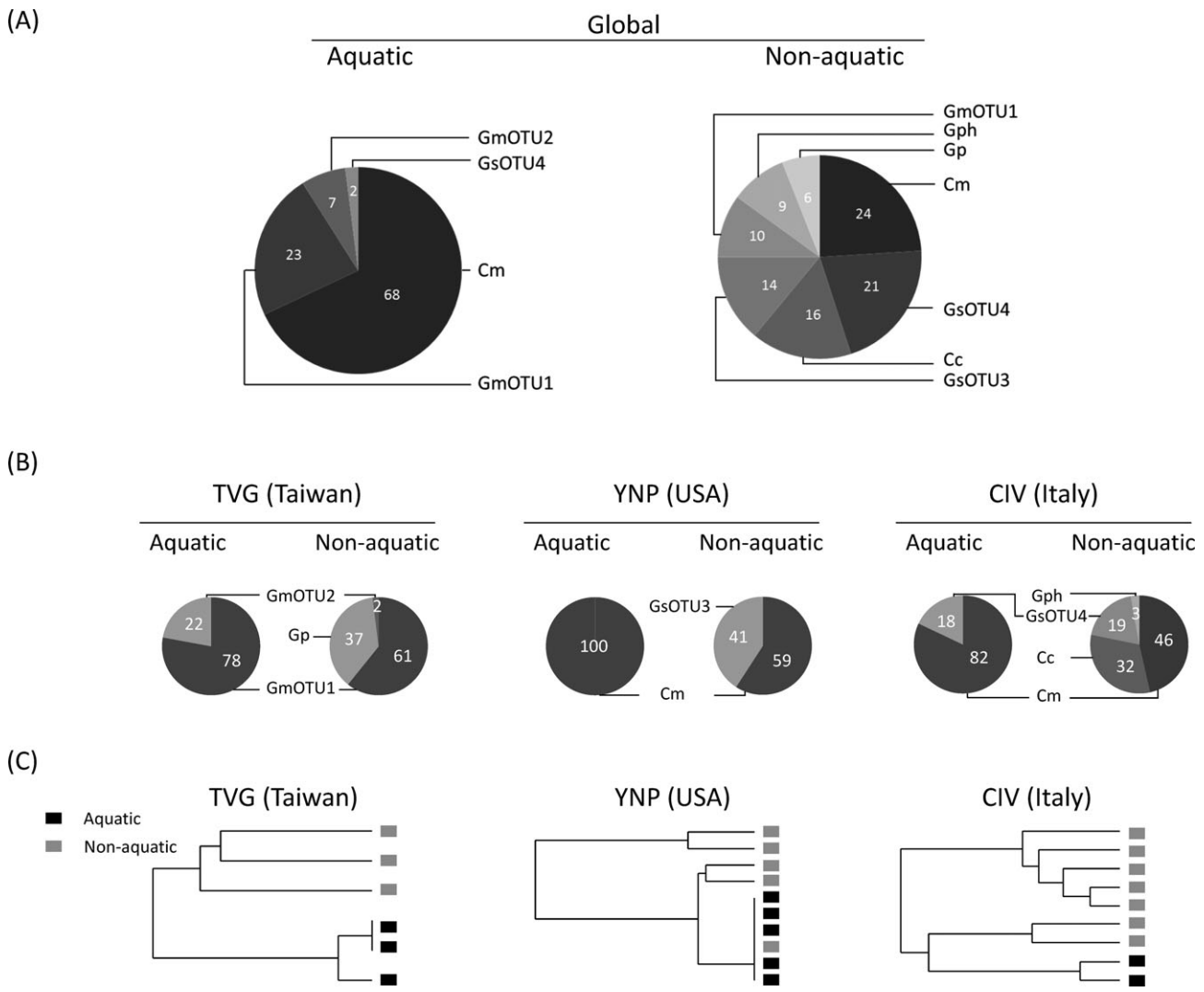

FIG. 3. Species composition of Cyanidiales in the aquatic habitats and the non-aquatic habitats of the Tatun Volcano Group area (TVG), Central Italy of Volcanic area (CIV) and Yellowstone National Park (YNP) based on the eDNA data set. (A) Global species composition of Cyanidiales in aquatic and non-aquatic habitats with all available data from three geothermal areas (TVG, CIV and YNP) summarized. The numbers in the pie graph indicate relative abundance. (B) Species composition of Cyanidiales in the aquatic and non-aquatic habitats for each of three geothermal areas, TVG, CIV and YNP. Number in the pie graph indicates the relative abundance. (C) Results of the picante cluster analysis, showing similarity among communities from different microhabitats in non-aquatic and aquatic environments. References for the data sources are provided in Figure S4A.

conducted a phylogenetic clustering analysis for samples from each geothermal area. In the TVG area, Taiwan, communities from the non-aquatic microhabitats grouped together and differed from those from the aquatic microhabitats (Fig. 3C). Similarly, communities from the aquatic habitats grouped together and differed from those in the non-aquatic habitats in the CIV area, Italy (Fig. 3C). Likewise, communities from aquatic and non-aquatic microhabitats generally formed two separated groups in YNP, USA (Fig. 3C), except for one sample from Mountain Basin where only Cm occurred in the endolithic microhabitat (Fig. S4A in the Supporting Information). Overall, our results supported the argument that moisture and type of habitats have substantial effects on community similarities in Cyanidiales.

