


2024

**TAXONOMIC REVISION OF THE SOIL GENUS TRICHOCOLEUS
(TRICHOCOLEUSACEAE, CYANOBACTERIA), INCLUDING FOUR
NEW SPECIES FROM WHITE SANDS NATIONAL PARK, NEW
MEXICO**

Amanda Szinte

Follow this and additional works at: <https://collected.jcu.edu/masterstheses>

 Part of the **Biology Commons**

TAXONOMIC REVISION OF THE SOIL GENUS *TRICHOCOLEUS*
(TRICHOCOLEUSACEAE, CYANOBACTERIA), INCLUDING FOUR NEW SPECIES
FROM WHITE SANDS NATIONAL PARK, NEW MEXICO

A Thesis Submitted to the
Graduate School of
John Carroll University
in Partial Fulfillment of the Requirements
for the Degree of
Master of Science

By
Amanda L. Szinte
2024

ABSTRACT

Cyanobacteria have never received molecular characterization in the gypsum soils at White Sands National Park. These gypsiferous soils support distinctive plant communities and considerable microbial taxonomic diversity in their biological soil crusts. Employing a polyphasic approach characterizing both morphology and sequencing of the 16S rRNA gene and associated 16S-23S ITS region, six strains of *Trichocoleus* isolated from White Sands biocrusts were fully characterized. Phylogenetic analyses including other as-yet-unnamed *Trichocoleus* strain sequences and the three previously described species, *T. desertorum*, *T. caatingensis*, and *T. badius*, led to a revision of the genus and identification of 17 unnamed species-level groups within *Trichocoleus*. Upon publication of this work, *Trichocoleus* will be one of the most species-rich cyanobacterial desert soil genera. *Trichocoleus* is a widespread genus in arid and semi-arid lands around the world, containing multiple species in a number of different desert regions, and will undoubtedly be found in future studies of arid land biocrusts.

KEY WORDS: *Trichocoleus*, Cyanobacteria, White Sands, Gypsum, Polyphasic Approach

INTRODUCTION

Arid Lands Cyanobacteria

Even though soils host a wide variety of major taxonomic groups (bryophytes, arthropods, annelids, fungi, bacteria, archaea) and hold over 25% of the world's biodiversity, there is still a lack of taxonomic knowledge of the pedosphere, especially in semi-arid and arid biomes. Cyanobacteria are essential components of arid soils, but are overall understudied (Johansen 1993, Belnap & Lange 2013). Several early papers that reported cyanobacterial species present in arid soils, but these were all based on morphology as the single criterion for identification (Cameron 1960, Novichkova-Ivanova 1980, Johansen et al. 1981, 1984, 1993, Ashley et al. 1985). Characterization of cyanobacteria using a polyphasic approach first occurred in 2002 (Boyer et al. 2002, Flechtner et al. 2002), papers which presaged the use of molecular data to characterize cyanobacteria from arid land soil crusts. Subsequently, numerous genera and species have been described from desert soils (Řeháková et al. 2007, Perkerson et al. 2011, Mühlsteinová et al. 2014a, 2014b, Osorio-Santos et al. 2014, Pietrasiak et al. 2014, 2019, Bohunická et al. 2015).

Biological soil crusts or biocrusts (BSC) hold intricate and complex links between cyanobacteria, eukaryotic algae, lichens, fungi, mosses, and bacteria that can provide homes for invertebrates and reptiles (Maestre et al. 2011, Pietrasiak et al. 2011, Bu et al. 2015). BSC were first recognized as distinctive soil features in the 1950's and have been called cryptogamic crusts, microphytic crusts, cryptobiotic crusts and microbiotic crusts since intensive research began in the 1980's in middle latitude deserts (early work reviewed in Johansen 1993, Evans and Johansen 1999). None of these terms were an exact fit for the components of the crusts, and finally the most generic term, biological soil crusts, became widely used and is the term most

applied today. BSC are found all over the globe but are most frequently found in semi-arid to arid environments (Maestre et al. 2011), which cover 35% of Earth's continental surface and over 70% of arid regions' surfaces (Belnap & Lange 2013). In arid regions, much of the biodiversity is found in BSC microbial communities, all within the uppermost 2 cm of the soil surface (Maestre et al. 2011, Bu et al. 2015).

Cyanobacterial research in the deserts and arid lands of the western United States is becoming more prevalent. Over 50 cyanobacterial species have been identified in North American deserts (Bu et al. 2015) and with further research, unquestionably, more will be discovered. Boyer et al. (2002) isolated putative *Microcoleus* strains from the Great Basin, Colorado Plateau, Mojave Desert and Chihuahuan Desert, and through examination of the 16S rRNA gene and associated 16S-23S internal transcribed spacer (ITS) region, discovered that there were at least two different species of *Microcoleus* in desert soils. Variation in the sequences of *Microcoleus* indicated the presence of several cryptic taxa beyond the two taxa recognized at that time. *Nostoc* spp. strains were subsequently collected from arid regions in California and Utah and were also sequenced for the 16S gene and associated 16S-23S ITS region. *Nostoc* is a polyphyletic genus, meaning that the name has been applied to a number of clades that do not have a common ancestor for all members assigned to the genus (the definition of a monophyletic taxon). This polyphyly is likely due to the fact that *Nostoc* is poor in morphological characters. However, phylogenetic analysis resulted in three new species of BSC *Nostoc* and the new genus *Mojavia* (Řeháková et al. 2007). This work defined the clade containing *Nostoc sensu stricto*, thus allowing workers to begin to revise the genus by naming *Nostoc*-like strains outside of the genus. The first cryptic species of cyanobacteria were subsequently described in Osorio-Santos et al. (2014). In their study, seven new species of *Oculatella* were described from diverse soil and

subaerial habitats, and most of the species were morphologically cryptic. Osorio-Santos et al. (2014) was the first study in which percent dissimilarity of aligned ITS regions was used to justify recognition of cryptic lineages, following the lead of the pioneering work of Erwin and Thacker (2008), who first suggested ITS dissimilarity thresholds as a species-recognition tool in cyanobacteria.

Polyphyly is not just a problem in *Nostoc* and *Microcoleus*, but has been repeatedly observed in many morphologically defined desert soil genera once molecular approaches have been used (Pietrasiak et al. 2019). Many species and genera have been assumed to be cosmopolitan based on morphological studies, but convergent evolution in desert species has led to similar morphology between distantly related taxa (Belnap 2003, Mai et al. 2018).

Trichocoleus, the focal genus in this study, and *Microcoleus*, one of the most common genera in BSC, are examples of morphologically similar desert genera that are not phylogenetically related (Komárek & Anagnostidis 2005, Mai et al. 2018). In fact, these taxa are so morphologically similar that all current *Trichocoleus* species used to belong in *Microcoleus* (Komárek and Anagnostidis 2005). Although morphology can be helpful, and it is interesting to analyze, molecular methods have been shown to more accurately reflect evolutionary relationships (Belnap 2003, Mai et al. 2018), and thus allow detection of polyphyly requiring taxonomic revision.

Another recurring problem in cyanobacterial taxonomy is the undersampling of isolated strains. In previous cyanobacterial taxonomic studies, it was rare to have multiple strains of the same species sequenced. Consequently, intraspecific variation in desert cyanobacteria is far from being understood (Curtis et al. 2006, Pietrasiak et al. 2019). Molecular data is needed to accurately assess cyanobacterial diversity in all habitats, especially in desert soils, and multiple

strain sequences are important to evaluate intraspecific variation within those species. Therefore, we are using multiple sequences of *Trichocoleus* both new, and previously reported, to better understand intrageneric and intraspecific variation.

Gypsiferous Soils

The Chihuahuan Desert, U.S.A., contains large gypsum dune fields in White Sands National Park and surrounding areas. Gypsum ($\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$) covers more than 200 million hectares of Earth's surface and is found in many semi-arid to arid biomes (Buck & Hoesen 2002). Gypsum soils are restricted to these areas for two main reasons. The first is that gypsum is water soluble and therefore, would not be present in its crystalized structure in mesic soils. The second reason is that dry air creates a high evaporitic demand that in turn causes capillary uplift of the gypsum towards the soil surface (Verheye & Boyadgiev 1997, Herrero et al. 2009, Muller et al. 2017). These soil properties have previously been found to be responsible for assembly of unique plant communities (Escudero et al. 2015). This finding makes it plausible that this unusual soil will also support distinct microbial communities.

Shields et al. (1957) studied microbial communities in the gypsiferous soils of White Sands prior to the development of molecular taxonomic methods. They identified common soil species in the cyanobacterial genera *Microcoleus*, *Nostoc*, *Schizothrix*, and *Scytonema* from sand dunes at White Sands. However, they did not document the morphology of the species they reported. These morphological identifications were made before *Trichocoleus* was split out from *Microcoleus* and established as a genus. The work of Shields et al. (1957) is dated, and further study of this interesting site is warranted.

While more recent studies of cyanobacteria in gypsum soils have been conducted, it is typical that morphology, metagenomics, or only the 16S rRNA gene sequence is used, leading to

misidentification or identification only to the family or genus level (Belnap & Gardner 1993, Dong et al. 2007, Menéndez-Serra et al. 2019). Although Menéndez-Serra et al. (2019) only identified bacterial OTUs to high level taxonomic groups (phylum, class, and/or order), they found that about 10% of sequences had high genetic novelty for the 16S rRNA gene, indicating that further molecular studies should be performed and will likely reveal new genera, as the 16S rRNA gene is the primary molecular indicator of new genera (Boyer et al. 2002). The lack of 16S rRNA data for comparison highlights the need to continue systematic studies, especially in saline and gypsiferous desert soils.

Trichocoleus

In this study, we are specifically researching the genus *Trichocoleus* in the family Trichocoleusaceae (Cyanobacteria). Komárek & Anagnostidis 2005 described this genus as a filamentous cyanobacteria that can be separated from related taxa by the presence of multiple trichomes in a common sheath. *Trichocoleus* was essentially defined as “not-*Microcoleus*,” which is another genus with multiple trichomes in a common sheath. *Trichocoleus*, however, has thinner trichomes and parietal thylakoids (Komárek & Anagnostidis 2005), compared to *Microcoleus* which has radial or fasciculated thylakoid arrangements (Siegesmund et al. 2008).

The type species of *Trichocoleus* is *Trichocoleus delicatulus* and, unfortunately, this species has an insufficiently written description; there are no photos, drawings, or type materials (e.g., an herbarium packet). No sequence attributable to this species exists at present. The absence of data on this species creates a great deal of uncertainty in the morphology and taxonomy of the genus. Since the original description, there have been reports of *Trichocoleus* that differ in cell size and number of trichomes in a common sheath, but all the strains identified

morphologically from soil as *Trichocoleus* fall into a single clade when characterized molecularly (Mühlsteinová et al. 2014, Machado de Lima et al. 2020, Mehda et al. 2022).

Trichocoleus has been molecularly identified in the Sahara Desert in Africa (Mehda et al. 2022), the Atacama Desert in Chile, the Mojave Desert and Colorado Desert in the U.S. (Mühlsteinová et al. 2014), Great Smoky Mountains National Park in the U.S. (Johansen et al. 2008), and semi-arid regions of Brazil (Machado de Lima et al. 2020). With *Trichocoleus* being described all over the world, and all throughout the United States, we believed that White Sands National Park would also be home to this genus.

The aim of this study is to further develop the *Trichocoleus* genus both morphologically and molecularly using newly sequenced strains from White Sands National Park. In this paper, we recognize 17 potential species-level *Trichocoleus*. Unfortunately, some of these strains were sequenced in other labs and are not available to name. Furthermore, some strains were lost from culture before completion of this study. We provide descriptions for 11 *Trichocoleus* species that we anticipate will be published in a subsequent manuscript. With the previously named species *T. desertorum*, *T. caatingensis*, and *T. badius*, this will bring the number of polyphasically described species to 14 species, making *Trichocoleus* one of the most species-rich cyanobacterial genera in desert soils.

MATERIALS AND METHODS

Field Collection

Samples were collected at White Sands National Park, New Mexico, during June and July of 2019. The study area (denoted by the black circle in Fig. 1) was split into six landforms based on an elevation gradient. Each landform was further split into three plots, except for Landform 3, which had four plots. *Trichocoleus* strains were isolated from four of the landforms

(Table 1). From each GPS plot point (Table 1), a 50 m 61° West to East transect was sampled every 5 m by obtaining 1.0-3.0 cm of topsoil using a 30 mm diameter centrifuge tube as a coring and sampling device. Each collection along the transect was combined into a composite soil sample whirl Pak or Ziplock bag to minimize the effects of microheterogeneity in microbial communities (Grondin & Johansen 1993, Flechtner et al. 2008).

