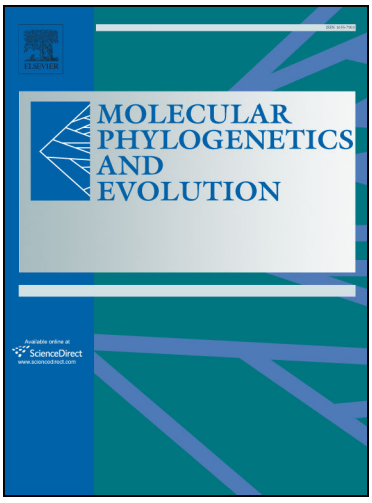
Journal Pre-proofs



Formally described species woefully underrepresent phylogenetic diversity in the common lichen photobiont genus *Trebouxia* (Trebouxiophyceae, Chlorophyta): An impetus for developing an integrated taxonomy

Lucia Muggia, Matthew P. Nelsen, Paul M. Kirika, Eva Barreno, Andreas Beck, Hanna Lindgren, H. Thorsten Lumbsch, Steven D. Leavitt, Trebouxia working group,

PII: S1055-7903(20)30093-2
DOI: https://doi.org/10.1016/j.ympev.2020.106821
Reference: YMPEV 106821

To appear in: *Molecular Phylogenetics and Evolution*

Received Date: 16 July 2019
Revised Date: 30 March 2020
Accepted Date: 1 April 2020

Please cite this article as: Muggia, L., Nelsen, M.P., Kirika, P.M., Barreno, E., Beck, A., Lindgren, H., Thorsten Lumbsch, H., Leavitt, S.D., Trebouxia working group, Formally described species woefully underrepresent phylogenetic diversity in the common lichen photobiont genus *Trebouxia* (Trebouxiophyceae, Chlorophyta): An impetus for developing an integrated taxonomy, *Molecular Phylogenetics and Evolution* (2020), doi: https://doi.org/10.1016/j.ympev.2020.106821

Formally described species woefully underrepresent phylogenetic diversity in the common lichen photobiont genus *Trebouxia* (Trebouxiophyceae, Chlorophyta): An impetus for developing an integrated taxonomy

Lucia Muggia[1], Matthew P. Nelsen[2], Paul M. Kirika[3], Eva Barreno[4], Andreas Beck[5], Trebouxia working group[#], Hanna Lindgren[2], H. Thorsten Lumbsch[2], Steven D. Leavitt[6]*

[1] University of Trieste, Department of Life Sciences, via Giorgieri 10, 34127-Trieste, Italy
[2] Science & Education, The Field Museum, Chicago, Illinois, USA
[3] Botany Department, EA Herbarium, National Museums of Kenya, P. O Box 40658-00100, Nairobi, Kenya
[4] Botánica, Instituto Cavanilles de Biodiversidad y Biología Evolutiva, Fac. CC. Biológicas, Universitat de València, C/ Dr. Moliner, 50. 46100-Burjassot, Valencia, Spain.
[5] Botanische Staatssammlung München, SNSB-BSM, Menzinger Str. 67, D-80638 Munich, Germany
[6] Department of Biology and M. L. Bean Life Science Museum, Brigham Young University, Provo, Utah, USA

*corresponding author e-mail: steve_leavitt@byu.edu

[#] Fernando Fernandez-Mendoza[7], Sergio Perez-Ortega[8], Martin Grube[7], Patricia Moya[4], Arantzazu Molins[4], Anna Sadowska-Deś[9,10], Beata Guzowa-Krzemińska[11], Ulrike Ruprecht[12], Francesco Dal Grande[13,14], Garima Singh[14], Anna Voytsekhovich[15]

[7] University of Graz, Institute of Biology, Holteigasse 6, 8010 Graz, Austria
[8] Real Jardín Botánico, CSIC, Plaza de Murillo, 2, 28014 Madrid, Spain
[9] Department of Biological Sciences, Institute of Ecology, Evolution and Diversity, Goethe Universität, Max-von-Laue-Str. 13, D-60438 Frankfurt, Germany
[10] Biodiversity and Climate Research Centre (BiK-F), Senckenberg Gesellschaft für Naturforschung, Senckenberganlage 25, D-60325 Frankfurt, Germany

[11] University of Gdańsk, Faculty of Biology, Department of Plant Taxonomy and Nature Conservation, Wita Stwosza 59, PL-80-308 Gdańsk, Poland

[12] University of Salzburg, Department of Biosciences, Hellbrunnerstrasse 34, 5020 Salzburg, Austria

[13] Senckenberg Biodiversity and Climate Research Centre (SBiK-F), Senckenberganlage 25, 60325 Frankfurt am Main, Germany

[14] LOEWE Centre for Translational Biodiversity Genomics (TBG), Senckenberganlage 25, 60325 Frankfurt am Main, Germany

[15] Bush Habitat Restoration Co-op, 120 Fitzgerald Ave, Maroubra NSW 2035, Australia

Abstract

Lichens provide valuable systems for studying symbiotic interactions. In lichens, these interactions are frequently described in terms of availability, selectivity and specificity of the mycobionts and photobionts towards one another. The lichen-forming, green algal genus *Trebouxia* Puymaly is among the most widespread photobiont, associating with a broad range of lichen-forming fungi. To date, 29 species have been described, but studies consistently indicate that the vast majority of species-level lineages still lack formal description, and new, previously unrecognized lineages are frequently reported. To reappraise the diversity and the evolutionary relationships of species-level lineages in *Trebouxia*, we assembled DNA sequence data from over 1600 specimens, compiled from a range of sequences from previously published studies, axenic algal cultures, and lichens collected from poorly sampled regions. From these samples, we selected representatives of the currently known genetic diversity in the lichenized *Trebouxia* and inferred a phylogeny from multi-locus sequence data (ITS, *rbcL*, *cox*2). We demonstrate that the current formally described species woefully underrepresent overall species-level diversity in this important lichen-forming algal genus. We anticipate that an integrative taxonomic approach, incorporating morphological and physiological data from axenic cultures with genetic data, will be required to establish a robust, comprehensive taxonomy for *Trebouxia*. The data presented here provide an important impetus and reference dataset for more reliably characterizing diversity in lichenized algae and in using lichens to investigate the evolution of symbioses and holobionts.

Key words: algae, biodiversity, fungi, holobiont, multigene, species delimitation, symbiosis.

1. Introduction

Among symbiotic associations, those that incorporate photosynthetic partners often coincide with major transformative events in evolutionary history, such as the photosynthetic-centered symbioses in the evolutionary lineage that emerged as plants (Moreira et al., 2000) and associations that led to lichens (Gargas et al., 1995). Algae-associated symbiotic interactions continue to reveal crucial insights into central processes of symbioses (Zook, 2015) and species diversification, where interactions often span both evolutionary and ecological time scales (Peksa and Škaloud, 2011; Rowan and Powers, 1991; Werth and Sork, 2014). Studying these interactions can lead to reconfiguring insight into diverse patterns of mutualistic associations.

Lichens represent iconic examples of symbiosis (Honegger, 2009) and are a valuable system for studying mutualistic interactions (Peksa and Škaloud, 2011; Spribille, 2018). These obligate associations of a predominant filamentous fungus, the mycobiont, with a population of unicellular photosynthesizing green alga and/or a cyanobacterium, the photobiont (i.e. the chlorobiont and the cyanobiont, respectively), coupled with a suite of other associated algae, bacteria and fungi (Arnold et al., 2009; Casano et al., 2010; Grube et al., 2009; Moya et al. 2017; Muggia and Grube, 2018; Spribille et al., 2016; Voytsekhovich and Beck, 2016), were suggested to represent one of the most successful nutritional strategies among Fungi (Lücking et al., 2016). In contrast to the high diversity known for the lichen-forming fungi, about 19,000 species (27% of the known Ascomycota; Lücking et al., 2016), only about 200 species of photobionts, from a limited number of green algal and cyanobacterial genera, have formally been described (Muggia et al., 2018; Singh et al., 2017; Tschermak-Woess, 1988; Voytsekhovich et al., 2011). However, relationships and interactions among lichen symbionts are often masked by uncertain species boundaries or the inability to reliably identify symbionts.