In other geothermal areas, Cm was found to occur predominantly in the aquatic habitats, whereas Gs (= GsOTU3 and GsOTU4) and Gph were observed primarily in the non-aquatic habitats, such as endolithics and sulfur fume (Ciniglia et al. 2004). Our synthetic analyses showed that GsOTU3 and Gph might indeed primarily occur in the non-aquatic habitats (Figs. 3B and S4A), consistent with the observations of Ciniglia et al. (2004) and Skorupa et al. (2013). In contrast, we did not find that Cm and GsOTU4 show habitat-specific preference (Figs. 3B and S4A), consistent with the observations of Yoon et al. (2006) and Skorupa et al. (2013).

OTU composition and biogeographical affinities of global cyanidia communities. We determined the biogeographical affinity and community structures of cyanidia among different geothermal areas using the more conservative OTU estimates. We compared the cyanidia communities using only the eDNA data. We applied two methods to analyze the cyanidia community structure among different geothermal areas worldwide, first using the clustering method implemented in picante and then a PCA. The picante analysis showed that the TVG area communities grouped together (Fig. 4A). Most communities from the CIV area (Italy) and YNP (USA) were mainly divided into two separate groups (Fig. 4A), except for two samples (one from a stream and the other from a pool from Pisciarelli, Italy, that clustered with Cm-dominant communities from YNP). These observations indicated that the cyanidia community structure is influenced by geographic affinities. Consistent with the results of the picante analysis, the PCA showed that the cyanidia communities could be roughly segregated into three different groups, each strongly associated with a geothermal area (Fig. 4B). The first two principal components (PCs) from the PCA explain about 71% of the total variance of the data. This segregation was largely driven by differences in the dominant OTU in each area (Fig. 4B). In the TVG area, GmOTU1 was more abundant, in the CIV area (Italy), GsOTU4 was more abundant, and in YNP (USA), Cm was more abundant (Table 2).

We broadened the comparison of the cyanidia community structure by combining the eDNA and culture-based methods into a single data set. To mitigate potential bias in diversity estimation, we summarized the data as presence or absence of cyanidia OTUs to produce the cluster dendrogram. Results showed that the global cyanidia communities could be split into two groups. The first group consisted of communities from the CIV area in Italy and YNP in USA (North America-Southern Europe community), and the second group (West Pacific-Iceland community) comprised communities from Iceland, the Honshu area in Japan, the TVG area in Taiwan, and the North Island area in New Zealand (Fig. 5), reflecting their geographic proximity. Such global community structures might be ascribed to the regional endemism of some cyanidia OTUs. Based on the current sampling efforts, OTUs in the "Gm" assemblage were only found in the West Pacific-Iceland community, and Gd was only found in Kunashir Island, Russia and Iceland (Fig. S4). This observation suggests that the dispersal of some cyanidia OTUs may be hampered by great distances. However, the possibility remains that rare and/or lowly abundant OTUs have been missed by present sampling efforts. For instance, GsOTU3 was detected in Taiwan only through laboratory culturing (Fig. S4B).

DISCUSSION

Our analysis of *rbc*L sequence data using species delimitation methods revealed at least 20 cyanidia OTUs (interpreted as hypothetical species in this

TABLE 2. Factor loadings for variables, eigenvalues, and percentage of total variance explained for principal components.

Species	PC 1	PC 2
GmOTU1	−0.482	−0.658
GmOTU2	0.000	0.000
Cm	0.851	−0.232
Cc	−0.102	0.266
Gph	0.000	0.000
Gp	0.000	0.000
GsOTU3	0.000	0.000
GsOTU4	−0.143	0.653
Eigenvalue	0.459	0.246
% Variance explained	45.9	24.6

Principal components were derived from the correlation matrix of relative abundance data of different Cyanidiales species.

PC1, the first principal component; PC2, the second principal component; GmOTU1, OTU1 in the *G. maxima* assemblage; GmOTU2, OTU2 in the *G. maxima* assemblage; Cm, *Cyanidioschyzon merolae*; Cc, *Cyanidium caldarium*; Gph, *Galdieria phlegrea*; Gp, *Galdieria partita*; GsOTU3, OTU3 in the "*G. sulphuraria*" assemblage; GsOTU4, OTU4 in the "*G. sulphuraria*" assemblage.