Laboratory Analysis

From each sample bag, 1.0 g of soil was added to 100 mL of liquid Z8 media (Carmichael 1986) and then shaken at room temperature for 30 minutes on a rotary shaker. Dilution plate methods were used to create a total of six Z8 agar plates per sample. Three plates at 10^{-3} sample dilution were made by taking 0.1 mL directly out of the mixed sample and spreading the sample on the plate. Three plates at 10^{-4} sample dilution were made by taking 1.0 mL out of the mixed sample and adding that to 9.0 mL of Alga Grow liquid media (Plagron, Ospel, The Netherlands), mixing those together and spreading 0.1 mL onto the plate. Plates were left to sit overnight and then sealed with parafilm, placed under a 12:12 light:dark cycle at 25°C, and allowed to grow for six weeks. Cyanobacteria were isolated from the plates and placed into liquid Z8 media to grow for another six weeks before being transferred onto Z8 agar slants and deposited into the JCU Culture Collection. Strain designations were made for each isolate by using the site code (WS) followed by the landform number, plot number, initials of researcher (AS or JRJ) and isolate number (e.g., WS6-4-AS1). During isolation, photographs were taken with an Olympus BH2 photomicroscope with Nomarski DIC optics and an Olympus D-25 digital camera with Cell-Sense software. Before entering the JCU culture collection, all isolates were photographed. Sequenced strains were characterized by taking at least ten images each and immobilized on herbarium packets subsequently deposited in the Herbarium of the University of

South Bohemia in České Budějovice, Czech Republic. In addition to the strains from White Sands National Park, strains previously isolated and sequenced by other researchers in the Johansen lab were also morphologically characterized and their sequences included in my analyses.

Molecular Analysis

DNA was extracted with the DNeasy PowerSoil Kit (Qiagen, Carlsbad, CA, U.S.A) according to the manufacturers protocol. To amplify the 16S rRNA gene, primers VRF1R: 5'–CTC TGT GTG CCT AGG TAT CC–3' and VRF2F: 5'–GGG GAA TTT TCC GCA ATG GG–3' (Boyer et al. 2001,2002) were used in a polymerase chain reaction (PCR). The PCR mixture contained 1.0 µL of each primer (10 µM concentration), 12.5 µL LongAmp™ Taq 2x Master Mix (NEB, Ipswich MA), 1.0 µL template DNA (50 ng/µL), and nuclease free water to a final volume of 25 µL. Using a Bio-Rad S1000 or C1000 Thermal Cycler (Hercules, CA), the PCR parameters were as follows: 35 cycles of denaturing at 94°C for 45 s, annealing at 57°C for 45 s, and extension at 72°C for 2:15 m, followed by a final extension at 72°C for 5:00 m, and then an indefinite dwell at 4°C. The PCR products were visualized with gel electrophoresis on an ethidium bromide stained 1.5% agarose gel. If there was a presence of the 16S rRNA band (~1500 base pairs), the PCR product was cloned using the StrataClone PCR cloning kit (Agilent, Santa Clara, CA) according to manufacturer protocols. Bacteria containing putative clones were plated on LB plus ampicillin agar plates with 40 µL of 2% X-gal for 12-14 hours at 37°C. Two white colonies per strain were transferred to LB plus ampicillin broth and incubated for 12 hours at 37°C. The QIAprep® Spin Miniprep kit (Qiagen, Carlsbad, CA) was used for plasmid purification before EcoRI (NEB, Ipswich, MA) digestion to select successful clones. Cloned products were sent to Functional Biosciences, Inc. (Madison, WI) for Sanger sequencing with the

plasmid primers M13 forward and M13 reverse and internal primers 5: 5'–TGT ACA CAC CGG CCC GTC–3' (Willmotte et al. 1993), 7: 5'–AAT GGG ATT AGA TAC CCC AGT AGT C–3' (Nübel et al. 1997) and 8: 5'–AAG GAG GTG ATC CAG CCA CA–3' (Willmotte et al. 1993).

Phylogenetic Analysis

Phylogenetic analysis was done using the protocol from Pietrasiak et al. (2019). Raw sequences were arranged, proofread, and assembled into contigs in Sequencher (Gene Codes Corp, Ann Arbor, MI, U.S.A). All sequences were submitted to NCBI GenBank. Sequences from this study as well as other previously submitted *Trichocoleus* sequences in GenBank (Mühlsteinová et al. 2014, Machado de Lima and Branco 2020) were aligned with ClustalW. First, the 16S sequences were separated from the ITS and aligned to assess 16S similarity and phylogenetic relationships. Bayesian inference (BI) and maximum likelihood (ML) analyses were performed on the CIPRES Science Gateway supercomputer cluster using MrBayes on XSEDE 3.2.6 and RAxML-HPC2 on XSEDE 8.2.10, respectively. For both analyses the GTR+G+I evolutionary model was used.

Bayesian analysis was run for 95 million generations discarding the first 25% of samples as burn-in, with parameters set for the GTR+G+I model. The final average standard deviation of split frequencies was 0.0244, the average potential scale reduction factor (PSRF) for parameter values was = 1.000, and Estimated Sample Size >4000 for all parameters. These convergence factors indicated that convergence between the two chains was achieved (Gelman and Rubin 1992). For ML, we ran rapid bootstrapping by choosing a fixed random number seed (12,345) and requesting 1,000 bootstrap iterations. All other options in CIPRES (Miller et al. 2015) were left as default. Bootstrap values from the ML analysis were mapped on to nodes in the Bayesian Inference tree, which was then post-edited in Adobe Illustrator CS5.1.

Sequences of the 16S-23S ITS region of the operon containing both tRNA genes were aligned using secondary structure linear sequences, with indels coded (0 for gap, 1 for nucleotide), and an unrooted phylogenetic tree was obtained with a BI analysis with 20 million generations, discarding the first 25% of samples as burn-in, choosing NST = MIXED, and applying the GTR+I+G evolutionary model. Average standard deviation of split frequencies was = 0.0106 and the average PSRF for this analysis was = 1.000. Estimated Sample Size was > 19,000 for all parameters, which indicates that the chains converged and each parameter was well sampled. A maximum parsimony analysis was also run on the ITS alignment using default parameters with the exception that indels were coded as a fifth base. A heuristic search with 1000 bootstrap replicates was performed. This analysis was conducted using PAUP on XCEDE 2.0 in the CIPRES Science Gateway. Posterior probabilities from the Bayesian Analysis were mapped on to nodes of the maximum parsimony analysis and the phylogeny was post-edited in Adobe Illustrator CS5.1.

P-distance of 16S rRNA sequences was determined in PAUP to calculate percent similarity ($100*[1-p\text{-distance}]$) among our strains of interest. Percent dissimilarity of aligned ITS regions was also calculated using the SHOWDIST command in PAUP, and multiplying p by 100 to give results in percent. The hypothetical ITS secondary structures of helices D1-D1', Box-B and V3 were derived using M-fold (Zuker 2003) and re-drawn in Adobe Illustrator CS5.1.

RESULTS

Phylogenetic Analysis

All strains in the genus *Trichocoleus* fell into a monophyletic clade when using 16S rRNA data (Fig. 2). The node including *Trichocoleus desertorum sensu stricto*, *Trichocoleus* sp. 2, sp. 3, sp. 4, sp. 6, sp. 14, and sp. 18 had a bootstrap value of 85, indicating that every strain in that clade is different than anything outside of it. *T. badius* contained three strains each with

different operons. Two *T. badius* strains were grouped together within many of the other *Trichocoleus* sequences and the third strain was separated from all other *Trichocoleus* strains (posterior probability 0.76, bootstrap 55). This third sequence may have represented horizontal gene transfer, or an ancient duplication event in the genome with subsequent divergence. The 16S rRNA phylogeny did not resolve species, likely due to the fact that the 16S rRNA sequences were highly similar among species and could not resolve the nodes sufficiently. However, with the addition of the strains in this study, relationships between species within the tree were better resolved than past trees (Mühlsteinová et al. 2014).

The 16S-23S ITS phylogeny resolved 21 species-level groups (Fig. 3). The four strains of previously described *Trichocoleus desertorum sensu stricto* fell into a well-supported clade (bootstrap 86, posterior probability 0.92). *Trichocoleus caatingensis* CAT-CD2 had a long branch length, separating it clearly from its sister taxon *Trichocoleus* sp. 16 WS7-2-JRJ1. *Trichocoleus badius* CRS-1 represented by two sequences, each from a different operon, was sister to *Trichocoleus* sp. 1. None of the new sequences added as part of this study could be placed into any of the previously described species.

The 16S-23S ITS phylogeny did not clearly establish where to draw the distinction between species when there was only a single representative. However, it did resolve some of the species we recognize, as evidenced by the fact that when there were multiple strains within a single species, they formed a monophyletic species. *Trichocoleus* sp. 6 (from Mojave Desert soils) and *Trichocoleus* sp. 9 (from Sahara Desert soils) both had seven strains. The Mojave Desert also had two other species (sp. 13 and sp. 18) both of which were from different landforms than each other and from the other Clark Mountain strain (CMT-1BRIN-NPC4B). The Sahara Desert also had two additional species well separated from each other and sp. 9, sp.

7, and sp. 10. A very interesting finding was that the two strains of *Trichocoleus* sp. 15, were biogeographically separated, one from San Nicolas Island in California (SNI-TA9-AZ2) and the other from desert soil in China (CXA025). The strains from the Atacama Desert (all ATA strains) fell into five distinct species, including the previously described *T. desertorum*. These all belonged to *T. desertorum* in the work where that taxon was described (Mühlsteinová et al. 2014). However, the additional taxon sampling achieved in this study showed that more species were represented in this group. All Atacama species were isolated from sites at least 100 km from each other, thus geographical separation exists even though the desert forms a continuum for its entire 1000+ km length.

Six strains of *Trichocoleus* were isolated from the restricted access part of White Sands National Park. These fell into five different species (spp. 1, 11, 14, 16, 17). Unfortunately, one strain was lost before full characterization and voucher preparation was completed (sp. 11), so we will only be able to describe four species from White Sands. None of the strains were isolated from the same sample, although two of the strains were isolated from the same landforms (WS4 and WS7). The White Sands species are all well-separated phylogenetically.

All species we recognize in our phylogeny had 16S-23S ITS dissimilarity among species differing by more than 5%, while members of the same species were mostly <3.0% dissimilar. Some strains fell in the ambiguous area of 3-5%, and these were usually placed in the same species, a conservative approach. Some of our strains had more than one operon, and these operons generally differed by 3-5%. These included strains in sp. 4 (ATA-14-RM36), sp. 7 (SIK24), sp. 9 (SIK23, SIK114), and *T. badius* (CRS-1).

16S-23S ITS structures

The D1-D1' helix in the 16S-23S ITS was sequenced and assembled for all strains (Fig. 4). The sequences were almost all the same length (62 nt), with only *T. caatingensis* differing (61 nt). They also all had the same basal clamp, 5'-GACCU—AGGUC-3'. The middle helix had a very similar structure and sequence, only differing in the mismatched bases in the upper third of the helix. The terminal loop region shows possible base pairings, represented by the lines connecting them in the figure (Fig. 4). Although these pairings did not form in M-fold, the presence of them in nearly every species group indicates this may be a pattern for the genus, and these may form in interaction with proteins in the cytoplasm. Alternative pairings for *T. desertorum* and *T. badius* are shown as examples of alternative structures (Fig 4, A,Q and B,R, respectively). The D1-D1' helix in *Trichocoleus* is conserved in sequence length and structure but does vary in sequence between species.

The BoxB structures all show the same 4 bp basal clamp, 5'-AGCA—UGCU-3' and are identical above that clamp to the U—A pairing two nucleotides above the basal bilateral bulge (Fig. 5). The BoxB helices differed in length by up to 4 nt. The primary differences in structure were in the terminal loop, which differed in number of nucleotides due to the occasional formation of an extra pairing in the helix (Fig. 5). *Trichocoleus* sp. 9 (Fig. 5, J) has high intraspecific variability on the 3' side of the terminal loop. One strain of sp. 9 had an identical structure to sp. 10 (Fig. 5, K). SIK24 is very different from all other SIK strains in the terminal loop; it is shorter and has a different sequence (Fig. 5, H). *Trichocoleus* sp. 6 and sp. 9 have identical BoxB helices, but differ notably in other parts of the 16S-23S ITS.

V3 region structures are represented in Figure 6. There is high variability among all species; however, there is consistency in the basal clamp through the U—A pairing one base pair

above the second bulge, indicating that *Trichocoleus* has a highly conserved V3 region throughout the base of the helix. Above that portion, there is significant variation. Many species have a consistent “UCAAA” loop, but many did not. *Trichocoleus* sp. 9 and sp. 17 both have a small bulge where there is an unpaired C across from an AU, which is not seen in any other sequence (Fig. 6M, R). Interestingly, these two species are not near each other in the 16S-23S ITS phylogeny. *Trichocoleus* sp. 7, sp. 11, and sp. 15 are identical other than sp. 15 having one base pair difference (Fig. 6J). Sp. 7 and sp. 11 are sister taxa, but sp. 15 is separated in the 16S-23S ITS phylogeny. *Trichocoleus* sp. 5 and sp. 6 are identical other than one nucleotide change (Fig. 6H and I), and these strains are very separated on the 16S-23S ITS tree. *Trichocoleus* sp. 13 and sp. 18 show this same pattern of identical V3 regions, but separate species-level placement on the 16S-23S ITS phylogeny (Fig. 9P). Variation of *Trichocoleus* sp. 9 is shown in Fig. 9L and M. Analysis of the hypothetical secondary ITS structures provided qualitative evidence of lineage separation at the species level with relatively high conservation at the genus level; however, the phylogenetic placement of the taxa did not necessarily reflect similarity in the V3 region, suggesting the importance of other data (such as overall ITS dissimilarity) in separating these taxa.