Accurate taxonomic characterization of members of symbiotic guilds in symbioses is essential to properly describe the specificity patterns of these associations. Relationships among lichen symbionts are not random, as ecological, environmental and evolutionary factors influence symbiont associations (Beck et al., 2002; Chagnon et al., 2019; Grube et al., 2016; Leavitt et al., 2015; Miadlikowska et al., 2006). These intimate relationships among organisms involved in symbioses create conditions for strong interactions and have far-ranging implications (Chagnon et al., 2018; Magain et al., 2017; Rikkinen et al., 2002; Thrall et al., 2007). Interactions among potential lichen symbionts are frequently described in terms of availability, selectivity and

specificity (Beck et al., 2002; Yahr et al., 2004); each has been shown to have significant impacts on structuring lichen associations across evolutionary and ecological scales (Muggia et al., 2014; Steinova et al., 2019; Werth and Sork, 2014). This is particularly evident when the diverse reproductive strategies of the lichen holobionts are considered and results in varying patterns of fungal-algal specificity (Beck et al., 2019; Cao et al., 2015; Fedrowitz et al., 2011; Pardo De la Hoz et al., 2018; Rolshausen et al., 2018; Steinova et al., 2019; Wornik and Grube, 2010). Specificity in symbiotic systems can have an impact on community-level responses to disturbance and environmental change (Hester et al., 2016). In coral symbioses, patterns of coral-*Symbiodinium* associations are known to influence community stability under climate change (Fabina et al., 2013). Simulations of loss of coral species or the ability to engage in new symbiotic interactions have been performed to measure the community stability. These have shown that potential symbiotic unions are maximized by the survival of generalist hosts and symbionts and that compatible symbiotic assemblages increase the potential for local recolonization of certain niches (Fabina et al., 2013). Indeed, lichen distributions have also been hypothesized to be influenced by the ecological specialization and the physiological responses of the photobionts (e.g., Beck et al., 2002; Casano et al., 2011; Darnajoux et al., 2019; Lu et al., 2018; Magain et al., 2017; Peksa and Škaloud, 2011; Yahr et al., 2006). The fungal-photobiont specificity, however, can reflect also the specific habitat preference of a certain associations. This has been observed in those selectivity patterns described as geographic mosaics and detected in cyanolichens, in which there is higher selectivity locally than globally (Fedrowitz et al., 2011). Inconsistent and/or taxonomically biased resolution can impair cross-network comparisons and inferences of interaction patterns (Igns et al., 2007; Raffaelli, 2007).

A woefully incomplete understanding of the basic features of lichenized algae, including diversity and evolutionary relationships, currently limits our ability to effectively study these systems in a broader context. This is often due to the limited availability of lichen material and the constraints which promote the research on those associations represented by the most conspicuous macrolichen lineages. The lichen-forming photobiont (chlorobiont) genus *Trebouxia* Puymaly (Trebouxiophyceae, Chlorophyta) comprises coccoid, unicellular green algae, reproducing primarily by autospores (Fig. 1 A-C), that occur in diverse terrestrial ecosystems. The genus is seldom found free living (Bubrick et al., 1984; Yung et al., 2014) and associates with an estimated 80% of lichen-forming fungi in temperate regions and more than 20% of all lichen-forming fungi worldwide – far more than any other lichenized algal genus (Rambold et al.,

1998). *Trebouxia* cells contain a massive, axial, lobed chloroplast with usually one pyrenoid (Fig. 1 D-F; Archibald, 1975; Friedl, 1989). Differences in chloroplast lobe patterns and pyrenoid structures, have traditionally been used as diagnostic phenotypic traits for species differentiation and generic delimitation (as reviewed in Muggia et al., 2018). The genus *Trebouxia* is one of the three major lineages, together with *Chlorella* and *Oocystis*, into which the class Trebouxiophyceae is subdivided by molecular approaches (Fučíková et al., 2014; Leliaert et al., 2012; Lemieux et al., 2014; Martínez-Alberola et al., 2019); to date, it comprises 29 formally described species (Guiry and Guiry, 2017).

However, studies consistently indicate that the vast majority of species-level lineages in *Trebouxia* still lack formal description, and new, previously unrecognized lineages have been frequently reported (Leavitt et al., 2015, 2016; Muggia et al., 2014; Sadowska-Deś et al., 2014). For the vast majority of these species-level lineages, evolutionary relationships, biogeographic and ecological patterns, physiological properties, specificity, and selectivity have not yet been characterized. Although ongoing interest in understanding diversity and diversification in *Trebouxia* has resulted in an improved perspective into evolutionary relationships and diversity (Dal Grande et al., 2014; Helms, 2003; Muggia et al., 2014, 2018; Sadowska-Deś et al., 2014; Voytsekhovich and Beck, 2016), a comprehensive genus-wide multi-locus phylogeny has not been presented yet for this important group of symbiotic algae. Here, we provide a genus-wide, multi-locus phylogenetic hypothesis for *Trebouxia*. We assembled the most comprehensive taxon sampling including representatives of formally described species while identifying new monophyletic lineages that potentially merit formal species recognition.

2. Materials and Methods

2.1 Taxon sampling

We assembled a representative sampling of the currently known genetic diversity in the lichenized green algal genus *Trebouxia*. Data were compiled from a range of sequences from previously published studies, axenic algal cultures, and lichens collected from a poorly sampled region in East Africa. Sequence data were compiled from previous studies or newly generated for up to three genetic markers: the nuclear ribosomal internal transcribed spacer region (ITS), a fragment of the chloroplast gene for the large subunit of the ribulose-bisphosphate carboxylase (*rbcL*), and a fragment of the mitochondrial cytochrome oxidase subunit 2 (*cox*2) gene. Data from previously published studies included: (i) up to five individuals representing each of the

putative 67 species-level lineages circumscribed in Leavitt et al. (2015); (ii) ITS and *cox*2 sequences from *Trebouxia* associating with the mycobiont genus *Bryoria* (Parmeliaceae; Lindgren et al., 2014); (iii) ITS and *rbcL* sequences from *Trebouxia* associating with members of Teloschistaceae (Nyati et al., 2014); (iv) ITS, *cox*2, and *rbcL* sequences from *Trebouxia* associating with the mycobiont taxon *Umbilicaria pustulata* (Umbilicariaceae; Sadowska-Deś et al., 2014); (v) ITS and *rbcL* sequences from *Trebouxia* associating with the mycobiont genus *Tephromela* (Mycoblastaceae; Muggia et al., 2014); (vi) ITS sequences from *Trebouxia* associating with soil-crust forming lichens (Ruprecht et al., 2014); (vii) twenty-six ITS sequences generated from UTEX (http://www.utex.org) and SAG (http://www.uni-goettingen.de/en/184982.html) cultures, representing 20 of the 29 currently described *Trebouxia* species; and (viii) ITS sequences representing three recently described species, *T. cretacea* Voytsekhovich & Beck, *T. solaris* Voytsekhovich & Beck, and *T. vagua* Voytsekhovich & Beck (Voytsekhovich and Beck, 2016). New *Trebouxia* sequences were generated from 144 lichen specimens representing the family Parmeliaceae collected in East Africa (Supplementary Table S1).