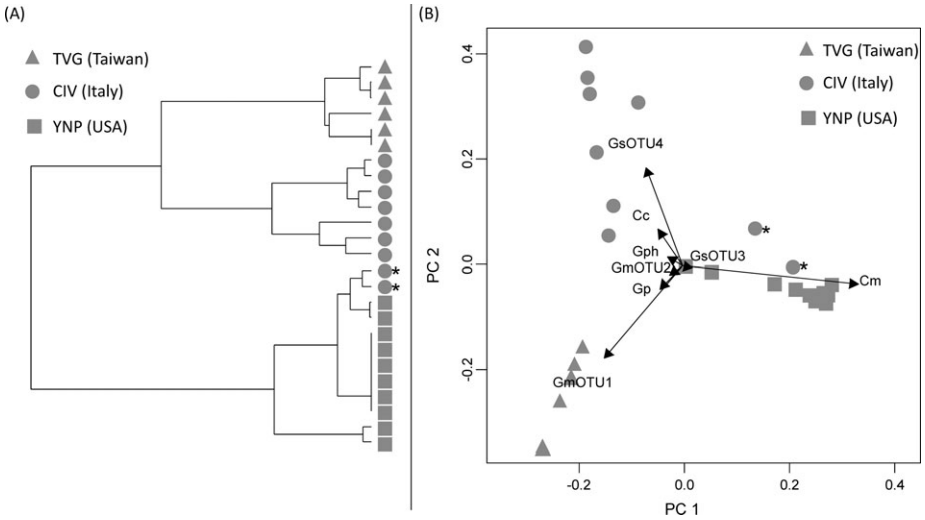

FIG. 4. Community similarity analyses of the eDNA-based data set from the Tatun Volcano Group area (TVG), Central Italy of Volcanic area (CIV) and Yellowstone National Park (YNP) using picante (A) and principal components analysis (B). Arrows indicate the degree of correlation between Cyanidiales species and two principal components (PC1 and PC2). Arrows indicate the correlation between the OTU and the site. Abbreviation: Csp, *Cyanidium* sp.; Gm, *Galdieria maxima*; Gp, *Galdieria partita*; OTU, operational taxonomic unit. Asterisk indicates the aquatic samples from Pisciarelli, Italy. References for the data sources are provided in Figure S4A.

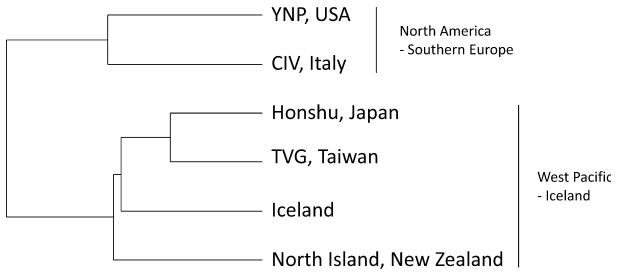

FIG. 5. Community similarity analysis of the eDNA and culture-based data sets combined using picante. Abbreviation: TVG, the Tatun Volcano Group area; CIV, the Central Italy Volcanic area; YNP, Yellowstone National Park. References for the data sources are provided in Figure S4B.

article) worldwide, more than three times the number of species recognized in the current taxonomy (e.g., Pinto et al. 2007). Instead of using some arbitrary sequence divergence cut-off to determine OTU boundaries (which has been a common practice), we applied statistical species delimitation methods that allow for more objective inference of OTUs (reviewed in Leliaert et al. 2014). The number of OTUs predicted using the four DNA-based algorithmic methods ranges from 20 to 29. Here, we took the most conservative estimate to be the species richness of Cyanidiales. Species delimitation methods can underestimate species diversity if the representation of intraspecific genetic variation is poor (reviewed in Carstens et al. 2013), such as in our *rbc*L data set. However, species delimitation from an ultrametric tree has been theoretically and empirically shown to yield higher estimates of OTUs (Zhang et al. 2013), which was the case in our study. In our study, the two phylotypes (IIIA and IIIB) of the "*G. maxima* (Gm)" assemblage identified by Toplin et al. (2008) were merged into a single OTU (GmOTU1), suggesting that our estimate of OTUs may be more conservative than estimates made by previous studies.

As Ciniglia et al. (2004) demonstrated, multilocus data should be collected in order to delimit the number of OTUs more accurately. We caution that the OTUs predicted in this study should be treated as hypothetical species that require further systematic examination. Nonetheless, the OTUs discovered in this study are potentially biologically meaningful. For example, our results support the separation of two closely related sister species, *G. partita* (Gp) and *G. daedala* (Gd), as described by Sentsova (1994). Among the 20 OTUs, we discovered a putative species which may be a novel member of acidic mesophilic cyanidia (CspOTU2) and is found in low temperature microhabitats. This novel cyanidia lineage is placed phylogenetically close to *Cyanidium caldarium* (Cc) rather than other non-acidic mesophilic cyanidia. Therefore, characterization of morphological and physiological traits of this alga may yield interesting information regarding acid and heat tolerance in Cyanidiales.