Morphological analysis

All *Trichocoleus* strains had several common morphological characteristics. All had parietal thylakoids and trichomes in either diffluent mucilage or firm sheaths (Figs. 7, 8, 9, 10). One of the defining features of *Trichocoleus* is the presence of multiple trichomes in a common sheath; however, we had some samples without any sheath material at all (Figs. 7A-G, N-Q; 9 A-E). There was variation in apical cell length, width, and shape. *Trichocoleus* sp. 3 ATA3-4Q-KO9 has long bent apical cells (Fig. 7N, Q); *Trichocoleus* sp. 8 ATA12-5-KO5 has long, thick, and

pointed apical cells (Fig. 9B-C), whereas most species possess bluntly conical apical cells (Figs. 7-10). Some strains differed in size ranges, but most size ranges had overlap, so dimensions are less useful as a defining characteristic of this set of species.

Taxonomic proposals

***Trichocoleus* sp. 1 WS7-5-JRJ8 (Fig. 7, A-G)**

Colony a flat blue-green mat. Filaments with one to several trichomes in diffluent mucilage. Sheath colorless, without firm layer, often absent. Trichomes nonmotile, lacking false branching, constricted at the crosswalls, mostly 3.0-3.3 μm wide, but ranging 2.8-3.6 μm wide. Cells shorter than wide, with parietal thylakoids along all cell walls encompassing a lighter-colored centroplasm, nongranular, mostly 1.9-2.2 μm long, but ranging 1.4-2.5 μm long. End cells bluntly rounded to bluntly conical, with conical cells up to 3.1 μm wide. Hormogonia present, few-celled to many-celled.

***Trichocoleus* sp. 2 ATA2-1-CV2 (Fig. 7 H-M)**

Filaments with one to several trichomes in a visible sheath. Sheath colorless, firm, and some exhibit structure. Filaments 5.8-6.0 μm wide but ranging from 2.3-8.7 μm wide. Trichomes motile, lacking false branching, constricted at crosswalls, mostly 3.3-3.6 μm wide, but ranging 3.0-6.3 μm wide. Cells vary in length greatly, but mostly almost isodiametric with parietal thylakoids around all cell walls, granular, 1.8-3.6 μm long. End cells long, tapered, and bluntly to acutely conical, 5.0-9.5 μm long. Hormogonia present, few-celled to many-celled. Necridia present.

***Trichocoleus* sp. 3 ATA3-4Q-KO9 (Fig. 7 N-Q)**

Filaments with one trichome, 7.5 μm wide. Sheath colorless, firm, and often absent. Trichomes nonmotile, lacking false branching, slightly constricted at crosswalls, and 2.9-3.2 μm wide. Cells

mostly almost isodiametric with parietal thylakoids along all cell walls, and granules, mostly 2.5-3.0 μm long, but ranging 2.2-5.0 μm long. End cells long, tapered, acutely conical, and bent, 6.1-7.5 μm long. Hormogonia present, few-celled to many-celled.

***Trichocoleus* sp. 4 ATA14-3-RM35 and ATA14-3-RM36 (Fig. 8, A-F)**

Filaments contain multiple trichomes wrapped around each other in a common sheath. Sheath colorless, firm, and 5.3-15.6 μm wide. Trichomes nonmotile, lacking false branching, constricted at crosswalls, 2.8-4.0 μm wide. Cells typically longer than wide with parietal thylakoids along all cell walls, granular, mostly 2.0-3.5 μm long, but ranging 1.5-7.5 μm long. End cells long, conical, slightly tapered, 1.8-3.5 μm wide, 2.0-7.0 μm long. Hormogonia present, few-celled to many-celled.

***Trichocoleus* sp. 5 ATA1-4-CV12 and ATA1-4-KO5 (Fig. 8, G-M)**

Filaments with one to multiple trichomes in a sheath. Sheath colorless, firm, often absent, 4.0-11.9 μm with sheath present. Trichomes nonmotile, lacking false branching, constricted at crosswalls, 2.6-3.6 μm wide. Cells mostly shorter than wide with parietal thylakoids along all cell walls, granular, 2.1-4.2 μm long. End cells tapered, bluntly to acutely conical, 4.2-16.6 μm long. Hormogonia present, few-celled to many-celled.

***Trichocoleus* sp. 6 WJTs and CMT-1BRIN-NPC4B (Fig. 8, N-S)**

Filaments with one to several trichomes in diffluent mucilage. Sheath colorless, without firm layer, often absent, and some exhibit structure. Trichomes motile, lacking false branching, constricted at crosswalls, 2.0-4.0 μm wide. Cell lengths vary, 2.2-15.8 μm long, all with parietal thylakoids along all cell walls, granular. End cells bent and bluntly conical, tapered, 3.0-5.0 μm long. Necridia present.

***Trichocoleus* sp. 8 ATA12-5-KO5 (Fig. 9, A-E)**

Filaments with one to several trichomes in a diffluent mucilage. Sheath colorless, often absent. Trichomes nonmotile, lacking false branching, slightly constricted at crosswalls, 2.9-3.6 μm wide. Cells length varies, 2.3-6.3 μm long, parietal thylakoids along all cell walls, granular. End cells taper to a point, vary in length, 6.9-21.9 μm long. Hormogonia present, few-celled to many-celled. Necridia and involution cells present.

***Trichocoleus* sp. 14 WS6-3-AS2 and WS4-1-AS1 (Fig. 9, F-J)**

Filaments with one to several trichomes in diffluent mucilage, 3.9-7.4 μm wide. Sheath colorless, firm, often absent. Trichomes nonmotile, without false branching, constricted to slightly constricted at crosswalls, 1.8-3.6 μm wide. Cells mostly isodiametric to shorter than wide, parietal thylakoids along all cell walls, granular, 1.8-5.5 μm long. End cells bluntly conical to rounded, without tapering, 3.4-5.2 μm long. Hormogonia present, few-celled to many-celled.

***Trichocoleus* sp. 15 CXA25 and SNI-TA9-AZ2 (Fig. 9, K-Q)**

Filaments with one to two trichomes in a firm, persistent sheath, 4.9-12.3 μm wide. Sheath colorless, firm. Trichomes nonmotile, without false branching, slightly constricted at crosswalls, 2.2-3.5 μm wide. Cells mostly almost isodiametric, parietal thylakoids along all cell walls, nongranular, 1.6-4.9 μm long. End cells are long, tapered conical cells, 5.8-11.9 μm . Hormogonia present, few-celled to many-celled.

***Trichocoleus* sp. 16 WS7-2-JRJ1 (Fig. 10, A-F)**

Filaments with one to several trichomes in diffluent mucilage or firm, colorless sheath, 3.9-13.2 μm wide. Trichomes nonmotile, exhibit false branching, constricted at cell walls, 1.9-4.5 μm wide. Cell length varies widely, from shorter than wide to longer to wide or with parietal thylakoids along all cell walls, granular, 1.0-2.6 μm long. End cells bluntly conical, 3.0-5.1 μm long. Hormogonia present, few-celled to many-celled.

***Trichocoleus* sp. 17 WS4-3-AS3 (Fig. 10, G-K)**

Filaments with one to several trichomes tangled in a diffluent mucilage or in a firm, blue green sheath material, 3.9-13.2 μm wide. Trichomes nonmotile and highly constricted at cell walls, 1.9-3.8 μm wide. Cells shorter than wide with parietal thylakoids along all cell walls, granular, 1.0-2.3 μm long. End cells bluntly conical, 3.0-3.9 μm long and often wider than other cells in the trichome 2.5-4.3 μm wide. Hormogonia present, few-celled to many-celled.

DISCUSSION

Trichocoleus was only molecularly characterized as a genus in 2014 (Mühlsteinová et al. 2014). Previously, all accounts were morphological descriptions (Komárek & Anagnostidis 2005). The “ATA”, “CMT”, and “WJT” strains from Mühlsteinová et al. (2014) were used for our molecular analysis. The reference strain designated from that study was *Trichocoleus desertorum* ATA4-8CV12, and it is the strain upon which the holotype is based. Our findings supported the initial placement of *Trichocoleus* in the 16S rRNA tree (Mühlsteinová et al. 2014); however, we were able to further differentiate potential species groups within the 16S rRNA tree as well (Fig. 2). Our phylogenetic tree supports the placement of *Trichocoleus* in the family Trichocoleusaceae, together with *Pinocchia* and the undescribed species “*Leptolyngbya hensonii*” (which will need to be placed in a new genus) (Fig. 2). Strunecký et al. (2023) additionally placed the unsequenced genera *Geitleribactron*, *Schizothrix*, and *Sirocoleum* into the Trichocoleusaceae.

Although 16S rRNA data cannot be used to characterize species alone, it is helpful to analyze the similarities of the 16S rRNA gene when trying to separate sequence differences due to species versus sequence differences due to operons. For example, the 16S-23S ITS sequence in different operons of *Trichocoleus* sp. 4 ATA-14-RM36 were distinct and could be confused

with species-level differences if we did not know they were from the same strain. Furthermore, their 16S rRNA sequences are identical, indicating that the sequenced strains belong to the same species. In the instance where only a single operon is obtained for a strain, and multiple strains in the species are recovered but the 16S-23S ITS operons indicate species-level separation, this can only be resolved by looking at the 16S rRNA sequence data as a cross-check.

Mühlsteinová et al. (2014) had very little grouping in their 16S rRNA tree, likely due to insufficient sequence data, which is why almost all sequences in their study, except for *T. badius*, were named *T. desertorum*. We were able to add 6 new sequences from our study and others to help parse out the 16S rRNA data relationships. With the D1-D1' data added from our study, we were able to recognize the pattern of the C—G and A—U potential pairings in the upper loop. This data was present in the 2014 study and at that time there was not enough data to clearly see the pattern. D1-D1' structures were very consistent throughout all species, even those with multiple operons. The 16S rRNA data grouped species together as well, although not as unequivocally.

Johansen et al. (2008) and Mühlsteinová et al. (2014) previously recognized the presence of multiple operons in *T. badius*. Operons 1 and 2 of *T. badius* are in the same clade on the 16S rRNA tree, and only differ in the ITS. We found those same two operons grouped together in our 16S rRNA tree as well, indicating that they are different operons and not due to the presence of a second contaminant. This same pattern is also seen with strains ATA13-3-RM36, SIK24, SIK23 and SIK114, which were included in our analyses.

Machado de Lima and Branco (2020) were the next researchers to analyze *Trichocoleus* molecularly and named a new species *Trichocoleus caatingensis*. The “SIK” strains used in our analysis from Mehda et al. 2022 were previously named *T. desertorum* because they were not

separated into their own clade in the phylogeny. Based upon our new 16S-23S ITS analyses, we consider that these strains represent their own species, and therefore, have designated most of them as *Trichocoleus* sp. 9, whereas two strains are grouped together into sp. 7, and a third strain is grouped with SBC54 as sp. 10. These will remain as *Trichocoleus* sp. 7, sp. 9, and sp. 10 in the eventual publication because we do not have the authority to name species from sequences outside of our lab. They do show clear separation from *Trichocoleus desertorum sensu stricto* in both the 16S rRNA gene and 16S-23S ITS phylogenies, and the curators of those strains could describe them as new species.

The 16S-23S ITS phylogeny provides evidence for the species groups characterized in this study (Fig. 3). Although there are many highly supported clades within the tree, there is low backbone support throughout. The highly supported clades are those of the species at the tips of the tree, while the backbone and some basal lineages remain unsupported. Having low backbone support indicates that we cannot fully understand the relationship between the species to each other, but we are able to conclude that certain groups are their own species. *Trichotorquatus* also has low backbone support in the 16S-23S ITS phylogeny (Pietrasiak et al. 2021), and it is a genus with multiple operons, possibly causing the uncertainty. *Oculatella*, another cyanobacterial genus with multiple operons, had low backbone support in the 16S-23S ITS phylogeny (Osorio-Santos et al. 2014); however, with more sequences added, the 16S-23S ITS was able to resolve the species and produce a supported backbone in the 16S-23S ITS phylogeny (Becerra-Absalón et al. 2020). As more strains of *Trichocoleus* are sequenced, we would hope to see a clearer picture of species groups in relation to each other.

Exploring the details of taxonomic relationships of an entire genus can impact many types of future research. Since the *Trichocoleus* clade is supported and well defined, new

sequences can be added and identified easily. The work we did in this study will prevent taxonomic issues in the future as well as prevent this genus from becoming a paraphyletic or polyphyletic assemblage, a common occurrence in other cyanobacterial genera (Johansen et al. 2011, Bauersachs et al. 2019). It is also important for metabarcoding studies which use available sequences to identify large amounts of DNA based on a small portion of a gene. If more sequences become available for comparison, it can strengthen the identification power of metabarcoding (Li et al. 2019, MacKeigan et al. 2022).