2.2 DNA extraction, amplification and sequencing

For all new sequences generated for this study, total genomic DNA was extracted from a small piece of thallus material free from visible damage or contamination by symptomatic fungi or superficial algae using the USB PrepEase Genomic DNA Isolation Kit (USB, Cleveland, OH – product discontinued) following the manufacturer's recommendations. Alternatively, DNA was extracted from axenic *Trebouxia* cultures following the C-TAB protocol of Cubero et al. (1999). In this case the algal colonies were taken from the culture with a sterile inoculating loop and placed in a 1.5 ml tube, frozen and grinded with beads for three minutes using a bead-beater. We attempted to generate multi-locus sequence data from up to five individuals representing each candidate *Trebouxia* species (see section 2.4), although we used the maximum number of samples available in cases where candidate species were comprised of fewer than five individuals. The primers ITS1T and ITS4T (Kroken and Taylor, 2000) were used to amplify the algal ITS region (ITS1, 5.8S, ITS2); the *rbcL* was amplified using primers rbcL151f and rbcL986R (Nelsen et al., 2011); and primers COXIIf2 and COXIIr to amplify the *cox*2 locus (Lindgren et al., 2014). Ready-To-Go PCR Beads (GE Healthcare) were used for all polymerase chain reactions (PCR), following the manufacturer's recommendations. PCR cycling parameters

followed a 66–56 °C touchdown cycle described in Lindblom and Ekman (2006). All products were cleaned using ExoSAP-IT (USB, Cleveland, Ohio, USA) following the manufacturer's instructions, and complementary strands were sequenced from cleaned PCR products with the same primers used for amplification. Sequencing reactions were performed using BigDye v3.1 (Applied Biosystems, Foster City, CA, USA), and products were then run on an ABI 3730 automated sequencer (Applied Biosystems) at the Pritzker Laboratory for Molecular Systematics and Evolution at the Field Museum, Chicago, IL, USA. New sequences were assembled and edited using Sequencher v4.10 (Gene Codes Corporation, Ann Arbor, MI), and sequences with more than two ambiguous nucleotides were excluded.

2.3. Sequence data analyses

An initial multiple sequence alignment (MSA) for all *Trebouxia* ITS data assembled for this study was performed using the program MAFFT v7 (Katoh et al., 2005; Katoh and Toh, 2008). We implemented the FFT-NS-i alignment algorithm and '200PAM / K=2' scoring matrix, with an offset value of 0.0, unalignlevel = 0.4 using the 'Leave gappy regions' setting, and the remaining parameters set at default values. Based on the results of the initial alignment and exploratory phylogenetic reconstructions, sequences were divided into four groups corresponding to the four major *Trebouxia* clades – clade 'A', clade 'C' (previously provisionally called clade 'G'; Leavitt et al., 2015), clade 'I' and clade 'S'. Clade-specific MSAs of ITS sequences were performed in MAFFT using the G-INS-I alignment algorithm. The overall quality of the ITS sub-alignments was assessed using the Guidance2 web server (http://guidance.tau.ac.il/ver2/), and an overall reliability score was calculated for each ITS MSA (Sela et al., 2015). GUIDANCE2 alignments were performed using the MAFFT MSA algorithm with 'localpair' pairwise alignment strategy and 'Max-Iterate' set to 100 bootstrap replicates. Two separate MSAs were generated using GUIDANCE2: one masking residues with confidence scores <0.95 and the second masking columns with confidence scores <0.95. Multiple sequence alignments of the protein-coding *cox*2 and *rbcL* sequences were relatively straightforward, and both regions were aligned in MAFFT using the G-INS-I alignment algorithm. All alignments were deposited in FigShare (https://doi.org/10.6084/m9.figshare.12049086).

2.4 Assigning sequences to candidate species

We used the Automatic Barcode Gap Discovery method (ABGD; Puillandre et al., 2012) and a Bayesian implementation of the coalescent-based General Mixed Yule Coalescent (bGMYC; Reid and Carstens, 2012) to infer candidate species from the ITS sequence alignments (see Leavitt et al., 2015). ABGD infers a model-based confidence limit for intraspecific divergence, and then detects the barcode gap as a first significant gap beyond this limit to infer primary partitions. The primary data partitions are then recursively split to get finer partitions using the same approach until no further gaps can be detected (Puillandre et al., 2012). We used the ABGD web server (http://wwwabi.snv.jussieu.fr/public/abgd/abgdweb.html) and ABGD analyses were independently performed for each MSA of the four main *Trebouxia* clades (Table 1). Genetic distances were calculated for each subclade ITS alignment using the JC69 model, and other model parameters in ABGD were set using default parameter values, with the exception of the P_{max} value: P_{min} = 0.001, P_{max} = 0.01, steps = 100, and Nb bins = 100. Puillandre et al. (2012) suggested that implementing a Pmax value of 0.01 provides the most accurate estimate for the number of groups based on empirical comparisons of groupings inferred using ABGD with data from previous studies where species groups are well-characterized. We implemented a range of values for the gap width (X), beginning with X =1.5 and decreasing at 0.1 intervals until the first shift from one group to multiple groups was observed. Additional ABGD analyses implementing X values between 0.1 and 1.5 were run to assess the consistency of the inferred groups under varying gap width values. ABGD analyses were run for the complete ITS alignment generated using MAFFT and for GUIDANCE alignments after removing residues and columns below 0.95 to compare impact of different alignment strategies. Names for previously recognized candidates species followed Leavitt et al. (2015), and new candidate species-level OTUs circumscribed using ABGD were provisionally named following the same criterion, e.g., annotated with a clade-specific prefix and followed by the next available number. Provisional names were replaced by formal taxonomic nomenclature in cases where ABGD-delimited OTUs included sequences representing SAG or UTEX cultures or ITS sequences representing the three species *T. solaris*, *T. vagua* and *T. cretacea* (Voytsekhovich and Beck, 2016).

For the bGMYC analyses, ultrametric ITS gene trees were estimated from each ITS matrix representing clade ‘A’, ‘C’, ‘I’ and ‘S’, independently, using the program BEAST v1.8.0 (Drummond et al., 2012; Drummond and Rambaut, 2007). We ran two independent Markov Chain Monte Carlo (MCMC) chains for 25 million generations, implementing a relaxed lognormal clock and a constant coalescent speciation process prior. Default values were used for

the remaining priors. Chain mixing and convergence were evaluated in Tracer v1.5 (Rambaut and Drummond, 2003), considering effective sample size (ESS) values > 200 as a good indicator; and the first 5 million generations were discarded as burn-in. As input for the bGMYC analysis, a total of 100 trees were evenly sampled from the throughput of the posterior distribution of trees (50 trees from each independent run). The starting number of species was set to half the total number of tips, and the remaining settings were left at their default values. The bGMYC package was run in R v3.0.2 (R Core Team, 2014) for 50,000 generations, sampling every 200th generation, and discarding the first 40,000 trees as burn-in.

2.5 Phylogenetic reconstructions

Phylogenetic relationships were inferred from concatenated multi-locus sequence data for each of the four *Trebouxia* subclades and for a reduced data matrix representing the entire genus. A total of 870 specimens representing the range of genetic variation observed in the exploratory ITS phylogeny, including those represented by additional loci, were selected for phylogenetic analyses. Specimens were selected to minimize the amount of missing data in the final, reduced dataset. Clades 'A' and 'S' were comprised of specimens represented by up to three loci, ITS, *rbcL* and *cox*2. However, we were not able to consistently amplify the *cox*2 locus for clades 'C' and 'I', and the few successfully amplified sequences were often recovered with spurious relationships in exploratory analyses. Therefore, clades 'C' and 'I' were represented only by the ITS and *rbcL* markers. To minimize bias from ambiguously aligned regions of the ITS MSA, we used the GUIDANCE2 alignment excluding columns with GUIDANCE2 scores < 95%, which was concatenated with the *rbcL*, and *cox*2 MSAs (Table 2). We used the program RAxML v8.1.11 (Stamatakis, 2006; Stamatakis et al., 2008) to reconstruct maximum likelihood (ML) phylogenies from the four clade-specific alignments and a reduced genus-wide dataset (n=399). Sub-alignments of the ITS region from each of the four *Trebouxia* clades were combined using the 'merge' feature in MAFFT using the E-INS-I algorithm. We used the GTRGAMMA model, which includes a parameter (Γ) for rate heterogeneity among sites and chose not to include a parameter for estimating the proportion of invariable sites following the authors' recommendation, when the datasets were partitioned by locus. For each data matrix, a search combining 200 separate ML searches was conducted to find the optimal tree, and 1000 'fast bootstrap' replicates were performed to evaluate nodal support. The concatenated ITS, *cox*2, and *rbcL* dataset was partitioned by marker, and substitution models for each were selected using

jModelTest v2.1.1 (Darriba et al., 2014) using the Corrected Akaike Information Criterion (AICc) (Sugiura, 2007)

3. Results and Discussion

3.1 Species recognition

In the present study, between 109–113 candidate species distributed across the four main *Trebouxia* lineages were circumscribed using the ABGD approach (Table 1). The candidate species inferred using ABGD were largely consistent with those inferred by using bGMYC (Figs. 2–5) and consistent between MAFFT- and GUIDANCE-based alignments of the ITS sequence data (Table 1). The species-level diversity inferred here represents an approximate four-fold increase from the formally described species within *Trebouxia*. The absence of a formal, phylogeny-based taxonomy for the bulk of *Trebouxia* isolates and sequences has resulted in the application of provisional names. These names are associated with varying levels of phylogenetic diversity, and have been inconsistently applied across studies, thereby limiting the comparison, synthesis and effective communication in the scientific community (Muggia et al., 2018). Given the overall prominence of *Trebouxia*-associated lichens, this 'classification gap' is particularly problematic and limits our understanding of species interactions in lichen symbioses.