Besides the discovery of a new acidic mesophilic cyanidia, another contribution of our study is to produce molecular evidence for the species composition of Cyanidiales dwelling in the geothermal areas in Taiwan. Our analyses revealed that the "Gm" assemblage and Gp are the most predominant species in the TVG geothermal areas. Our phylogenetic analyses support the polyphyly of *Galdieria*. Considering that the genus name "*Galdieria*" was

first proposed to circumscribe the "*G. sulphuraria* (Gs)" assemblage (type species), a new genus name should be proposed for the "Gm" assemblage. Because we cannot differentiate these two cyanidia lineages with the present morphological and physiological information, we defer any taxonomic proposal for the "Gm" assemblage until systematic morphological and physiological examinations are conducted.

The phylogenetic data in this study allowed us to study the diversification dynamics of Cyanidiales, which has never been investigated to-date. Visual inspection of the phylogenetic trees revealed that the two *Galdieria* assemblages may exhibit greater OTU diversity than the other lineages. This may be more pronounced under the more liberal OTU estimates (e.g., 29 ABGD OTUs). Although we found no statistical evidence for this hypothesis under the conservative OTU estimate (20, 21, and 22 OTUs predicted by PTP, GMYC, and SPN, respectively), we did start to see significant increases in diversification rate in the "Gm" assemblage if we took a more liberal OTU estimate (i.e., 29 ABGD OTUs). Nonetheless, we believe that the current data may be inadequate to rigorously test diversification hypotheses in this algal clade (given small phylogenies with only 20 to 29 tips), and that the diversity dynamics of Cyanidiales should be revisited once more sequence data are accumulated.

Despite a lack of significant results from the diversification tests, there are intriguing observations that have led us to speculate that the two *Galdieria* assemblages might have enjoyed elevated diversification. We hypothesize that the acquisition of horizontally transferred prokaryotic genes and/or the evolution of mixotrophic ability might have spurred adaptive radiation in *Galdieria*. In stark contrast to the obligatory photosynthesis in *Cyanidium* and *Cyanidioschyzon* (Gross 1999, Barbier et al. 2005), Schönknecht et al. (2013) and Qiu et al. (2013) revealed that Gs and Gph could adapt to extreme environments and grow mixtrophically by horizontally acquiring genes from prokaryotes. These horizontally transferred genes (HGTs) could help the algae to tolerate high temperatures and high concentrations of heavy metals, to utilize urea as a nitrogen source, and to metabolize glycerol as an organic carbon source. Although it is unclear whether HGTs play important role in Gm, Gm was shown to have heterotrophic growth ability (Gross and Oesterhelt 1999). The evolution of these physiological abilities might have allowed the genus *Galdieria* to undergo niche differentiation, which in turn might have facilitated their diversification.

Consistent with the popular view that microbial diversity is largely shaped by environmental factors (Meyer-Dombard et al. 2005, Hou et al. 2013), we found that the PD of Cyanidiales is higher in non-aquatic habitats than in aquatic habitats. One possible explanation is greater microhabitat heterogeneity in non-aquatic habitats. Species diversity has been shown to increase with habitat heterogeneity (e.g., Tylianakis et al. 2008). The non-aquatic habitats covered in this study comprise at least four different types of microhabitats—endolithics, epilithics, soils, and sulfur fumes. In aquatic habitats, we only explored two different types of microhabitats: pools and streams. Another explanation is that non-aquatic habitats experience a higher degree of disturbances, leading to reduced competition and more space for the colonization of different cyanidia species (Gause 1934, Connell 1978). In general, non-aquatic microhabitats are exposed to a broader range of fluctuations in environmental factors—such as temperature and pH—than aquatic habitats, where environmental factors may be homogenized via mixing by water currents. In addition, previous studies and our own analyses revealed that Cm is mainly distributed in aquatic environments (e.g., Yoon et al. 2006, Skorupa et al. 2013), probably due to the reduction in its osmotic adaptation by the lack of vacuoles, cell wall, and osmolyte synthesis (Barbier et al. 2005).

Our global comparison of cyanidia communities showed that some cyanidia may be endemic to certain geographic areas. For instance, the West Pacific-Iceland restricted distribution of the "Gm" assemblage and Gd suggests that these algae may be unable to disperse globally. But, in the event that they do disperse across oceans, they might be selected against by different volcanic environments. In contrast, the distribution patterns of some cyanidia species are widespread around the globe (e.g., GsOTU3 and Cm). How do Cyanidiales disperse over such a long distance? The *rbc*L phylogenies suggest that identical or closely related strains of some cyanidia species can be found in distantly separated continents (e.g., Ciniglia et al. 2004, Toplin et al. 2008, and this study). Based on these data, we speculate that a continuous gene flow exists between cyanidia populations on different continents, and that it is maintained by some unidentified dispersal mechanisms. One possibility is that some cyanidia species might be dispersed via the migration of water birds, which often carry with them a small amount of acidic mud containing viable cyanidia cells (e.g., Castenholz and McDermott 2010). The other possibility is that some cyanidia species are aerophytic algae and so wind flow might be a likely dispersal vector (Ciniglia et al. 2014). Another possible mechanism is air dispersal through volcanic activities (e.g., Reeb and Bhattacharya 2010). A fourth mode of dispersal might be mediated by thermobiologists or tourists who unintentionally transport cyanidia species from one geothermal area to another (e.g., Castenholz and McDermott 2010). However, dispersal by thermobiologists or tourists might more likely happen within a geothermal area or a country. Intercontinental transport of Cyanidiales might be unlikely because these organisms are

intolerant to long-term desiccation (Toplin et al. 2008). But species belonging to the "Gs" assemblage possess much higher metabolic flexibility. So, it might be not surprising to see that GsOTU3 is found on distant continents, for example.