Trichocoleus seems to be another problematic cryptic genus due to the insufficient characterization of the generitype, *T. delicatulus*. This is a bit of an unusual problem because *Trichocoleus* is a relatively new genus, but an older species belonging to another genus (*Microcoleus*) was chosen as type due to that species having taxonomic priority. Incomplete descriptions can lead to difficulties when moving forward with sequencing classical taxa. *Schizothrix* is another example of a genus without genetic data associated with the type species (Komárek et al. 2006) and given its morphological similarity to *Trichocoleus*, it is difficult to determine whether they are truly different genera. Epitypification utilizing molecular data for recently obtained isolates of classical taxa can solve this problem and assist in the harmonization of classical taxonomy and modern taxonomy (Shalygin et al. 2019). Many researchers today are using molecular data to describe new genera and species without consideration of the classical taxa (Moreira C. Fernandes et al. 2021), and we hope this practice will end in the near future.

Not only is *Trichocoleus* a cryptic genus, the species within the genus are also cryptic. Morphological variation was present within *Trichocoleus*, but most species show a similar set of characteristics: multiple trichomes in a common sheath, presence of hormogonia, and cells constricted at the crosswalls with parietal thylakoids. These traits should be considered essential

genus traits for *Trichocoleus*, as they were present in almost every strain. There were three species groups without sheath material, but this could be due to phenotypic plasticity and not because they lack a common sheath. These traits agree with the original description of *T. delicatulus*; however, there are also many differences observed. Since we have expanded the number of species groups identified in *Trichocoleus*, we have found that there are some species that exhibit morphological characteristics that vary from the protologue of *T. desertorum*. Specialized cells, like necridia, have been observed (Fig. 7K), as well as false branching and much wider trichomes than originally observed. Some species can be differentiated based on morphology, and from our samples, apical cell shape is the most distinctive and taxonomically informative trait. *Trichocoleus* sp. 3 ATA3-4Q-KO9 possess long apical cells with a bent notch, and no sheath material present (Fig. 7N,Q). *Trichocoleus* sp. 8 ATA12-5-KO5 possess long, thick, and pointed apical cells (Fig. 9B-C). However, these end cells are also very similar to *Trichocoleus* sp. 15 (Fig. 9K-Q). The distinction between these two is that sp. 8 has no visible sheath material, whereas sp. 15 possesses very firm and thick sheaths as well as necridia.

Often times, cryptic species can be separated based upon ecological differences, particularly in habitat. Geographical proximity is not an important factor for lineage separation in this genus. The White Sands strains sequenced herein are scattered throughout the phylogenetic trees and are not explicitly sister to strains occurring in nearby desert sites. If habitat proximity were a driving factor, we would expect White Sands strains to not only group together, but to also group with strains isolated from the nearby Clark Mountains. However, only WS4-1-AS1 and WS6-3-AS2 place close to each other, and the rest are interspersed throughout the tree with a variety of sister taxa. This same pattern is observed with the samples from the Sahara Desert, Atacama Desert, and Clark Mountains.

Identifying *Trichocoleus* populations in nature will prove difficult, as the full life cycle and range of morphological variability often cannot be observed in field collections. Furthermore, soil samples contain many morphologically similar Synechococcalean taxa which, when present in the same soil community, make confident identification of genera almost impossible. Samples from lakes and hot springs also have complex communities causing populations of simple filamentous taxa to be difficult to identify based on morphology alone in these habitats as well. Consequently, identifying thin filamentous cyanobacteria in the Synechococcales from isolated strains in a culture collection is the best practice. Even then, many strains look too alike to tell apart and most have overlapping size ranges which highlights the critical importance of molecular data, especially for recognition of cryptic species. Without molecular data, there would be many species lacking recognition. The use of DNA sequencing in the present study has allowed for identification of numerous species in a previously limited genus, and is likely only the beginning in the discovery of diversity in *Trichocoleus* and other Synechococcalean genera.

ACKNOWLEDGMENTS

I am very grateful to Abigail Perrino, who collected soil samples for me during her field work in White Sands National Park and to David Bustos for providing logistical support for her during sampling. I'd also like to thank Nicole Pietrasiak for her help in setting up the study design as well as providing guidance on my thesis work. Field collections at White Sands National Park were made possible through Grant no. P21AC11241-01 and Permit no. WHSA-2018-SCI-0012. Some of the strains appearing in my phylogenies were isolated and sequenced by past and present lab members Karina Osorio-Santos, Nicole Pietrasiak, Radka Mühlsteinová, and Brian Jusko. Thank you, Dr. Ashley Wain, for coordinating the Biology 1 laboratory sequencing of

several strains from White Sands (*Trichocoleus* and others). I would also like to thank my committee members, Drs. Rebecca Drenovsky and Michael Martin, for providing helpful recommendations for improving the thesis as well as my fellow lab members for their support: Joy Jackson, Brian Jusko, and Chelsea Villanueva. Lastly, thank you to Dr. Jeff Johansen for providing unending assistance and mentorship for this entire work.

LITERATURE CITED

- Anagnostidis, K. 2001. Nomenclatural changes in cyanoprokaryotic order Oscillatoriales. *Preslia* 73: 359–75.
- Ashley, J., Rushforth, S. R., & Johansen, J. R. (1985). Soil algae of cryptogamic crusts from the Uintah Basin, Utah, U.S.A. *Great Basin Naturalist*, 45(3), 432–442.
- Bauersachs, T., Miller, S. R., Gugger, M., Mudimu, O., Friedl, T., & Schwark, L. (2019). Heterocyte glycolipids indicate polyphyly of stigonematalean cyanobacteria. *Phytochemistry*, 166, 112059. <https://doi.org/10.1016/j.phytochem.2019.112059>
- Becerra-Absalón, I., Johansen, J. R., Osorio-Santos, K., & Montejano, G. (2020). Two new *Oculatella* (Oculatellaceae, Cyanobacteria) species in soil crusts from tropical semi-arid uplands of México. *Fottea*, 20(2), 160–170. <https://doi.org/10.5507/fot.2020.010>
- Belnap, J. (2003). The world at your feet: desert soil crusts. *Front Ecol Environ*, 1(5), 181–189.
- Belnap, J., & Lange, O. L. (2013). Structure and Functioning of Biological Soil Crusts: a Synthesis (Vol. 150; J. Belnap & O. L. Lange, eds.). https://doi.org/10.1007/978-3-642-56475-8_33
- Bohunická, M., Pietrasiak, N., Johansen, J. R., Gomez, E. B., Hauer, T., Gaysina, L. A., & Lukesova, A. (2015). *Roholtiella*, gen. nov. (Nostocales, Cyanobacteria)-- a tapering and branching cyanobacteria of the family Nostocaceae. *Phytotaxa*, 197(2), 84–103.
- Boyer, S. L., Flechtner, V. R., & Johansen, J. R. (2001). Is the 16S-23S rRNA internal transcribed spacer region a good tool for use in molecular systematics and population genetics? A case study in cyanobacteria. *Molecular Biology and Evolution*, 18(6), 1057–1069. <https://doi.org/10.1093/oxfordjournals.molbev.a003877>

- Boyer, S. L., Johansen, J. R., Flechtner, V. R., & Howard, G. L. (2002). Phylogeny and genetic variance in terrestrial *Microcoleus* (Cyanophyceae) species based on sequence analysis of the 16S rRNA gene and associated 16S-23S ITS region. *Journal of Phycology*, 38(6), 1222–1235. <https://doi.org/10.1046/j.1529-8817.2002.01168.x>
- Bu, C., Wu, S., Zhang, K., Yang, Y., & Gao, G. (2015). Biological soil crusts: An eco-adaptive biological conservative mechanism and implications for ecological restoration. *Plant Biosystems*, 149(2), 364–373. <https://doi.org/10.1080/11263504.2013.819820>
- Buck, B. J., & Van Hoesen, J. G. (2002). Snowball morphology and SEM analysis of pedogenic gypsum, southern New Mexico, U.S.A. *Journal of Arid Environments*, 51(4), 469–487. <https://doi.org/10.1006/jare.2001.0849>
- Cameron, R. E. (1960). Communities of soil algae occurring in the Sonoran Desert in Arizona. *J. Ariz. Acad. Sci*, 1, 85–88.
- Carmichael, W. W. (1986). Isolation, culture, and toxicity testing of toxic freshwater cyanobacteria (blue-green algae). In Shilov, V. [Ed.] *Fundamental Research in Homogenous Catalysis*, 3.
- Curtis, T. P., Head, I. M., Lunn, M., Woodcock, S., Schloss, P. D., & Sloan, W. T. (2006). What is the extent of prokaryotic diversity? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 361(1475), 2023–2037. <https://doi.org/10.1098/rstb.2006.1921>
- Dong, H., Rech, J. A., Jiang, H., Sun, H., & Buck, B. J. (2007). Endolithic cyanobacteria in soil gypsum: Occurrences in Atacama (Chile), Mojave (United States), and Al-Jafr Basin (Jordan) Deserts. *Journal of Geophysical Research*, 112(G2), G02030. <https://doi.org/10.1029/2006JG000385>

- Erwin, P.M. & Thacker, R.W. (2008). Cryptic diversity of the symbiotic cyanobacterium *Synechococcus spongiarum* among sponge hosts. *Molecular Ecology* 17: 2937–2947.
- Escudero A, Palacio S, Maestre FT, Luzuriaga AL (2015) Plant life on gypsum: a review of its multiple facets. *Biological Reviews of the Cambridge Philosophical Society* 90(1): 1–18. <https://doi.org/10.1111/brv.12092>
- Evans, R.D. & Johansen, J.R. 1999. Microbiotic crusts and ecosystem processes. *Critical Reviews in Plant Sciences* 18:183-225.
- Flechtner, V. R., Boyer, S. L., Johansen, J. R., & DeNoble, M. L. (2002). *Spirirestis rafaensis* gen. et sp. nov. (Cyanophyceae), a new cyanobacterial genus from arid soils. *Nova Hedwigia*, 74(1–2), 1–24. <https://doi.org/10.1127/0029-5035/2002/0074-0001>
- Flechtner, V. R., Johansen, J. R., & Belnap, J. (2008). The Biological Soil Crusts of the San Nicolas Island: Enigmatic Algae from a Geographically Isolated Ecosystem. *Western North American Naturalist*, 68(4), 405–436. <https://doi.org/10.3398/1527-0904-68.4.405>
- Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical science*, 457-472.
- Grondin, A., & Johansen, J. (1993). Microbial spatial heterogeneity in microbiotic crusts in Colorado National Monument. I. Algae. *Great Basin Naturalist*, 53(1), 4.
- Herrero, J., Artieda, O., & Hudnall, W. H. (2009). Gypsum, a tricky material. *Soil Science Society of America Journal*, 73(6), 1757–1763. <https://doi.org/10.2136/sssaj2008.0224>
- Johansen, J. R. (1984). Response of soil algae to a hundred-year storm in the Great Basin Desert, U.S.A. *Phycos*, 23(1), 51–54.
- Johansen, J. R. (1993). Cryptogamic Crusts of Semiarid and Arid Lands of North America. *Journal of Phycology*, 29(2), 140–147. <https://doi.org/10.1111/j.0022-3646.1993.00140.x>

- Johansen, J. R., Kovacik, L., Casamatta, D. A., Iková, K. F., & Kaštovský, J. (2011). Utility of 16S-23S ITS sequence and secondary structure for recognition of intrageneric and intergeneric limits within cyanobacterial taxa: *Leptolyngbya corticola* sp. nov. (*Pseudanabaenaceae*, *Cyanobacteria*). *Nova Hedwigia*, 92(3–4), 283–302. <https://doi.org/10.1127/0029-5035/2011/0092-0283>
- Johansen, J. R., Olse, C. E., Lowe, R. E., Fučíková, K., & Casamatta, D. A. (2008). *Leptolyngbya* species from selected seep walls in the Great Smokey Mountains National Park. *Algological Studies*, 126, 21–36.
- Johansen, J. R., Rushforth, S. R., & Brotherson, J. D. (1981). Sub-aerial algae of Navajo National Monument, Arizona. *Great Basin Nat.*, 41, 433–439.
- Kocurek, G., Carr, M., Ewing, R., Havholm, K. G., Nagar, Y. C., & Singhvi, A. K. (2007). White Sands Dune Field, New Mexico: Age, dune dynamics and recent accumulations. *Sedimentary Geology*, 197(3–4), 313–331. <https://doi.org/10.1016/j.sedgeo.2006.10.006>
- Komárek, J. & Anagnostidis, K. 2005. *Cyanoprokaryota -2. Teil/2nd Part: Oscillatoriales*. Elsevier/Spektrum, Heidelberg, 759.
- Komárek, J., Taton, A., Sulek, J., Wilmotte, A., & Kaštovská, K. (2006). Ultrastructure and taxonomic position of two species of the cyanobacterial genus *Schizothrix*. *Cryptogamie Algologie*, 27(1), 53–62.
- Li, X., Huo, S., Zhang, J., Ma, C., Xiao, Z., Zhang, H., Xi, B., & Xia, X. (2019). Metabarcoding reveals a more complex cyanobacterial community than morphological identification. *Ecological Indicators*, 107, 105653. <https://doi.org/10.1016/j.ecolind.2019.105653>