The detected classification gap is further exacerbated by the substantial lack of culture isolates for the majority of the detected species-level lineage. The analysis of morpho-anatomical traits in axenically cultured photobionts is in fact key for a complete reliable definition of species. This drawback has been overcome in the sister genus to *Trebouxia*, *Asterochloris*, for which species delimitation and formal taxonomy has been appropriately integrated using genetic, ecological and morphological analyses of environmental samples, and morpho-anatomical analyses of reference cultured strains (Skaloud et al., 2015). The study on *Asterochloris* also demonstrated that a clear phylogenetic characterization is essential for the correct estimation of species diversity as morphologically and ecologically diverse species can frequently be grouped into a single taxonomic unit (Skaloud et al. 2015).

A recently proposed provisional naming system for *Trebouxia* has improved effective communication and consistency among subsequent studies (Leavitt et al., 2015; Moya et al., 2017), and has been applied here across the entire genus. However, species-level lineages circumscribed within an integrative framework, additionally complemented with corresponding

algal cultures, will increase our ability to investigate and explain patterns of symbiotic interactions in lichens. Previous research with *Trebouxia* highlighted the importance of recognizing species-level diversity to understand interactions among lichen symbionts and their evolutionary relationships, and the ecological and biogeographic factors which trigger the establishment of certain, specific fungal-algal associations (Chen et al., 2016; Dal Grande et al., 2012, 2014; Leavitt et al., 2015, 2016; Nelsen and Gargas, 2009; Piercey-Normore and DePriest, 2001; Werth and Sork, 2014).

3.2 Species diversity of Trebouxia

Our multi-locus phylogenetic reconstructions recovered four major clades within *Trebouxia* (Fig. 6; Supplementary Fig. S1), which are named after the species characterized by certain major morphoanatomical differences of the chloroplast and the pyrenoid type, i.e. clade 'A' *arboricola/gigantea*-type, clade 'C' *corticola*-type, clade 'I' *impressa/gelatinosa*-type and clade 'S' *simplex/jamesii*-type. Clade 'C', the '*corticola*-type' clade, was formerly and provisionally recognized as clade 'G', the '*galapagensis*/*usneae*' clade (Helms, 2003; Leavitt et al., 2015); however, it is here recognized as clade 'C', to better reflect the *corticola*-type pyrenoid common among its members (Beck, 2002; Friedl, 1989). These major clades correspond to previously recognized clades recovered using the ITS data alone (e.g. Cordeiro et al., 2005; Helm, 2003; Leavitt et al., 2015; Molins et al., 2018; Muggia et al., 2014; Nyati et al., 2014; Pérez-Ortega et al., 2012; Voytsekhovich and Beck, 2016;). Here, clade 'A' comprises about 50 species-level lineages, in contrast to clade 'I' in which only 14 lineages are present.

Clade 'A' encompasses the majority of previously recognized *Trebouxia* species and includes most axenically isolated taxa characterized for their physiological traits (as revised by Muggia et al., 2018). Alternatively, the highest proportion of previously unrecognized species-level diversity was identified within the clade 'C' *corticola*-type, in which our sampling of Parmeliaceae-associated *Trebouxia* in East Africa resulted in 22 to 23 previously unsampled species-level clades. This confirms that there is a high level of the still unknown, potentially high genetic diversity of *Trebouxia* in geographic regions which are so far poorly studied.

While previous studies support the combination of the ITS, *rbcL*, and *cox*2 loci for phylogenetic reconstructions in *Trebouxia* (Lindgren et al., 2014; Sadowska-Deś et al., 2014), multi-locus sequence data generated from total genetic DNA from lichen thalli has limitations. Indeed, multiple *Trebouxia* species have been reported in individual thalli of several lichen

species (Casano et al., 2011; Del Campo et al., 2010, 2012; Dal Grande et al., 2018; Muggia et al., 2014). This problem might be mitigated by the fact that Sanger sequencing consistently yields the most abundant algal sequence in lichen thalli, especially when the intrathalline diversity is skewed towards the predominant photobiont (Paul et al., 2018). Intrathalline algal diversity has been correlated with the different physiological properties tested for in *in vitro* culture conditions (Casano et al., 2011). However, it is unclear whether this might be correlated with the lichen growth form, which could potentially influence the capacity of the mycobiont to acquire different photobionts during thallus establishment and later to retain different photobionts in certain parts of the thallus or during certain periods of its life. Moya et al. (2017) showed evidence for algal zonation in the branching thallus of the lichen *Ramalina farinacea*, although more data is required to validate this trend.

Issues related to the utility of ITS region as a DNA barcode and central molecular marker for species delimitation may also bias our perspective of diversity in *Trebouxia*. Intragenomic variability of the ITS region has been reported for *Trebouxia* photobionts (Muggia et al., 2014). The potential for various copies of the ITS within a genome may also hamper an accurate assessment of the major photobiont in lichens. Furthermore, ambiguities in alignments of highly variable regions are reflected in the inferred topology of the phylogeny and may limit reliable species delimitation using a phylogenetic species criterion. The approach used here does delimit potential species based on their genetic diversity but at the same time stresses the importance to integrate these data with morphological, physiological and metabolic ones. Indeed, strains of *Trebouxia* might be genetically identical but show rather different ecophysiological responses, as was demonstrated for two strain isolated form Antarctic lichens under desiccation stress (Sadowsky et al., 2016). Thus, intraspectific variation in ecologically relevant traits may lead to an underestimation of *Trebouxia* functional diversity using ITS-based taxonomy alone.

In many cases, primer biases or availability of resources have not allowed the sequencing of multiple loci for the same lichen individual, and this has led to the disproportionate amount of sequences available for the ITS locus compared with *rbcL*, *cox*2 and actin type I genes. Consequently, multigene phylogenies for the entire genus *Trebouxia* are lacking. Multigene phylogenetic studies have treated only subsets of *Trebouxia* species or have investigated the population structure of single *Trebouxia* species, reported for a few lichens (Català et al., 2016; Chen et al., 2016; Kroken and Taylor, 2000; Nyati et al., 2014; Werth and Sork, 2010, 2014). This is exemplified by clades ‘C’ and ‘I’ which totally lack *cox*2 sequences despite multiple

sequencing attempts. Notwithstanding these missing data, the *Trebouxia* phylogeny reported here (Fig. 6; Supplementary Fig. S1) and the detailed clade-specific phylogenies (Figs. 2–5) are proposed as background references for future studies aiming at resolving the phylogenetic positions of newly discovered lineages within this genus.

Phylogenomic data generated from pure algal cultures, rather than total DNA extract from intact lichen thalli, will be essential to reconstruct robust phylogenies with confidence (thereby avoiding amplification of co-existing algae). At the same time, axenic algal cultures will provide a crucial resource for investigating morphological, developmental, and physiological properties of species-level lineages.

4. Conclusions

Despite the vital significance of algae associated in symbioses, disentangling the specific processes structuring interactions among symbionts is presently limited due to an incomplete perspective of diversity, inability to accurately identify microbial partners, and limited human resources. Elucidating and identifying species involved in symbiotic interactions is crucial to understanding highly integrated associations which can be major factors determining ecological and evolutionary dynamics (Darwell et al., 2014). Mounting evidence indicates that, in lichen symbioses in spite of the obligate associations of the major fungal and algal symbionts, these interactions are much more complex and malleable than simple co-diversification hypotheses (e.g., Chagnon et al., 2019; Nelsen and Gargas, 2008; Singh et al., 2017).