One caveat of this study is that the higher species diversity in non-aquatic habitats may be affected when the most liberal OTU estimate (29 OTUs) is considered instead of more conservative estimates. Our analyses showed that this is not the case. The subdivisions of Gp would lead to higher OTU numbers in non-aquatic habitats than aquatic habitats in TVG (Fig. 1; Fig. S2), whereas the subdivisions of GmOTU1 would contribute equally to both habitats. Another caveat concerning the results of our community similarity analyses is that they may be affected by horizontal transfers of *rbc*L. Horizontally transferred *rbc*L might obscure the global community structure of Cyanidiales. For instance, Toplin et al. (2008) demonstrated that some populations of Gm from New Zealand (phylotype IV) may possess *rbc*L transferred horizontally from species in the "Gs" assemblage. If we factor in these findings, the New Zealand community would be more closely affiliated with other communities in the West Pacific and Iceland, where members of the "Gm" assemblage are the dominant species. However, the extent to which *rbc*L is horizontally transferred among Cyanidiales remains unknown. This will be an important topic for future research. A third caveat is that our analyses of the phylogenetic community structure is limited by a small sample size, which would potentially lead to biased conclusions. For example, only three geothermal areas in which we have eDNA data sets (i.e., Italy, USA, and Taiwan) can be compared using PCA and Picante. Therefore, our finding that OTU diversity is greater in non-aquatic habitats than in aquatic habitats should be re-examined with an increased sample size. Nonetheless, our study provides a good basis for us to further study species diversity, community structure, and biogeography of Cyanidiales in the future.

The data set used in our study includes *rbc*L sequences sampled around the world (Japan, New Zealand, Taiwan, Iceland, Italy, Mexico, and USA) using eDNA and cultured-based methods. But how comprehensively have the current sampling efforts captured the species diversity of Cyanidiales? The species diversity of Cyanidiales in many geothermal areas around the world remains to be explored. For instance, the ring of fire in the West Pacific is still a "black box" without much known about the species diversity of their resident cyanidia. Likewise, the cyanidia species diversity in the geothermal areas in Central America, South America, and Africa is also unknown. Many of the above-mentioned geothermal areas have been shown to be inhabited by "*Cyanidium*" (Brock 1978). Our present knowledge about the global species diversity of Cyanidiales is still very limited and requires future diversity surveys to cover aforementioned regions.

Exploration of the species diversity of Cyanidiales has been hindered by the lack of the appropriate tools. At present, the only widely used methods to survey cyanidia biodiversity are the eDNA and culture-based approaches combined with the cloning-Sanger method introduced in this study. These two approaches cannot fully capture the species diversity of Cyanidiales. For instance, GsOTU3 cannot be detected using the eDNA method unless the time-consuming culture-based method was applied. This additional culturing step hampers progress in investigating the global species diversity of Cyanidiales. Recently, targeted next-generation sequencing (NGS) has been applied to sample species diversity at an unprecedented resolution (e.g., Cuvelier et al. 2010, Egge et al. 2013, Taylor and Cunliffe 2014). The studies using targeted NGS are revealing species diversity never appreciated before. Therefore, some cyanidia OTUs in diversity surveys may be missed due to poor sensitivity of the cloning-Sanger eDNA approach when species abundance is too low.

In this study, we demonstrated that DNA-based algorithmic species delimitation can reveal OTU diversity of unicellular algae that may be missed by morphological examination alone. We then introduced and tested several intriguing hypotheses about the biodiversity and biogeography of Cyanidiales (for example, adaptive radiation in *Galdieria* and the endemism of the "Gm" assemblage and Gd). However, we believe that these hypotheses need to be revisited with more appropriate tools (e.g., targeted NGS) and improved sampling effort. We anticipate that studies using similar approaches will shed light on the evolutionary and ecological mechanisms that shape and maintain the biodiversity of microalgae.

We thank Dr. Liang-Chi Wang (Collection Management Department, National Taiwan Museum, Taiwan) for sample collections. This study was financially supported by grants from the Ministry of Science and Technology in Taiwan (NSC101-2621-B-029-004 and NSC102-2621-B-029-002) and the Jiayi Foundation in Taiwan to SLL. The funding agencies played no role in the study design, data collection, data analyses, decision to publish, or preparation of this manuscript.