- Machado De Lima, N. M., & Branco, L. H. Z. (2020). Biological soil crusts: New genera and species of Cyanobacteria from Brazilian semi-arid regions. *Phytotaxa*, 470(4), 263–281. <https://doi.org/10.11646/phytotaxa.470.4.1>
- Maestre, F. T., Bowker, M. A., Cantón, Y., Castillo-Monroy, A. P., Cortina, J., Escolar, C., Escudero, A., Lazaro, R., & Martínez, I. (2011). Ecology and functional roles of biological soil crusts in semi-arid ecosystems of Spain. *Journal of Arid Environments*, 75(12), 1282–1291. <https://doi.org/10.1016/j.jaridenv.2010.12.008>
- MacKeigan, P. W., Garner, R. E., Monchamp, M. È., Walsh, D. A., Onana, V. E., Kraemer, S. A., Pick, F. R., Beisner, B. E., Agbeti, M. D., Barbosa da Costa, N., Shapiro, B. J. & Gregory-Eaves, I. (2022). Comparing microscopy and DNA metabarcoding techniques for identifying cyanobacteria assemblages across hundreds of lakes. *Harmful Algae*, 113, 102187.
- Mai, T., Johansen, J. R., Pietrasiak, N., Bohunická, M., & Martin, M. P. (2018). Revision of the Synechococcales (Cyanobacteria) through recognition of four families including Oculatellaceae fam. nov. and Trichocoleaceae fam. nov. and six new genera containing 14 species. *Phytotaxa*, 365(1), 1–59. <https://doi.org/10.11646/phytotaxa.365.1.1>
- Mehda, S., Muñoz-Martín, M. Á., Oustani, M., Hamdi-Aïssa, B., Perona, E., & Mateo, P. (2022). Lithic cyanobacterial communities in the polyextreme Sahara Desert: Implications for the search for the limits of life. *Environmental Microbiology*, 24(1), 451–474. <https://doi.org/10.1111/1462-2920.15850>
- Menéndez-Serra, M., Triadó-Margarit, X., Castañeda, C., Herrero, J., & Casamayor, E. O. (2019). Microbial composition, potential functional roles and genetic novelty in gypsum-

- rich and hypersaline soils of Monegros and Gallocanta (Spain). *Science of The Total Environment*, 650, 343–353. <https://doi.org/10.1016/j.scitotenv.2018.09.050>
- Miller, M. A., Schwartz, T., Hoover, P., Yoshimoto, K., Sivagnanam, S., & Majumdar, A. (2015). The CIPRES workbench: a flexible framework for creating science gateways. In *Proceedings of the 2015 XSEDE Conference: Scientific Advancements Enabled by Enhanced Cyberinfrastructure* (pp. 1-8).
- Moreira C. Fernandes, V., Giraldo-Silva, A., Roush, D. & Garcia-Pichel, F. (2021), Coleofasciculaceae, a Monophyletic Home for the *Microcoleus steenstrupii* Complex and Other Desiccation-tolerant Filamentous Cyanobacteria. *J. Phycol.*, 57: 1563-1579. <https://doi.org/10.1111/jpy.13199>
- Mühlsteinová, R., Johansen, J. R., Pietrasiak, N., Martin, M. P., Osorio-Santos, K., & Warren, S. D. (2014). Polyphasic characterization of *Trichocoleus desertorum* sp. nov. (Pseudanabaenales, Cyanobacteria) from desert soils and phylogenetic placement of the genus *Trichocoleus*. *Phytotaxa*, 163(5), 241–261. <https://doi.org/10.11646/phytotaxa.163.5.1>
- Muller, C. T., Moore, M. J., Feder, Z., Tiley, H., & Drenovsky, R. E. (2017). Phylogenetic patterns of foliar mineral nutrient accumulation among gypsophiles and their relatives in the chihuahuan desert. *American Journal of Botany*, 104(10), 1442–1450. <https://doi.org/10.3732/ajb.1700245>
- Novichkova-Ivanova, L. N. (1980). *Soil Algae of the Sahara-Gobi Desert Region*. Nauka, Lennigrad, 256 pp. (In Russian).
- Nübel, U., Garcia-Pichel, F., & Muyzer, G. (1997). PCR primers to amplify 16S rRNA genes from cyanobacteria. *Applied and environmental microbiology*, 63(8), 3327-3332.

- Osorio-Santos, K., Pietrasiak, N., Bohunická, M., Miscoe, L. H., Kováčik, L., Martin, M. P., & Johansen, J. R. (2014). Seven new species of *Oculatella* (Pseudanabaenales, Cyanobacteria): taxonomically recognizing cryptic diversification. *European Journal of Phycology*, 49(4), 450–470. <https://doi.org/10.1080/09670262.2014.976843>
- Perkerson, R. B., Johansen, J. R., Kováčik, L., Brand, J., Kaštovský, J., & Casamatta, D. A. (2011). A unique pseudanabaenalean (cyanobacteria) genus *Nodosilinea* gen. nov. based on morphological and molecular data. *Journal of Phycology*, 47(6), 1397–1412. <https://doi.org/10.1111/j.1529-8817.2011.01077.x>
- Pietrasiak, N., Johansen, J. R., & Drenovsky, R. E. (2011). Geologic composition influences distribution of microbiotic crusts in the Mojave and Colorado Deserts at the regional scale. *Soil Biology and Biochemistry*, 43(5), 967–974. <https://doi.org/10.1016/j.soilbio.2011.01.012>
- Pietrasiak, N., Mühlsteinová, R., Siegesmund, M. A., & Johansen, J. R. (2014). Phylogenetic placement of *Symplocastrum* (Phormidiaceae, Cyanophyceae) with a new combination *S. californicum* and two new species: *S. flechtnerae* and *S. torsivum*. *Phycologia*, 53(6), 529–541. <https://doi.org/10.2216/14-029.1>
- Pietrasiak, N., Osorio-Santos, K., Shalygin, S., Martin, M. P., & Johansen, J. R. (2019). When is a lineage a species? A case study in *Myxacorys* gen. nov. (Synechococcales: Cyanobacteria) with the description of two new species from the Americas. *Journal of Phycology*, (June). <https://doi.org/10.1111/jpy.12897>
- Pietrasiak, N., Reeve, S., Osorio-Santos, K., Lipson, D. A., & Johansen, J. R. (2021). *Trichotorquatus* gen. Nov. - a new genus of soil cyanobacteria discovered from American drylands 1. *Journal of Phycology*, 57(3), 886–902. <https://doi.org/10.1111/jpy.13147>

- Řeháková, K., Johansen, J. R., Casamatta, D. A., Xuesong, L., & Vincent, J. (2007). Morphological and molecular characterization of selected desert soil cyanobacteria: Three species new to science including *Mojavia pulchra* gen. et sp. nov. *Phycologia*, 46(5), 481–502. <https://doi.org/10.2216/06-92.1>
- Shalygin, S., Kavulic, K.J., Pietrasiak, N., Bohunická, M., Vaccarino, M.A., Chesarino, N.M. & Johansen, J.R. (2019). Neotypification of *Pleurocapsa fuliginosa* and epitypification of *P. minor* (Pleurocapsales): resolving a polyphyletic cyanobacterial genus. *Phytotaxa* 392 (4):245–263.
- Shields, L. M., Mitchell, C., & Drouet, F. (1957). Alga- and Lichen-Stabilized Surface Crusts as Soil Nitrogen Sources Author (s): Lora Mangum Shields , Charles Mitchell and Francis Drouet Published by : Botanical Society of America ALGA- AND LICHEN-STABILIZED SURFACE CRUSTS AS SOIL NITROGEN SOURCES ' L. American Journal of Botany, 44(6), 489–498.
- Siegesmund, M. A., Johansen, J. R., Karsten, U., & Friedl, T. (2008). COLEOFASCICULUS GEN. NOV. (CYANOBACTERIA): MORPHOLOGICAL AND MOLECULAR CRITERIA FOR REVISION OF THE GENUS MICROCOLEUS GOMONT 1. *Journal of Phycology*, 44(6), 1572–1585. <https://doi.org/10.1111/j.1529-8817.2008.00604.x>
- Strunecký, O., Ivanova, A. P., & Mareš, J. (2023). An updated classification of cyanobacterial orders and families based on phylogenomic and polyphasic analysis. *Journal of Phycology*, 59(1), 12–51. <https://doi.org/10.1111/jpy.13304>
- Verheye, W. H., & Boyadgiev, T. G. (1997). Evaluating the land use potential of gypsiferous soils from field pedogenic characteristics. *Soil Use and Management*, 13(2), 97–103. <https://doi.org/10.1111/j.1475-2743.1997.tb00565.x>

- Willmotte, A., Auwera, G. V. D., & Wachter, R. D. (1993). Structure of the 16S rRNA of the thermophilic cyanobacterium *Chlorogloeopsis* HTF ('*Mastigocladus laminosus* hTF') strain PCC7518, and phylogenetic analysis. *Federation of European Biochemical Societies*, 317, 96–100. [https://doi.org/10.1016/0014-5793\(93\)81499-P](https://doi.org/10.1016/0014-5793(93)81499-P)
- Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic acids research*, 31(13), 3406-3415.

Table 1. GPS coordinates and date of soil sample collection for each plot in White Sands National Park, New Mexico. The plot numbers are determined by the landform number, followed by the plot number. *Trichocoleus* strains were found in the plots marked with an “*”, and strains were named according to their plot number. For example, “WS4-1-AS1” uses the “WS4-1” to identify the plot from which it was collected.

Plot	Date finished	GPS coordinates	<i>Trichocoleus</i> strains collected
4-1*	6/20/2019	32.708425, -106.3620444	WS4-1-AS1
4-2	7/5/2019	32.7221028, -106.3522417	
4-3*	7/1/2019	32.7297278, -106.3438722	WS4-3-AS3
6-3*	7/4/2019	32.6927611, -106.3274694	WS6-3-JRJ2
6-4	6/20/2019	32.7005528, -106.3063972	
6-5	6/19/2019	32.7114667, -106.2865611	
7-2*	7/13/2019	32.6758472, -106.3463222	WS7-2-JRJ1
7-3	6/25/2019	32.6835028, -106.3228	
7-5*	6/15/2019	32.6912687, -106.3004474	WS7-5-JRJ8
8-2	6/23/2019	32.6811667, -106.3042444	
8-3*	6/26/2019	32.7035560, -106.2557220	WS8-3-JRJ2
8-4	7/3/2019	32.6912333, -106.2685556	

FIGURE DESCRIPTIONS

Figure 1. Map of the Tularosa Basin with White Sands National Park (previously a national monument). All samples were collected within the area defined by the black circle (Modified from Kocurek et al. 2007)

Figure 2. Bayesian inference phylogeny based on aligned 16S rRNA sequences with maximum likelihood values mapped onto the nodes. “*” indicates that full support was given for bootstrap (100) or posterior probability (1). “-” indicates that the value of support is less than half for bootstrap (<50) or posterior probability (<0.50). The genus *Trichocoleus* is expanded to show 16S rRNA relationships between all available sequences with proposed species groups named. Strains isolated from White Sands are indicated in red highlighting.

Figure 3. Maximum parsimony tree based on the 16S-23S ITS rRNA region with Bayesian inference values mapped onto the nodes. “*” indicates that full support was given for bootstrap (100) or posterior probability (1). “-” indicates that the value of support is less than half for bootstrap (<50) or posterior probability (<0.50). All sequences are in the genus *Trichocoleus* with proposed species groups named. Strains isolated from White Sands are indicated in red highlighting.

Figure 4. Putative secondary structures for the D1-D1' helix in *Trichocoleus* species. Circled nucleotides show variation in sequence in the taxon listed as having that structure. A. *T. desertorum sensu stricto*. B. *T. badius*, C. *Trichocoleus* sp. 1, sp. 8, and sp. 11. D. *Trichocoleus* sp. 2, E. *Trichocoleus* sp. 3 F. *Trichocoleus* sp. 4. G. *Trichocoleus* sp. 5, H. *Trichocoleus* sp. 6 and sp. 18, I. *Trichocoleus* sp. 7, J. *Trichocoleus* sp. 9 and sp. 10. K. *Trichocoleus* sp. 12, L. *Trichocoleus* sp. 13, M. *Trichocoleus* sp. 14 and sp. 17. N. *Trichocoleus* sp. 15, O. *Trichocoleus*

sp. 16 WS7-2-JRJ1, P. *T. caatingensis*, Q. Alternative structure for *T. desertorum*. R. Alternative structure for *T. badius*. See text for further explanation.