Here, the genus-wide phylogeny and individual-clade topologies for *Trebouxia* highlight the current state of understanding of this important symbiotic algal genus, while revealing high levels of previously hidden diversity. We believe that robust hypotheses of evolutionary relationships are central for a more complete characterization of diversification and species boundaries in *Trebouxia*, in particular within the dynamics of lichen symbioses. The phylogenetic hypothesis reported here will stimulate the development of an integrative taxonomy for this genus, while providing an evolutionary, reference framework for future studies to investigate photobiont diversity and diversification dynamics in symbiotic algae, among other research avenues. We anticipate that integrative approaches to understanding symbiotic species evolution in lichens will serve as reference accessions for microbial ecologists, while continuing to generate interest in utilizing these organisms as models and systems for the study of symbiotic interactions.

Acknowledgements

SPO is supported by the Spanish Ministry of Science, Innovation and Universities through a 'Ramón y Cajal' contract (RYC-2014-16784).

References

Althoff, D.M., Segraves, K.A., Johnson, M.T.J., 2014. Testing for coevolutionary diversification: linking pattern with process. Trends Ecol. Evol. 29, 82–89.

Archibald P.A. 1975. *Trebouxia* de Pulmaly (Chlorophyceae, Chlorococcales) and *Pseudotrebouxia* gen. nov. (Chlorophyceae, Chlorosarcinales). Phycologia 14, 125–137.

Arnold, A. E., Miadlikowska, J., Higgins, K. L., Sarvate, S. D., Gugger, P., Way, A., Hofstetter, V., Kauff, F. and Lutzoni, F., 2009. A phylogenetic estimation of trophic transition networks for ascomycetous fungi: Are lichens cradles of symbiotrophic fungal diversification? Syst. Biol. 58, 283–297.

Beck, A., 2002. Selektivität der Symbionten schwermetalltoleranter Flechten. PhD thesis, Ludwig-Maximilians-Universität München, Munich, ISBN: 3-9808102-0-8.

Beck, A., Bechteler, J., Casanova-Katny, A., Dzhilyanova, I., 2019. The pioneer lichen *Placopsis* in maritime Antarctica: Genetic diversity of their mycobionts and green algal symbionts, and their correlation with deglaciation time. Symbiosis 79, 1–24.

Beck, A., Kasalicky, T., Rambold, G., 2002. Myco-photobiontal selection in a Mediterranean cryptogam community with *Fulgensia fulgida*. New Phytol. 153, 317–326.

Bubrick, P., Galun, M., & Frensdorff, A., 1984. Observation on free living *Trebouxia* Archibald, and evidence that both symbionts from *Xanthoria parietina* (L.) Th. Fr. Can be found free living in nature. New Phytol. 97, 455–462.

Cao, S., Zhang, F., Liu, C., Hao, Z., Tian, Y., Zhu, L., Zhou, Q., 2015. Distribution patterns of haplotypes for symbionts from *Umbilicaria esculenta* and *U. muehlenbergii* reflect the importance of reproductive strategy in shaping population genetic structure. BMC Microbiol. 15, 212.

Casano, L.M., del Campo, E.M., García-Breijo, F.J., Reig-Armiñana, J., Gasulla, F., del Hoyo, A., Guéra, A., Barreno, E., 2011. Two *Trebouxia* algae with different physiological performances are ever-present in lichen thalli of *Ramalina farinacea*. Coexistence versus Competition? Environ. Microbiol. 13, 806–818.

Català S., del Campo E.M., Barreno E., García-Breijo F.J., Reig-Armiñana J., Casano L.M., 2016. Coordinated ultrastructural and phylogenomic shed light on the hidden phycobiont diversity of the *Trebouxia* microalgae in *Ramalina fraxinea*. Mol. Phylogenet. Evol. 94, 765–777.

Chagnon, P. L., Magain, N., Miadlikowska, J., and Lutzoni, F., 2018. Strong specificity and network modularity at a very fine phylogenetic scale in the lichen genus *Peltigera*. Oecologia 187, 767–782.

Chagnon, P. L., Magain, N., Miadlikowska, J., and Lutzoni, F., 2019. Species diversification and phylogenetically constrained symbiont switching generated high modularity in the lichen genus *Peltigera*. J. Ecol. 107, 1645-1661.

Chen, J. M., Werth, S., Sork, V. L., 2016. Comparison of phylogeographical structures of a lichen forming fungus and its green algal photobiont in western North America. J. Biogeogr. 43, 932–943.

Cordeiro, L.M.C, Reis, R.A., Cruz, L.M., Stöcker-Wörgötter, E., Grube, M., Iacomin, M., 2005. Molecular studies of photobionts of selected lichens from the coastal vegetation of Brasil. FEMS Microbiol. Ecol. 54, 381–390.

Cubero, O.F., Crespo, A., Fatehi, J., Bridge, P.D., 1999. DNA extraction and PCR amplification method suitable for fresh, herbarium stored and lichenized fungi. Plant Syst. Evol. 217, 243–249.

Dal Grande, F., Beck, A., Cornejo, C., Singh, G., Cheenacharoen, S., Nelsen, M.P., Scheidegger, C., 2014. Molecular phylogeny and symbiotic selectivity of the green algal genus *Dictyochloropsis* s.l. (*Trebouxiophyceae*): a polyphyletic and widespread group forming photobiont-mediated guilds in the lichen family Lobariaceae. New Phytol. 202, 455–470.

Dal Grande, F., Rolshausen, G., Divakar, P.K., Crespo, A., Otte, J., Schleuning, M., Schmitt, I., 2018. Environment and host identity structure communities of green algal symbionts in lichens. New Phytol. 217, 277–289.

Dal Grande, F., Widmer, I., Wagner, H.H., Scheidegger, C., 2012. Vertical and horizontal photobiont transmission within populations of a lichen symbiosis. Mol. Ecol. 21, 3159–3172.

Darnajoux, R., Magain, N., Renaudin, M., Lutzoni, F., Bellenger, J.-P., Zhang, X. 2019. Molybdenum threshold for ecosystem scale alternative vanadium nitrogenase activity in boreal forests. Proc. Natl. Acad. Sci. U.S.A. 116, 24682–24688.

Darriba, D., Taboada, G. L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9, 772.

Darwell, C.T., al-Beidh, S., Cook, J.M., 2014. Molecular species delimitation of a symbiotic fig-pollinating wasp species complex reveals extreme deviation from reciprocal partner specificity. BMC Evol. Biol. 14, 1–10.

del Campo E.M., Catalá S., Gimeno J., del Hoyo A., Martínez-Alberola F., Casano L.M., Grube M., Barreno E., 2012. The genetic structure of the cosmopolitan three-partner lichen *Ramalina farinacea* evidences the concerted diversification of symbionts. FEMS Microbiol. Ecol. 83, 310–323.

del Campo E.M., Gimeno J., de Nova J.P.G., Casano L.M., Gasulla F., Garcia-Breijo F., Reig-Arminana J., Barreno E., 2010. South European population of *Ramalina farinacea* (L.) share different *Trebouxia* algae. Bibl. Lichenol. 105, 247–256.

Fabina N.S., Putnam H.M., Franklin E.C., Stat M., Gate R., 2013. Symbiotic specificity, association patterns, and function determine community responses to global changes: defining critical research areas for coral-*Symbiodinium* symbioses. Global Change Biol. 19, 3306–3316.

Fedrowitz, K., Kaasalainen, U., Rikkinen, J., 2011. Geographic mosaic of symbiont selectivity in a genus of epiphytic cyanolichens. Ecol. Evol. 2: 2291–2303.

Friedl, T., 1989. Comparative ultrastructure of pyrenoids in *Trebouxia* (Microthamniales, Chlorophyta). Plant Syst. Evol. 164, 145-159.

Fučíková K., Leliaert F., Cooper E.D., Škaloud P., D'Hondt S., De Clerck O., Gurgel C.F.D, Lewis L.A., Lewis P.O., López-Bautista J.M., Delwiche C.F., Verbruggen H., 2014. New phylogenetic hypotheses for the core Chlorophyta based on chloroplast sequence data. Front. Ecol. Evol. 2, 63.