Azúa-Bustos, A., Gonzalez-Silva, C., Mancilla, R. A., Salas, L., Palma, R. E., Wynne, J. J., McKay, C. P. & Vicuna, R. 2009. Ancient photosynthetic eukaryote biofilms in an Atacama Desert coastal cave. *Microb. Ecol.* 58:485–96.

Barbier, G., Oesterhelt, C., Larson, M. D., Halgren, R. G., Wilkerson, C., Garavito, R. M., Benning, C. & Weber, A. P. 2005. Comparative genomics of two closely related unicellular thermo-acidophilic red algae, *Galdieria sulphuraria* and *Cyanidioschyzon merolae*, reveals the molecular basis of the metabolic flexibility of *Galdieria sulphuraria* and significant differences in carbohydrate metabolism of both algae. *Plant Physiol.* 137:460–74.

Brock, T. D. 1967. Life at high temperatures. *Science* 158:1012–9.

Brock, T. D. 1978. The genus Cyanidium. *In* Starr, P. M. [Eds] *Thermophilic Microorganisms and Life at High Temperatures.* Springer - Verlag, New York, pp. 255–301.

Carstens, B. C., Pelletier, T. A., Reid, N. M. & Satler, J. D. 2013. How to fail at species delimitation. *Mol. Ecol.* 22:4369–83.

Castenholz, R. W. & McDermott, T. R. 2010. The cyanidiales: ecology, biodiversity, and biogeography. *In* Seckbach, J. & Chapman, D. J. [Eds.] *Red Algae in the Genomic Age*. Springer, Dordrecht, The Netherlands, pp. 357–71.

Ciniglia, C., Yang, E. C., Pinto, G., Iovinella, M., Vitale, L. & Yoon, H. S. 2014. Cyanidiophyceae in Iceland: plastid rbcL gene elucidates origin and dispersal of extremophilic *Galdieria sulphuraria* and *G. maxima* (Galdieriaceae, Rhodophyta). *Phycologia* 53:542–51.

Ciniglia, C., Yoon, H. S., Pollio, A., Pinto, G. & Bhattacharya, D. 2004. Hidden biodiversity of the extremophilic Cyanidiales red algae. *Mol. Ecol.* 13:1827–38.

Clement, M., Posada, D. & Crandall, K. A. 2000. TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* 9:1657–9.

Connell, J. H. 1978. Diversity in tropical rain forests and coral reefs. *Science* 199:1302–10.

Cuvelier, M. L., Allen, A. E., Monier, A., McCrow, J. P., Messié, M., Tringe, S. G., Woyke, T. et al. 2010. Targeted metagenomics and ecology of globally important uncultured eukaryotic phytoplankton. *Proc. Natl. Acad. Sci. USA* 107:14679–84.

Drummond, A. J. & Rambaut, A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7:214.

Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32:1792–7.

Egge, E., Bittner, L., Andersen, T., Audic, S., de Vargas, C. & Edvardsen, B. 2013. 454 pyrosequencing to describe microbial eukaryotic community composition, diversity and relative abundance: a test for marine haptophytes. *PLoS ONE* 8: e74371.

Faith, D. P. 1992. Conservation evaluation and phylogenetic diversity. *Biol. Conserv.* 61:1–10.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39:783–91.

Fordyce, J. A., Shah, P. & Fitzpatrick, B. M. 2014. iteRates: an R package for implementing a parametric rate comparison on phylogenetic trees. *Evol. Bioinform.* 10:127.

Gause, G. 1934. The struggle for existence. *Soil Sci.* 41:159.

Geitler, L. 1933. Diagnosen neuer Blaualgen von den Sunda-Inseln. *Arch. Hydrobiol. Suppl.* 12:622–34.

Gimmler, H. 2001. Acidophilic and acidotolerant algae. *In* Rai, L. C. & Gaur, J. P. [Eds.] *Algal Adaptation to Environmental Stresses–Physiological, Biochemical and Molecular Mechanisms*. Springer, Berlin, pp. 259–90.

Gross, W. 1999. Revision of comparative traits for the acido- and thermophilic red algae *Cyanidium* and *Galdieria*. *In* Seckbach, J. [Ed.] *Enigmatic Microorganisms and Life in Extreme Environments*. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 437–46.

Gross, W., Heilmann, I., Lenze, D. & Schnarrenberger, C. 2001. Biogeography of the Cyanidiaceae (Rhodophyta) based on 18S ribosomal RNA sequence data. *Eur. J. Phycol.* 36:275–80.

Gross, W. & Oesterhelt, C. 1999. Ecophysiological studies on the red alga *Galdieria sulphuraria* isolated from southwest Iceland. *Plant Biol.* 1:694–700.

Guiry, M. D. & Guiry, G. M. 2014. AlgaeBase. World-wide electronic publication, National University of Ireland, Galway. Available at: http://www.algaebase.org (searched on 28 October 2014).

Hart, M. W. & Sunday, J. 2007. Things fall apart: biological species form unconnected parsimony networks. *Biol. Lett.* 3:509–12.

Hill, M. O. 1973. Diversity and evenness: a unifying notation and its consequences. *Ecology* 54:427–73.