Figure 5. Putative secondary structures for the BoxB helix in *Trichocoleus* species. Circled nucleotides show variation in sequence in the taxon listed as having that structure. A. *T. desertorum sensu stricto*. B. *T. badius* and sp. 5, C. *Trichocoleus* sp. 1, D. *Trichocoleus* sp. 2, E. *Trichocoleus* sp. 3 F. *Trichocoleus* sp. 4, sp. 11, sp. 14. G. *Trichocoleus* sp. 6 sp. 18, H. *Trichocoleus* sp. 7, I. *Trichocoleus* sp. 8, J. *Trichocoleus* sp. 9. K. *Trichocoleus* sp. 9 and 10, L. *Trichocoleus* sp. 12, M. *Trichocoleus* sp. 13., N. *Trichocoleus* sp. 15, O. *Trichocoleus* sp. 16, P. *Trichocoleus* sp. 17, Q. *T. caatingensis*. See text for further explanation.

Figure 6. Putative secondary structures for the V3 helix in *Trichocoleus* species. Circled nucleotides show variation in sequence in the taxon listed as having that structure. A. *T. desertorum sensu stricto* and sp. 16, B. *T. badius*, C. *Trichocoleus* sp. 1, D. *Trichocoleus* sp. 2, E. *Trichocoleus* sp. 2 F. *Trichocoleus* sp. 3, G. *Trichocoleus* sp. 4, H. *Trichocoleus* sp. 5, I. *Trichocoleus* sp. 6, J. *Trichocoleus* sp. 7, sp. 11, and sp. 15 (differs in circled nt). K. *Trichocoleus* sp. 8, L. *Trichocoleus* sp. 9, M. *Trichocoleus* sp. 9, N. *Trichocoleus* sp. 10, O. *Trichocoleus* sp. 12, P. *Trichocoleus* sp. 13 and 18, Q. *Trichocoleus* sp. 14, R. *Trichocoleus* sp. 17, S. *T. caatingensis*. See text for further explanation.

Figure 7. Images showing diagnostic morphological characteristics for proposed species (all at 1000X, scale = 10 μ m). A-G: *Trichocoleus* sp. 1 WS7-5-JRJ8, H-M: *Trichocoleus* sp. 2 ATA2-1-CV2, N-Q: *Trichocoleus* sp. 3 ATA3-4Q-KO9.

Figure. 8 Images showing diagnostic morphological characteristics for proposed species (all at 1000X, scale = 10 μ m). A-F: *Trichocoleus* sp. 4 ATA14-3-RM35 and ATA14-3-RM36, G-M:

Trichocoleus sp. 5 ATA1-4-CV12 and ATA1-4-KO5, N-S: *Trichocoleus* sp. 6 WJTs and CMT-1BRIN-NPC4B.

Figure 9 Images showing diagnostic morphological characteristics for proposed species (all at 1000X, scale = 10 μ m). A-E: *Trichocoleus* sp. 8 ATA12-5-KO5, F-J: *Trichocoleus* sp. 14 WS6-3-AS2 and WS4-1-AS1, F-J: CXA25 and SNI-TA9-AZ2, K-Q: *Trichocoleus* sp. 15.

Figure 10 Images showing diagnostic morphological characteristics for proposed species (all at 1000X, scale = 10 μ m). A-F: *Trichocoleus* sp. 16 WS7-2-JRJ1, G-K: *Trichocoleus* sp. 17 WS4-3-AS3.