Gargas, A., DePriest, P.T., Grube, M., Tehler, A., 1995. Multiple origins of lichen symbioses in fungi suggested by SSU rDNA phylogeny. Science 268, 1492–1495.

Group, T.N.H.W., Peterson, J., Garges, S., Giovanni, M., McInnes, P., Wang, L., Schloss, J.A., Bonazzi, V., McEwen, J.E., Wetterstrand, K.A., Deal, C., Baker, C.C., Di Francesco, V., Howcroft, T.K., Karp, R.W., Lunsford, R.D., Wellington, C.R., Belachew, T., Wright, M., Giblin, C., David, H., Mills, M., Salomon, R., Mullins, C., Akolkar, B., Begg, L., Davis, C., Grandison, L., Humble, M., Khalsa, J., Little, A.R., Peavy, H., Pontzer, C., Portnoy, M., Sayre, M.H., Starke-Reed, P., Zakhari, S., Read, J., Watson, B., Guyer, M., 2009. The NIH Human Microbiome Project. Genome Res.19, 2317–2323.

Grube, M., Aschenbrenner, I., Cernava, T., Berg, G., 2016. Lichen–Bacterial Interactions. In: Druzhinina, I.S., Kubicek, C.P. (Eds.), Environmental and Microbial Relationships. Springer International Publishing, Cham, pp. 179–188.

Grube, M., Cardinale, M., de Castro, J.V.Jr., Muller, H., Berg, G., 2009. Species-specific structural and functional diversity of bacterial communities in lichen symbioses. ISME J. 3, 1105–1115.

Guiry, M.D., Guiry, G.M., 2017. AlgaeBase, http://www.algaebase.org/. National University of Ireland, Galway.

Helms, G., 2003. Taxonomy and symbiosis in associations of Physciaceae and *Trebouxia*. Inauguraldissertation am Albrecht-von-Haller Institut für Pflanzenwissenschaften, Experimentelle Phykologie und Sammlung von Algenkulturen der Georg-August-Universität Göttingen, Göttingen, pp. 1–156.

Hester, E.R., Barott, K.L., Nulton, J., Vermeij, M.J., Rohwer, F.L., 2016. Stable and sporadic symbiotic communities of coral and algal holobionts. The ISME J. 10, 1157–1169.

Honegger, R., 2009. Lichen-Forming Fungi and Their Photobionts. In: Deising, H. (Ed.), Plant Relationships. Springer Berlin Heidelberg, pp. 307–333.

Ings, T.C., Montoya, J.M., Bascompte, J., Blüthgen, N., Brown, L., Dormann, C.F., Edwards, F., Figueroa, D., Jacob, U., Jones, J.I., Lauridsen, R.B., 2009. Ecological networks–beyond food webs. J. Animal Ecol. 78, 253–69.

Katoh, K., Kuma, K.-i., Toh, H., Miyata, T., 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res. 33, 511–518.

Katoh, K., Toh, H., 2008. Recent developments in the MAFFT multiple sequence alignment program. Briefings Bioinform.9, 286–298.

Kroken, S., Taylor, J.W., 2000. Phylogenetic species, reproductive mode, and specificity of the green alga *Trebouxia* forming lichens with the fungal genus *Letharia*. Bryologist 103, 645–660.

Leavitt, S.D., Kraichak, E., Nelsen, M.P., Altermann, S., Divakar, P.K., Alors, D., Esslinger, T.L., Crespo, A., Lumbsch, H.T., 2015. Fungal specificity and selectivity for algae play a major role in determining lichen partnerships across diverse ecogeographic regions in the lichen-forming family Parmeliaceae (Ascomycota). Mol. Ecol. 24, 3779–3797.

Leavitt, S.D., Kraichak, E., Vondrak, J., Nelsen, M.P., Sohrabi, M., Perez-Ortega, S., St Clair, L.L., Lumbsch, H.T., 2016. Cryptic diversity and symbiont interactions in rock-posy lichens. Mol. Phylogenet. Evol. 99, 261–274.

Leliaert, F., Smith, D.R., Moreau H., Herron, M.D., Verbruggen, H., Delwiche, C.F., De Clerck, O., 2012. Phylogeny and molecular evolution of the green algae. Crit. Rev. Plant Sci. 31, 1–46.

Lemieux, C., Otis, C., Turmel, M., 2014. Chloroplast phylogenomic analysis resolves deep-level relationships within the green algal class Trebouxiophyceae. BMC Evol. Biol.14, 211.

Lindgren, H., Velmala, S., Högnabba, F., Goward, T., Holien, H., Myllys, L., 2014. High fungal selectivity for algal symbionts in the genus *Bryoria*. Lichenologist 46, 681–695.

Lu, J., Magain, N., Miadlikowska, J., Coyle, J. R., Truong, C., and Lutzoni, F., 2018. Bioclimatic factors at an intrabiome scale are more limiting than cyanobiont availability for the lichen-forming genus *Peltigera*. Am. J. Bot. 105, 1198–1211.

Lücking, R., Hodkinson, B.P., Leavitt, S.D., 2016. The 2016 classification of lichenized fungi in the Ascomycota and Basidiomycota – Approaching one thousand genera. Bryologist 119, 361–416.

Magain, N., Miadlikowska, J., Goffinet, B., Sérusiaux, E., and Lutzoni. F., 2017. Macroevolution of specificity in cyanolichens of the genus *Peltigera* section Polydactylon (Lecanoromycetes, Ascomycota). Syst. Biol. 66, 74–99.

Margulis, L., Fester, R., 1991. Symbiosis as a source of evolutionary innovation: speciation and morphogenesis. MIT Press, Cambridge, MA.

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

Miadlikowska, J., Kauff, F., Hofstetter, V., Fraker, E., Grube, M., Hafellner, J., Reeb, V., Hodkinson, B. P., Kukwa, M., Lücking, R., Hestmark, G., Garcia Otalora, M., Rauhut, A., Büdel, B., Scheidegger, C., Timdal, E., Stenroos, S., Brodo, I., Perlmutter, G., Ertz, D., Diederich, P., Lendemer, J. C., May, P., Schoch, C. L., Arnold, A. E., Gueidan, C., Tripp, E., Yahr, R., Robertson, C., Lutzoni, F., 2006. New insights into classification and evolution of the Lecanoromycetes (Pezizomycotina, Ascomycota) from phylogenetic analyses of three ribosomal RNA- and two protein-coding genes. Mycologia 98, 1088–1103.

Moreira, D., Le Guyader, H., Philippe, H., 2000. The origin of red algae and the evolution of chloroplasts. Nature 405, 69–72.

Molins, A., Moya, P., Garcia-Breijo, F.J., Reig-Armiñana, J., Barreno, E., 2018. A multi-tool approach to assess microalgal diversity in lichens: isolation, Sanger sequencing, HTS and ultrastructural correlations. Lichenologist 50, 123–138.

Moya, P., Molins, A.A., Ânez-Alberola, F.M., Muggia, L., Barreno, E., 2017. Unexpected associated microalgal diversity in the lichen *Ramalina farinacea* is uncovered by pyrosequencing analyses. PloS ONE 12, e0175091.

Muggia, L., Grube, M., 2018. Fungal diversity in lichens: from extremotolerance to interaction with algae. Life 8, 15.

Muggia, L., Leavitt, S.D., Barreno, E., 2018. The hidden diversity of lichenized Trebouxiophyceae (Chlorophyta). Phycologia 57, 503–524.

Muggia, L., Pérez-Ortega, S., Kopun, T., Zellnig, G., Grube, M., 2014. Photobiont selectivity leads to ecological tolerance and evolutionary divergence in a polymorphic complex of lichenized fungi. Ann. Bot. 114, 463–475.

Nelsen, M.P., Gargas, A., 2008. Dissociation and horizontal transmission of codispersing lichen symbionts in the genus *Lepraria* (Lecanorales: Stereocaulaceae). New Phytol. 177, 264–275.