Horner-Devine, M. C. & Bohannan, B. J. M. 2006. Phylogenetic clustering and overdispersion in bacterial communities. *Ecology* 87:S100–8.

Hou, W., Wang, S., Dong, H., Jiang, H., Briggs, B. R., Peacock, J. P., Huang, Q., Huang, L., Wu, G. & Zhi, X. 2013. A comprehensive census of microbial diversity in hot springs of Tengchong, Yunnan Province China using 16S rRNA gene pyrosequencing. *PLoS ONE* 8:e53350.

Huelsenbeck, J. P. & Ronquist, F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17:754–5.

Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., Blomberg, S. P. & Webb, C. O. 2010. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 26:1463–4.

Kondo, A., Yokoyama, A. & Hara, Y. 2004. Phylogenetic analysis of cyanidian algae (Cyanidiales, Rhodophyta), using 18S rRNA gene sequences. *Jpn. J. Phycol.* 52:195–9.

Lehr, C. R., Frank, S. D., Norris, T. B., D'Imperio, S., Kalinin, A. V., Toplin, J. A., Castenholz, R. W., Timothy, R. & McDermott, T. R. 2007. Cyanidia (Cyanidiales) population diversity and dynamics in an acid-sulfate-chloride spring in Yellowstone National Park. *J. Phycol.* 43:3–14.

Leliaert, F., Verbruggen, H., Vanormelingen, P., Steen, F., López-Bautista, J. M., Zuccarello, G. C. & De Clerck, O. 2014. DNA-based species delimitation in algae. *Eur. J. Phycol.* 49:179–96.

Liu, C. M., Song, S. R., Chen, Y. L. & Tsao, S. 2011. Characteristics and origins of hot springs in the Tatun Volcano Group in northern Taiwan. *Terr. Atmos. Ocean Sci.* 22:475–89.

MacElroy, R. 1974. Some comments on the evolution of extremophiles. *Biosystems* 6:74–5.

Meyer-Dombard, D., Shock, E. & Amend, J. 2005. Archaeal and bacterial communities in geochemically diverse hot springs of Yellowstone National Park, USA. *Geobiology* 3:211–27.

Minoda, A., Sakagami, R., Yagisawa, F., Kuroiwa, T. & Tanaka, K. 2004. Improvement of culture conditions and evidence for nuclear transformation by homologous recombination in a red alga, *Cyanidioschyzon merolae* 10D. *Plant Cell Physiol.* 45:667–71.

Misumi, O., Matsuzaki, M., Nozaki, H., Miyagishima, S. Y., Mori, T., Nishida, K., Yagisawa, F., Yoshida, Y., Kuroiwa, H. & Kuroiwa, T. 2005. Cyanidioschyzon merolae genome. A tool for facilitating comparable studies on organelle biogenesis in photosynthetic eukaryotes. *Plant Physiol.* 137:567–85.

Morrison, D. A. 2005. Networks in phylogenetic analysis: new tools for population biology. *Int. J. Parasitol.* 35:567–82.

Oksanen, J. 2009. Multivariate Analysis of Ecological Communities in R: Vegan Tutorial. Available at: http://cran.r-project.org.

Pinto, G., Ciniglia, C., Cascone, C. & Pollio, A. 2007. Species composition of Cyanidiales assemblages in Pisciarelli (Campi Flegrei, Italy) and description of *Galdieria phlegrea* sp. nov. *In* Seckbach, J. [Ed.] *Algae and Cyanobacteria in Extreme Environments*. Springer, Heidelberg, New York, pp. 487–502.

Pons, J., Barraclough, T., Gomez-Zurita, J., Cardoso, A., Duran, D., Hazell, S., Kamoun, S., Sumlin, W. & Vogler, A. 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Syst. Biol.* 55:595–609.

Puillandre, N., Lambert, A., Brouillet, S. & Achaz, G. 2012. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Mol. Ecol.* 21:1864–77.

Qiu, H., Price, D. C., Weber, A. P., Reeb, V., Yang, E. C., Lee, J. M., Kim, S. Y., Yoon, H. S. & Bhattacharya, D. 2013. Adaptation through horizontal gene transfer in the cryptoendolithic red alga *Galdieria phlegrea*. *Curr. Biol.* 23:R865–6.

R Core Team. 2014. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at: http://www.R-project.org/.

Reeb, V. & Bhattacharya, D. 2010. The Thermo-acidophilic Cyanidiophyceae (Cyanidiales). *In* Seckbach, J. & Chapman, D. J. [Eds.] *Red Algae in the Genomic Age*. Springer, Dordrecht, The Netherlands, pp. 411–26.

Schönknecht, G., Chen, W. H., Ternes, C. M., Barbier, G. G., Shrestha, R. P., Stanke, M., Brautigam, A. et al. 2013. Gene transfer from bacteria and archaea facilitated evolution of an extremophilic eukaryote. *Science* 339:1207–10.