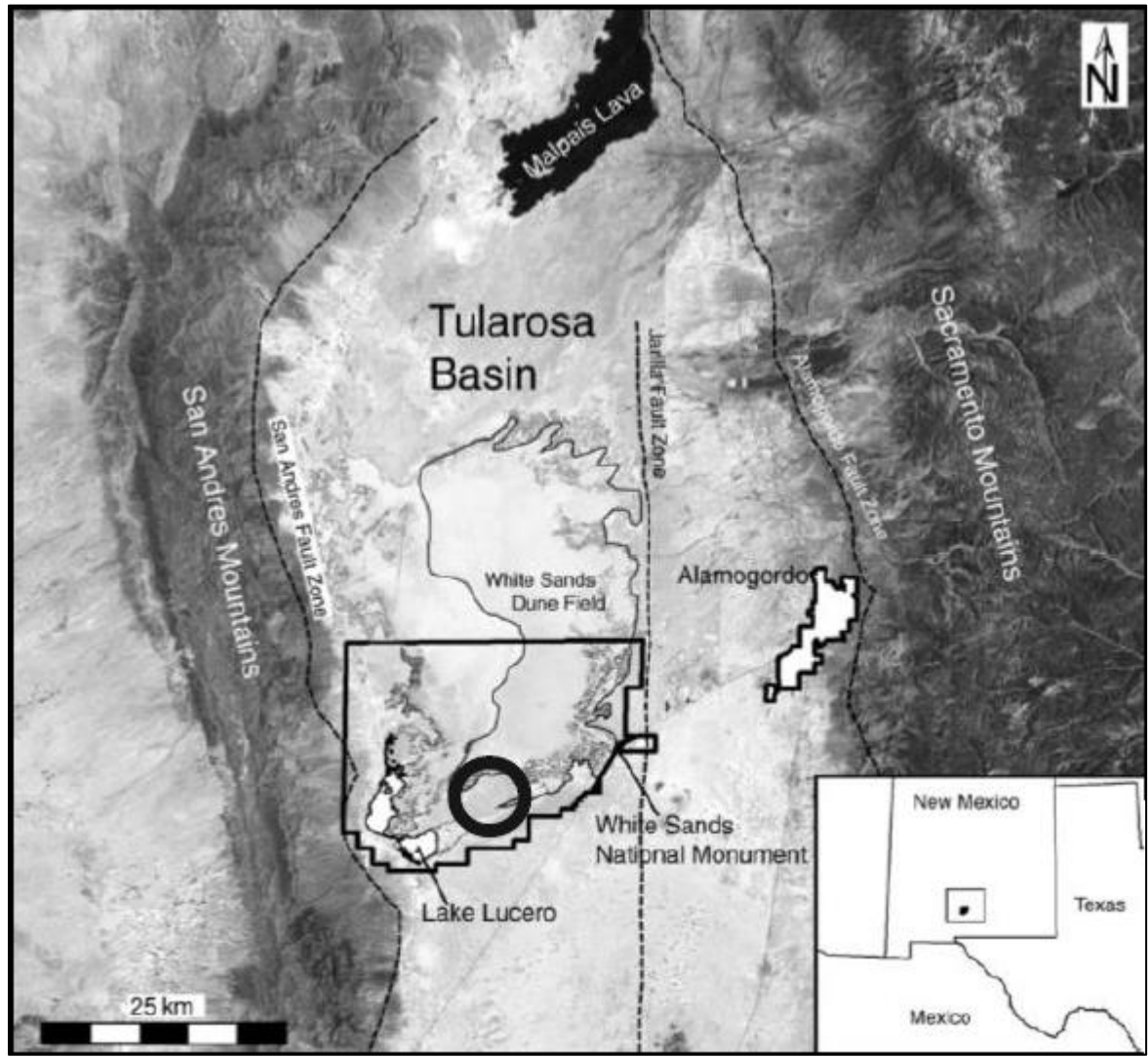


Figure 1

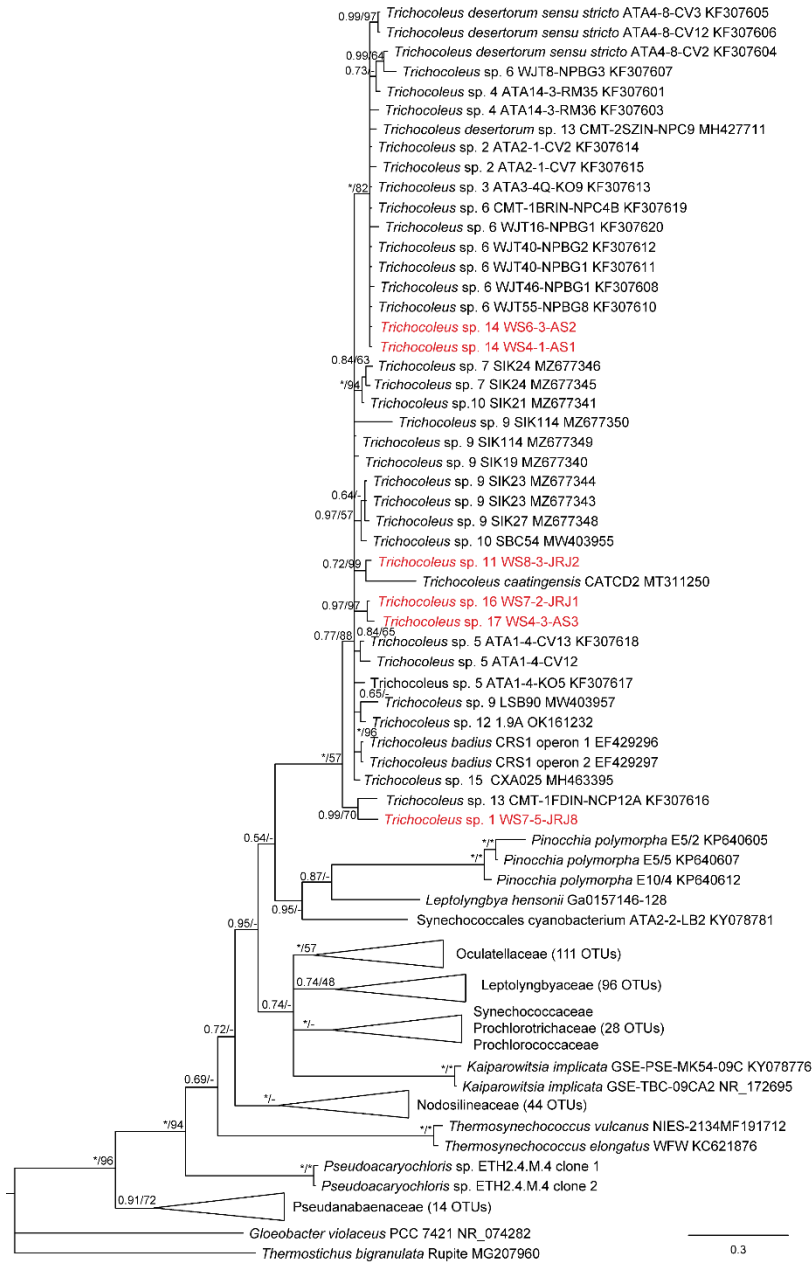


Figure 2

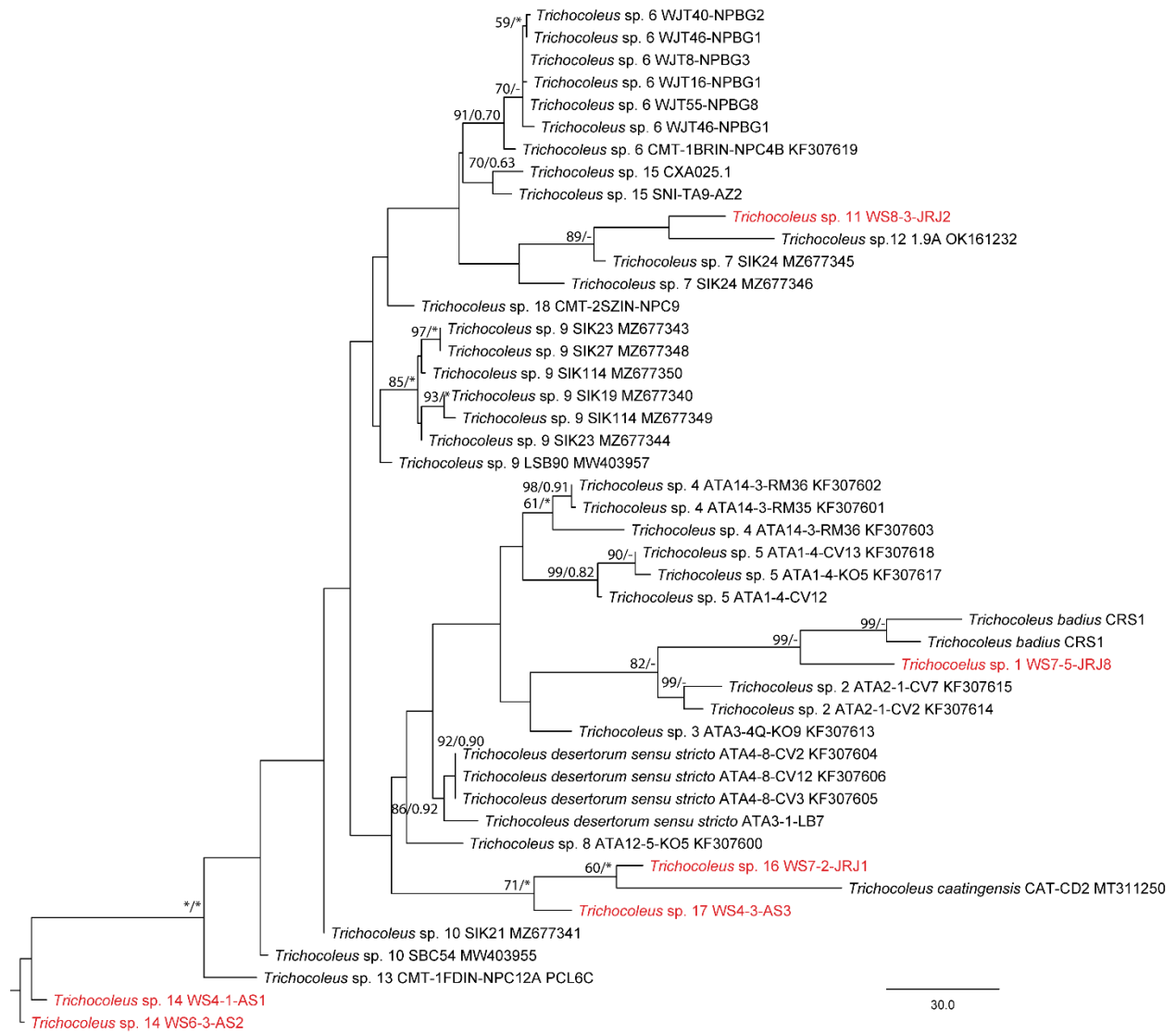


Figure 3

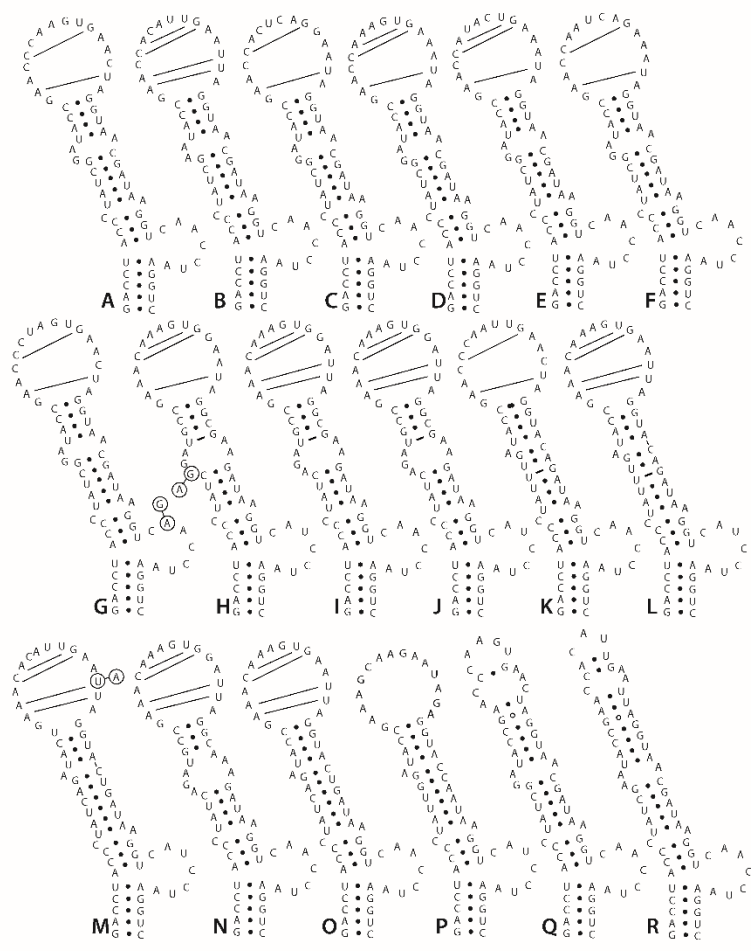


Figure 4

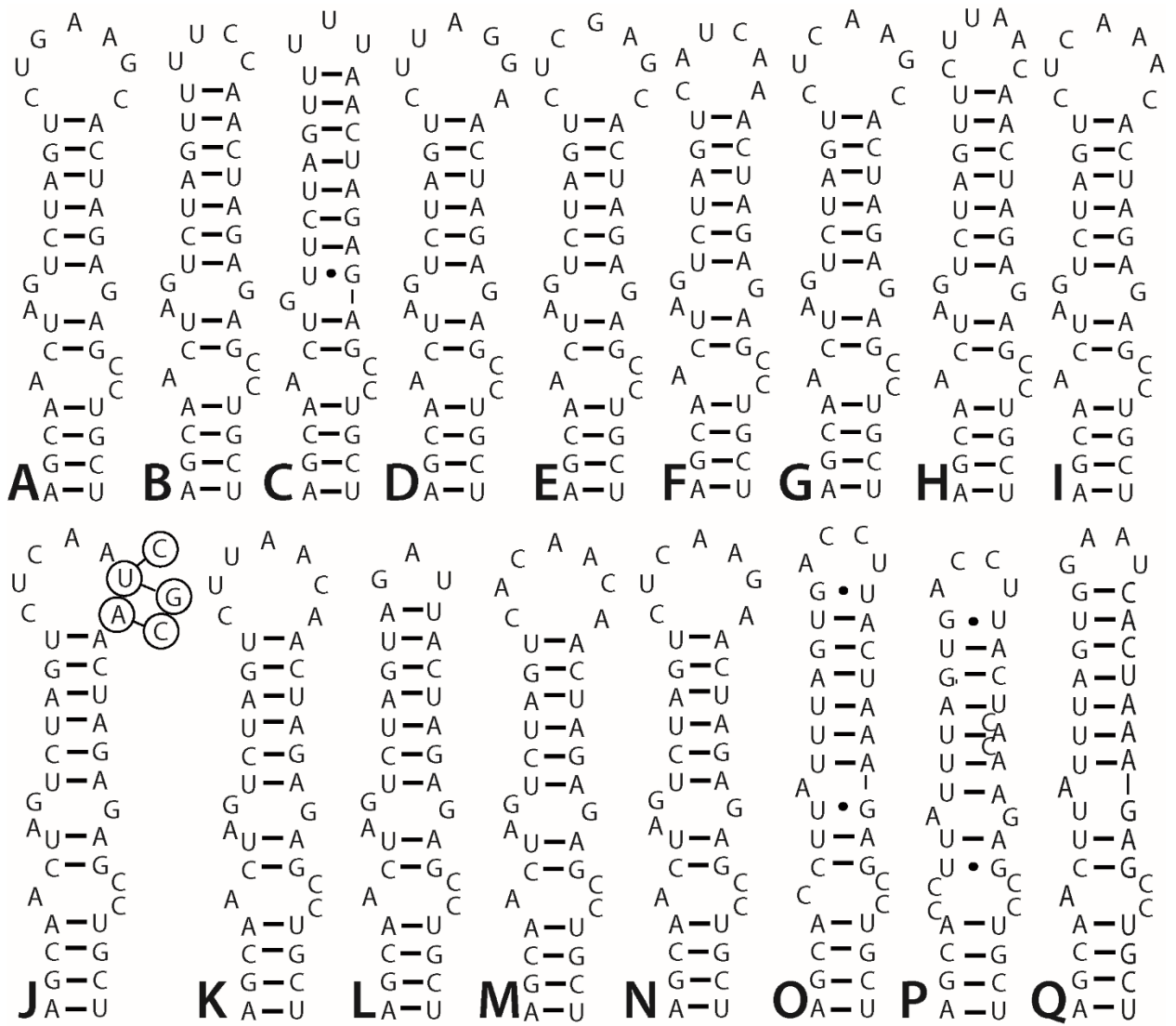


Figure 5