Nelsen, M.P., Gargas, A., 2009. Symbiont flexibility in *Thamnolia vermicularis* (Pertusariales: Icmadophilaceae). The Bryologist 112, 404–417.

Nelsen, M.P., Plata, E.R., Andrew, C.J., Lücking, R., Lumbsch, H.T., 2011. Phylogenetic diversity of Trentepohlialean algae associated with lichen-forming fungi. J. Phycol. 47, 282–290.

Nyati, S., Scherrer, S., Werth, S., Honegger, R., 2014. Green-algal photobiont diversity (*Trebouxia* spp.) in representatives of Teloschistaceae (Lecanoromycetes, lichen-forming ascomycetes). Lichenologist 46, 189–212.

Pardo De la Hoz, C. J., Magain, N., Lutzoni, F., Goward, T., Restrepo, S., and Miadlikowska, J. 2018. Contrasting symbiotic patterns in two closely related lineages of trimembered lichens of the genus *Peltigera*. Front. Microbiol. 9, 2770.

Paul, F., Otte, J., Schmitt, I., Dal Grande, F., 2018. Comparing Sanger sequencing and high-throughput metabarcoding for inferring photobiont diversity in lichens. Sci. Rep. 8, 8624.

Peay, K.G., 2016. The mutualistic niche: mycorrhizal symbiosis and community dynamics. Ann. Rev. Ecol. Evol. Syst. 47, 143–164.

Peksa, O., Škaloud, P., 2011. Do photobionts influence the ecology of lichens? A case study of environmental preferences in symbiotic green alga *Asterochloris* (Trebouxiophyceae). Mol. Ecol. 20, 3936–3948.

Pérez-Ortega, S., Ortiz-Álvarez, R., Green, T.A., de los Ríos, A., 2012. Lichen myco- and photobiont diversity and their relationships at the edge of life (McMurdo Dry Valleys, Antarctica). FEMS Microbiol. Ecol. 82, 429–448.

Piercey–Normore, M.D., DePriest, P.T., 2001. Algal switching among lichen symbioses. Am. J. Bot. 88, 1490–1498.

Puillandre, N., Lambert, A., Brouillet, S., Achaz, G., 2012. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. Mol. Ecol. 21, 1864–1877.

Raffaelli, D., 2007. Food webs, body size and the curse of the Latin binomial. In From energetics to ecosystems: the dynamics and structure of ecological systems (pp. 53-64). Springer, Dordrecht.

Reid, N., Carstens, B.C., 2012. Phylogenetic estimation error can decrease the accuracy of species delimitation: a Bayesian implementation of the general mixed Yule-coalescent model. BMC Evol. Biol. 12,196.

Rambold, G., Friedl, T., Beck, A., 1998. Photobionts in lichens: possible indicators of phylogenetic relationships. The Bryologist 101, 392–397.

Rikkinen, J., Oksanen, I. and Lohtander, K. 2002. Lichen guilds share related cyanobacterial symbionts. Science 297, 357.

Rolshausen, G., Dal Grande, F., Sadowska-Deś, A.D., Otte, J., Schmitt, I., 2018. Quantifying the climatic niche of symbiont partners in a lichen symbiosis indicates mutualist-mediated niche expansions. Ecography 41, 1380–1392.

Rowan, R.O.B., Powers, D.A., 1991. A molecular genetic classification of Zooxanthellae and the evolution of animal-algal symbioses. Science 251, 1348.

Ruprecht, U., Brunauer, G., Türk, R., 2014. High photobiont diversity in the common European soil crust lichen *Psora decipiens*. Biodivers. Conserv. 23, 1771–1785.

Sadowska-Deś, A.D., Dal Grande, F., Lumbsch, H.T., Beck, A., Otte, J., Hur, J.-S., Kim, J.A., Schmitt, I., 2014. Integrating coalescent and phylogenetic approaches to delimit species in the lichen photobiont *Trebouxia*. Mol.r Phylogenet. Evol. 76, 202–210.

Sadowsky, A., Mettler-Altmann, T., Ott, S., 2016. Metabolic response to desiccation stress in strains of green algal photobionts (*Trebouxia*) from two Antarctic lichens of southern habitats. Phycologia 55, 703–714.

Sela, I., Ashkenazy, H., Katoh, K., Pupko, T., 2015. GUIDANCE2: accurate detection of unreliable alignment regions accounting for the uncertainty of multiple parameters. Nucleic Acids Res. 43, W7–W14.

Singh, G., Dal Grande, F., Divakar, P.K., Otte, J., Crespo, A., Schmitt, I., 2017. Fungal-algal association patterns in lichen symbiosis linked to macroclimate. New Phytol. 214, 317–329.

Škaloud P., Steinova J., Ridka T., Vančurova L., Peksa O., 2015 Assembling the challenging puzzle of algal biodiversity: species delimitation within the genus *Asterochloris* (Trebouxiophyceae, Chlorophyta). J. Phycol, 51, 507–527.

Spribille, T., Tuovinen, V., Resl, P., Vanderpool, D., Wolinski, H., Aime, M.C., Schneider, K., Stabentheiner, E., Toome-Heller, M., Thor, G., Mayrhofer, H., Johannesson, H., McCutcheon, J.P., 2016. Basidiomycete yeasts in the cortex of ascomycete macrolichens. Science 353, 488.

Spribille, T., 2018. Relative symbiont input and the lichen symbiotic outcome. Curr. Opin. Plant Biol. 44, 57–63.

Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690.

Stamatakis, A., Hoover, P., Rougemont, J., 2008. A rapid bootstrap algorithm for the RAxML web servers. Syst. Biol. 57, 758–771.

Steinova, J., Skaloud, P., Bestová, H., Yahr, R., Muggia, L., 2019. Reproductive and dispersal strategies shape the diversity of mycobiontphotobiont association in *Cladonia* lichens. Mol. Phylogenet. Evol. 134, 226–237.

Sugiura, N. 2007. Further analysis of the data by Akaike's information criterion and the finite corrections. Commun Stat—Theory Methods 7, 13–26.

Thrall, P.H., Hochberg, M.E., Burdon, J.J., Bever, J.D., 2007. Coevolution of symbiotic mutualists and parasites in a community context. Trends Ecol. Evol. 22, 120–126.

Tschermak-Woess, E., 1988. New and known taxa of *Chlorella* (Chlorophyceae): Occurrence as lichen phycobionts and observations on living dictyosomes. Plant Syst. Evol. 159, 123–139.

Voytsekhovich, A., Beck, A., 2016. Lichen photobionts of the rocky outcrops of Karadag massif (Crimean Peninsula). Symbiosis 68, 9–24.

Voytsekhovich, A., Dymytrova, L:, Nadyeina, O., 2011. Photobiont composition of some taxa of the genera *Micarea* and *Plancynthiella* (Lecanoromycets, lichenized Ascomycota) from Ukraine. Folia Cryptog. Estonica 48, 135–148.

Werth S., Sork V., 2010. Identity and genetic structure of photobiont of the epiphytic lichen *Ramalina menziesii* on three oak species in Southern California. Am. J. Bot. 97, 821–830.

Werth, S., Sork, V.L., 2014. Ecological specialization in *Trebouxia* (Trebouxiophyceae) photobionts of *Ramalina menziesii* (Ramalinaceae) across six range-covering ecoregions of western North America. Am. J. Bot. 101, 1127–1140.

Wornik, S., Grube, G., 2010. Joint dispersal does not imply maintenance of partnerships in lichen symbioses. Microb. Ecol. 59, 150–157.

Yahr R., Vilgalys R., DePriest P.T., 2004. Strong fungal specificity and selectivity for algal symbionts in Florida scrub *Cladonia* lichens. Mol. Ecol. 13, 3367–3378.

Yahr R., Vilgalys R., DePriest P.T., 2006. Geographic variation in algal partners of *Cladonia subtenuis* (Cladoniaceae) highlights the dynamic nature of a lichen symbiosis. New Phytol. 171, 847–860.

Yung, C.C., Chan, Y., Lacap, D.C., Pérez-Ortega, S., de los Rios-Murillo, A., Lee, C.K., Cary, S.C., Pointing, S.B., 2014. Characterization of chasmoendolithic community in Miers Valley, McMurdo Dry Valleys, Antarctica. Microb. Ecol. 68, 351–359.