Schwabe, G. H. 1936. Über einige Blaualgen aus dem mittleren und südlichen Chile. *Verh.des Deutsch. Wiss. Ver. Santiago de Chile*. 3:113–74.

Sentsova, O. Y. 1994. The study of Cyanidiophyceae in Russia. *In* Seckbach, J. [Ed.] *Evolutionary Pathways and Enigmatic Algae: Cyanidium caldarium (Rhodophyta) and Related Cells*. Kluwer

Academic Publishers, Dordrecht, The Netherlands/Boston, MA/London, pp. 167–74.
Shah, P., Fitzpatrick, B. M. & Fordyce, J. A. 2013. A parametric method for assessing diversification-rate variation in phylogenetic trees. *Evolution* 67:368–77.
Skorupa, D. J., Reeb, V., Castenholz, R. W., Bhattacharya, D. & McDermott, T. R. 2013. Cyanidiales diversity in Yellowstone National Park. *Lett. Appl. Microbiol.* 57:459–66.
Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Bio. Evol.* 30:2725–9.
Taylor, J. D. & Cunliffe, M. 2014. High-throughput sequencing reveals neustonic and planktonic microbial eukaryote diversity in coastal waters. *J. Phycol.* 50:960–5.
Templeton, A. R., Crandall, K. A. & Sing, C. F. 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* 132:619–33.
Toplin, J. A., Norris, T. B., Lehr, C. R., McDermott, T. R. & Castenholz, R. W. 2008. Biogeographic and phylogenetic diversity of thermoacidophilic cyanidiales in Yellowstone National Park, Japan, and New Zealand. *Appl. Environ. Microb.* 74:2822–33.
Tylianakis, J. M., Rand, T. A., Kahmen, A., Klein, A. M., Buchmann, N., Perner, J. & Tscharntke, T. 2008. Resource heterogeneity moderates the biodiversity-function relationship in real world ecosystems. *PLoS Biol.* 6:e122.
Walker, J. J., Spear, J. R. & Pace, N. R. 2005. Geobiology of a microbial endolithic community in the Yellowstone geothermal environment. *Nature* 434:1011–4.
Webb, C. O., Ackerly, D. D. & Kembel, S. W. 2008. Phylocom: software for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics* 24:2098–100.
Yoon, H. S., Ciniglia, C., Wu, M., Comeron, J. M., Pinto, G., Pollio, A. & Bhattacharya, D. 2006. Establishment of endolithic populations of extremophilic Cyanidiales (Rhodophyta). *BMC Evol. Biol.* 6:78.
Zhang, J., Kapli, P., Pavlidis, P. & Stamatakis, A. 2013. A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* 29:2869–76.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web site:

Figure S1. Photographs showing different types of microhabitats collected, including non-aquatic habitats: (A) Endolithics, (B) Epilithics, (C) Sulfur Fumes, and (D) Soils, and aquatic habitats: (E) Streams and (F) Pools. Arrows indicate where the samples were collected.

Figure S2. Ultrametric maximum clade credibility *rbc*L tree from the Bayesian inference with the results of statistical species delimitation from the Poisson Tree Processes (PTP), Generalized Mixed-Yule Coalescent (GMYC), statistical parsimony network (SPN), and automated barcode gap discovery (ABGD) analysis based on 413 *rbc*L sequences. Outgroup taxa were pruned prior to analysis. The columns to the right of trees indicate evolutionarily independent lineages obtained under PTP (the first column), GMYC (the second column), SPN (the third column), and ABGD (the fourth column) method. ID at the terminal nodes of tree indicates the collection information and strain/GenBank accession number (in parentheses). The statistical support based on 1,000 bootstrapping replicates with maximum likelihood tree is shown on branches (>60%) whereas the statistical support from the Bayesian posterior probability is presented by bold lines (>0.85). Scale bar is the substitution per site.

Figure S3. Examples showing results of the Parametric Rates Comparison (PRC) test applied to the Cyanidiales *rbc*L phylogenies. Statistics are computed for each ultrametric MrBayes phylogeny based on the outcomes of four different OTU delineation approaches (A–D). The colored circle ranges from a relative rate increase (red) to a relative rate decrease (blue). The size of the circle is scaled to the statistical support for a diversification rate shift at that branch. *P*-value is shown on the colored circles.

Figure S4. Number of global *rbc*L sequences from the cloning libraries from the eDNA approach (A) and direct sequencing from the culture-based approach (B). Collection information is indicated in the table. Black circles indicate the presence of species and white circles indicate absence. The number inside the black circle is the number of *rbc*L sequences from the clone libraries or the cultures. The statistical support based on 1,000 bootstrap replicates with maximum likelihood tree inference is shown on branches (>60%), whereas the statistical support from the Bayesian posterior probability is presented by bold lines (>0.85). Scale bar is substitution per site.

Table S1. Collection information and Genbank number for samples used in the eDNA-based sequencing.

Table S2. Collection information and GenBank number for samples used in the culture-based sequencing.

Table S3. GenBank information for outgroup taxa.