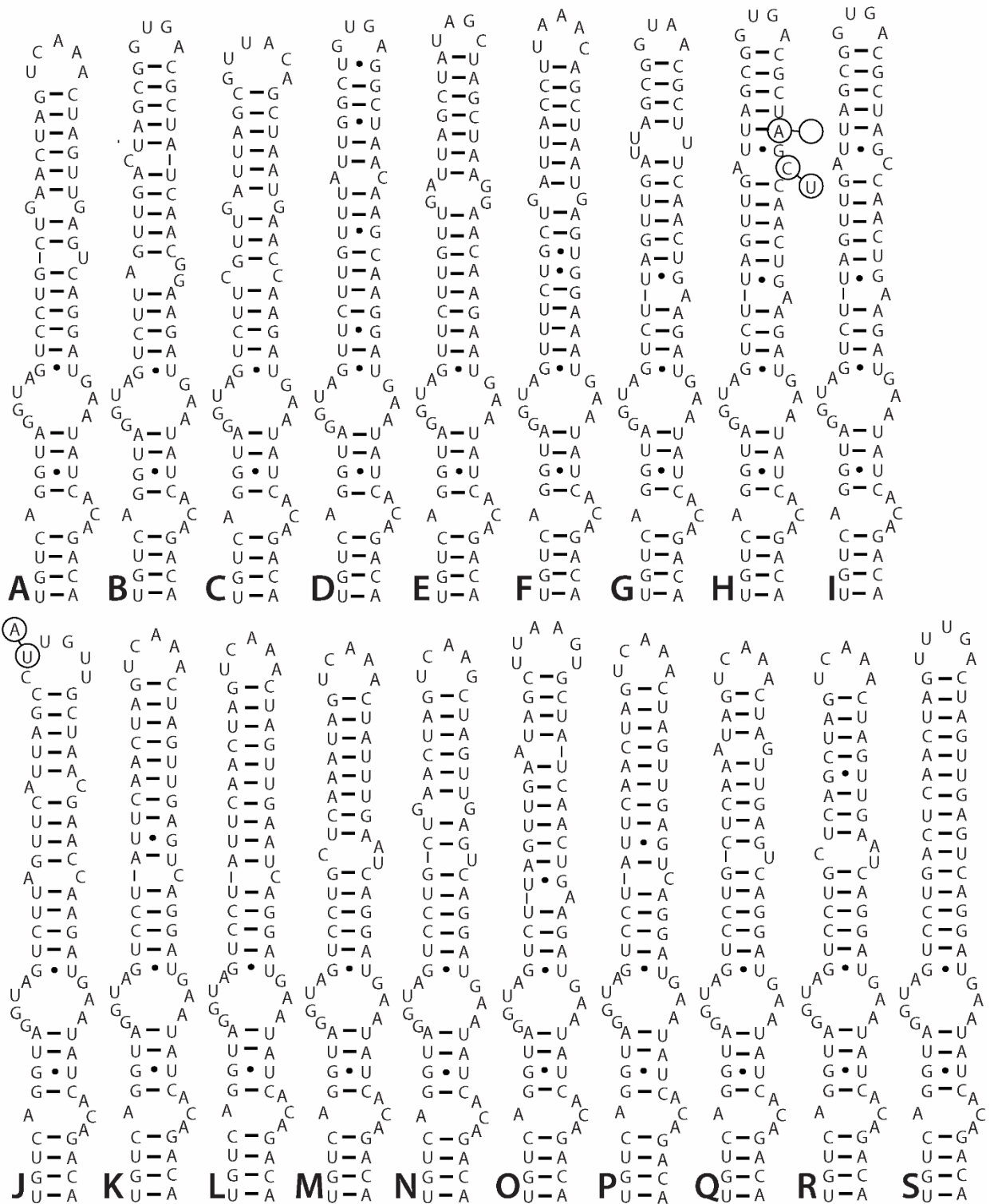


Figure 6

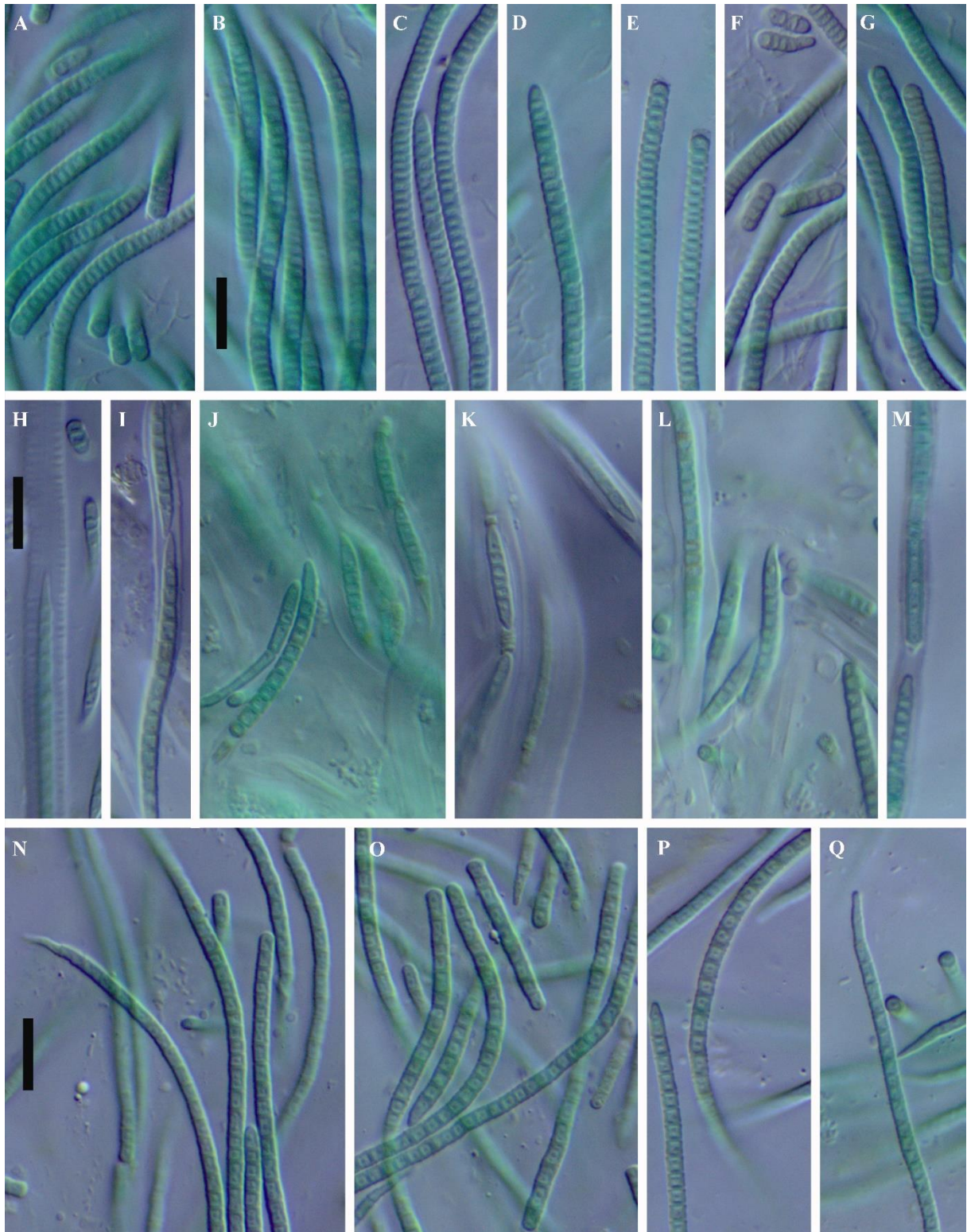


Figure 7

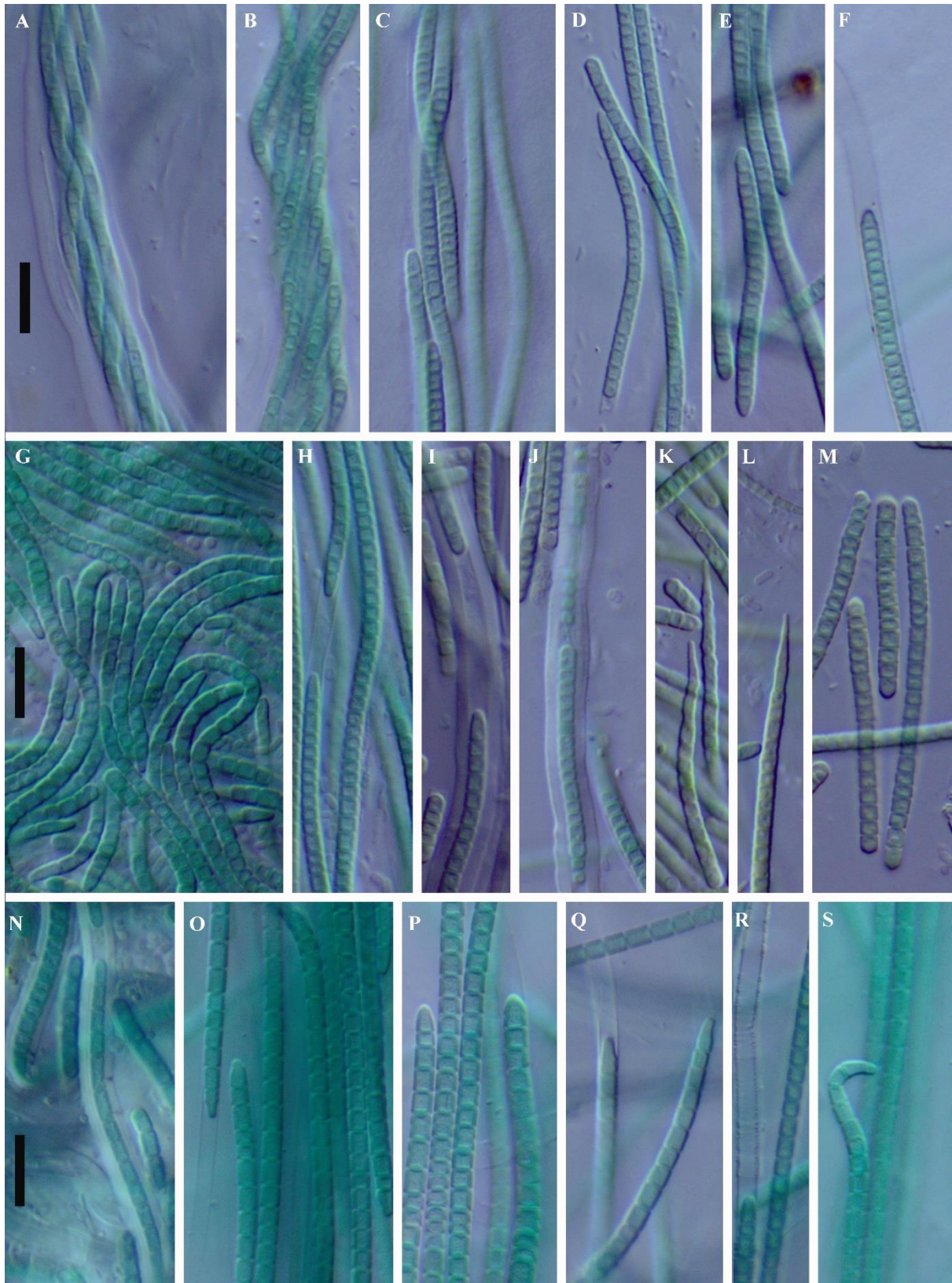


Figure 8

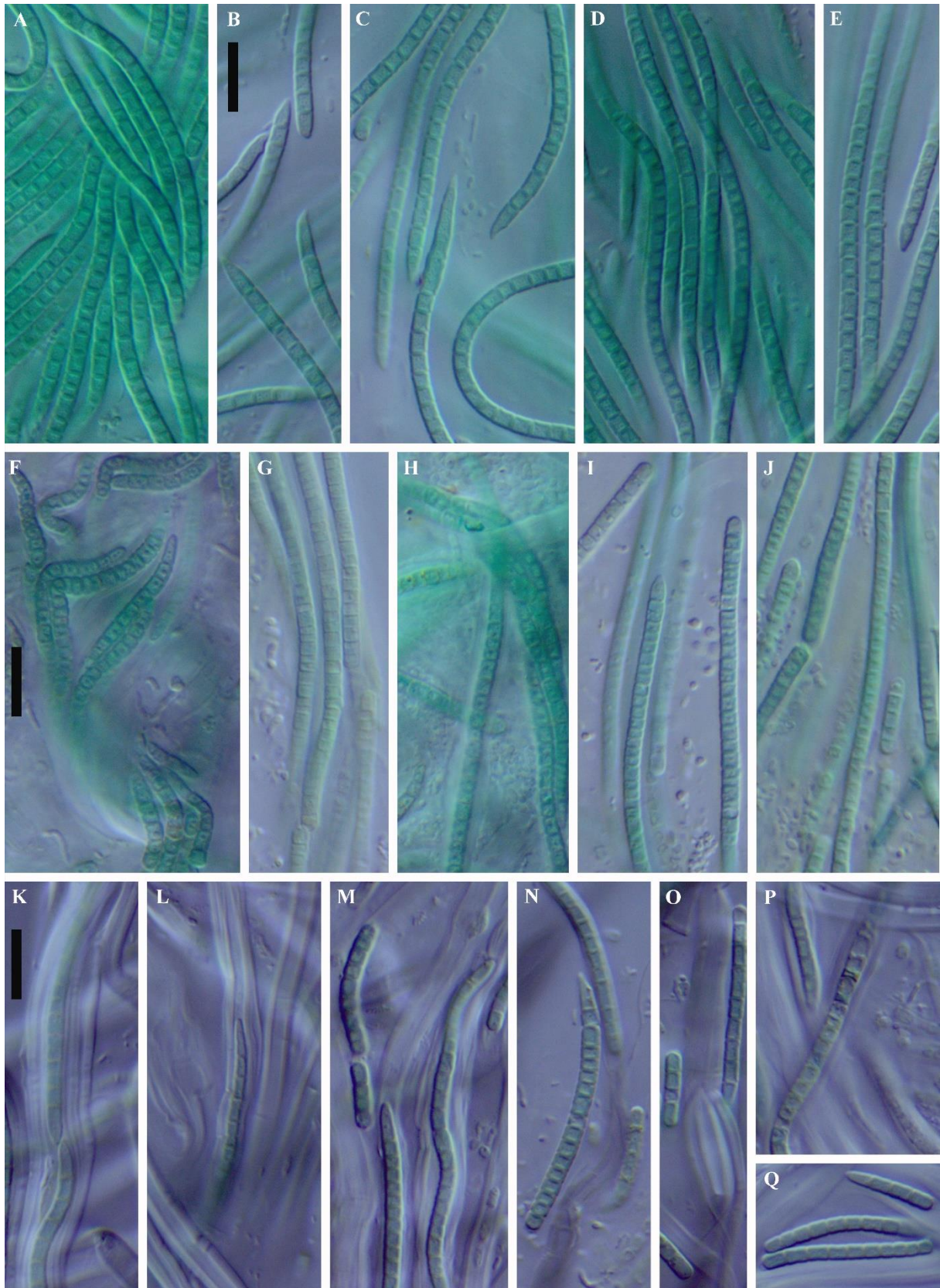


Figure 9

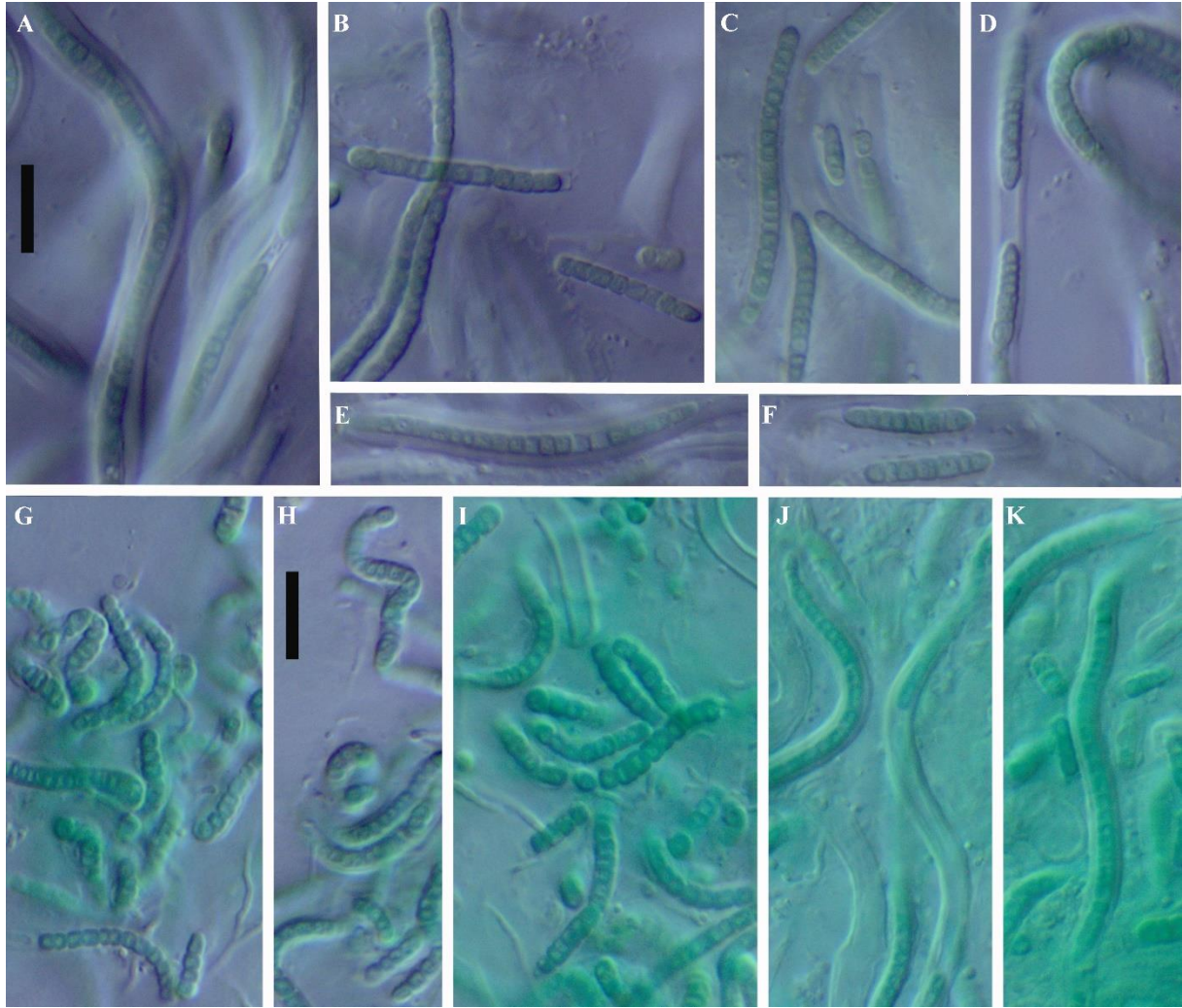


Figure 10