Zook, D., 2015. Symbiosis—Evolution's Co-Author. In: Gontier, N. (Ed.), Reticulate Evolution: Symbiogenesis, Lateral Gene Transfer, Hybridization and Infectious Heredity. Springer International Publishing, Cham, pp. 41–80.

Table 1. MAFFT alignment/residues removed (95%)/columns removed (95%) ABGD gap width (*X*) in parentheses ('ABGD_final_03June2019 data)

Algal clade	**No. of sequences**	**Total No. of candidate species**	**New**
'clade A' – *arboricola*/*gigantea* group	416	50(0.9)/48(0.9)/49(0.8)	15
'clade C' – *corticola*/*galapagensis/usneae* group	110	30(1.2)/29(0.9)/30(1.2)	20
'clade I' – *impressa*/*gelatinosa* group	146	14(0.5)/14(0.4)/13(0.5)	3
'clade S' – *simplex*/*letharii*/*jamesii* group	194	17(1.1)/22(0.3)/17(1.0)	14

Table 2. Summary of data remaining after GUIDANCE2 removal of ambiguous or difficult to align regions of the ITS multiple sequence data. Summaries based on confidence score for each residue-pair in the base MSA (residues) and for each column in the base MSA (columns). The overall GUIDANCE2 score reflects the robustness of the alignment to perturbations.

	Residues ≥ 0.95 remaining	Columns ≥ 0.95 remaining	GUIDANCE2 score
Clade A	96.6%	89.0%	0.988612
Clade G	95.4%	92.7	0.989769
Clade I	94.1%	93.2%	0.987630
Clade S	98.6%	97.4	0.995738

Figures captions

Figure 1. Morphological and anatomical traits of *Trebouxia* spp. A) Axenic culture of *Trebouxia impressa* (strain L1652); B) cultured algal cells of *Trebouxia* cf. *arboricola* (strain L1369) in light microscopy mounted in water: coccoid cells with a single central chloroplast; C) mature autospore of cultured *T.* cf. *arboricola* (strain L1379), microphoto in transmission electron microscopy (TEM); D) cultured *T.* cf. *arboricola* (strain L1379), mature cell with central, dividing chloroplast and pyrenoid, microphoto in TEM; E) chloroplast of *T. arboricola* (strain L1379) in three-dimensional view of a confocal laser scanning microscopy (CLSM) stack; F) pyrenoid of *gigantea*-type of cultured *Trebouxia* sp. (strain L1535), electron-dense spots are pyrenoglobules, short-branched thylakoids are delimiting the white spaces in the pyrenoid matrix. Scale bars: A) 2 mm, B) 15 μm, C-E) 5 μm, F) 0.5 μm.

Figure 2. Multi-locus phylogenetic inference of *Trebouxia* spp. belonging to clade 'A' – *arboricola/gigantea*. Bootstrap (BS) values ≥ 85% are indicated with thickened red branches; BS values ≥ 70 and < 85% with thickened orange branches; BS values ≥ 60% and < 70% with thickened yellow branches; and BS values < 60% with black branches. Candidate species recognized by the Automatic Barcode Gap Discovery method (ABGD) are highlighted at terminals. Corresponding results of bGMYC species delimitation, with the color scale is the probability scale of conspecificity (yellow = 0.95 to 1; orange = 0.5 to <0.95; red = < 0.5), are presented to the right of the topology. Provisional names 'TR1' (OTU 'A03') and 'TR9' (OTU 'A39') follow del Campo et al. (2010).

Figure 3. Multi-locus phylogenetic inference of *Trebouxia* spp. belonging to clade 'C' – *galapagensis*/*usneae*. Bootstrap (BS) values ≥ 85% are indicated with thickened red branches; BS values ≥ 70 and < 85% with thickened orange branches; BS values ≥ 60% and < 70% with thickened yellow branches; and BS values < 60% with black branches. Candidate species recognized by the Automatic Barcode Gap Discovery method (ABGD) are highlighted at terminals. Corresponding results of bGMYC species delimitation, with the color scale is the probability scale of conspecificity (yellow = 0.95 to 1; orange = 0.5 to <0.95; red = < 0.5), are presented to the right of the topology.

Figure 4. Multi-locus phylogenetic inference of *Trebouxia* spp. belonging to clade 'I' – *impressa*/*gelatinosa* Bootstrap (BS) values ≥ 85% are indicated with thickened red branches; BS values ≥ 70 and < 85% with thickened orange branches; BS values ≥ 60% and < 70% with thickened yellow branches; and BS values < 60% with black branches. Candidate species recognized by the Automatic Barcode Gap Discovery method (ABGD) are highlighted at terminals. Corresponding results of bGMYC species delimitation, with the color scale is the probability scale of conspecificity (yellow = 0.95 to 1; orange = 0.5 to <0.95; red = < 0.5), are presented to the right of the topology.

Figure 5. Multi-locus phylogenetic inference of *Trebouxia* spp. Belonging to clade 'S' – *simplex*/'*letharii*'/*jamesii*. Bootstrap (BS) values ≥ 85% are indicated with thickened red branches; BS values ≥ 70 and < 85% with thickened orange branches; BS values ≥ 60% and < 70% with thickened yellow branches; and BS values < 60% with black branches. Candidate species recognized by the Automatic Barcode Gap Discovery method (ABGD) are highlighted at terminals. Corresponding results of bGMYC species delimitation, with the color scale is the probability scale of conspecificity (yellow = 0.95 to 1; orange = 0.5 to <0.95; red = < 0.5), are presented to the right of the topology.

Figure 6. Simplified multi-locus phylogenetic inference of the lichen-forming algal genus *Trebouxia* based on the combined ITS, *cox*2 and *rbcL* datasets (n=399). The four major lineages, clades 'A', 'C', 'S', & 'I', are outlined. Bootstrap (BS) values ≥ 85% are indicated with thickened red branches; BS values ≥ 70 and < 85% with thickened orange branches; BS values ≥ 60% and < 70% with thickened yellow branches; BS values ≥ 50% and < 60% with thickened black branches; and BS <50% with grey branches.

Supplementary Material: captions to Figures and Tables

Table S1. ITS, *cox*2 and *rbcL* NCBI accessions included in the datasets of the phylogenetic analyses of Figures 2–6. Newly sequenced samples and the assigned NCBI accession numbers are highlighted in bold text.

Supplementary Fig. S1. Complete multi-locus phylogenetic inference of the lichen-forming algal genus *Trebouxia* based on the combined ITS, *cox*2 and *rbcL* datasets (n=399). Bootstrap

(BS) values ≥ 85% are indicated with thickened red branches; BS values ≥ 70 and < 85% with thickened orange branches; BS values ≥ 60% and < 70% with thickened yellow branches; BS values ≥ 50% and < 60% with thickened black branches. BS values < 50% are not shown.

- The vast majority of species-level lineages in *Trebouxia* lack formal description
- Four major *Trebouxia* lineages are recovered in our multi-locus phylogeny
- Integrative taxonomy will be essential for these important lichen photobionts
- Other under-sampled regions likely harbor additional species-level diversity

Lucia Muggia: Conceptualization, Methodology, Data curation, Data Analysis, Writing- Original draft preparation, Writing- Reviewing and Editing. **Matthew P. Nelsen:** Conceptualization, Methodology, Data Analysis, Writing- Reviewing and Editing. **Paul M. Kirika**: Specimen Sampling, Specimen Identification, Methodology, Writing- Reviewing and Editing. **Eva Barreno**: Methodology, Writing- Reviewing and Editing. **Andreas Beck**: Conceptualization, Methodology, Writing- Reviewing and Editing **Trebouxia working group**: Conceptualization, Methodology, Writing- Reviewing and Editing **Hanna Lindgren**: Methodology, Specimen Sampling, Data curation. **H. Thorsten Lumbsch**: Conceptualization, Methodology, Writing- Reviewing and Editing. **Steven D. Leavitt**: Conceptualization, Data curation, Data Analysis, Writing- Original draft preparation, Writing- Reviewing and Editing.