Fottea, Olomouc, 18(1): 130–136, 2018
DOI: 10.5507/fot.2017.023

Difference without distinction? Gaps in cyanobacterial systematics; when more is just too much

Petr Dvořák[1], Eva Jahodářová[1], Dale A. Casamatta[2*], Petr Hašler[1] & Aloisie Poulíčková[1]

[1] *Department of Botany, Faculty of Science, Palacký University Olomouc, Šlechtitelů 27, CZ–783 71 Olomouc, Czech Republic*
[2] *Department of Biology, University of North Florida, 1 UNF Drive, Jacksonville, Florida, USA;* *Corresponding author e–mail: dcasamat@unf.edu, phone: +1 904 620 1936

Abstract: Cyanobacteria are amongst the most abundant, ubiquitous, ecologically and evolutionarily significant microbes on Earth. Unique among the Bacteria in their capacity to be identified using morphology, understanding the evolutionary relationships and describing the diversity of this lineage is both important and challenging. The advent of modern sequencing technology has proven a boon to those studying cyanobacterial systematics as it has provided copious amounts of sequence data (mainly of the 16S rRNA gene sequence). However, this influx of data has also led to taxonomic confusion and recognition of polyphyly in many genera. Thus, the purpose of this paper is to describe this apparent paradox of increasing data yet poor phylogenetic resolutions by employing the Poisson Tree Process (PTP) algorithm and to propose some ameliorative efforts.

Key words: convergence, cyanobacteria, DNA database, metagenomics, phylogeny, species concept, taxonomy

Introduction

The cyanobacteria represent a large and diverse phylum of photo–oxygenic bacteria. They exist in aquatic, terrestrial and subaerial environments, inhabiting a wide variety of environments spanning the poles to the equator, from freshwaters to marine habitats. Being so widely distributed, they provide significant inputs to both the global oxygen and nitrogen cycles (Whitton & Potts 2000). Although cyanobacteria are an evolutionary and ecologically important group of organisms, significant gaps exist regarding their taxonomy and systematics. These problems are exacerbated due to difficulty of culturing, low sampling efforts outside the temperate zone, and problematic species concepts (for reviews see Castenholz 1992; Komárek 2003; Johansen & Casamatta 2005; Komárek et al. 2014; Dvořák et al. 2015a, b).

As a whole, the cyanobacteria exhibit an extensive amount of phenotypic variability. Their classification was initially based on features such as morphology of filaments, cells, sheaths, types of branching, cell differentiation, reproduction, etc., while employing botanical nomenclature due to their similarity to eukaryotic algae (Komárek & Anagnostidis 1998, 2005; Komárek 2013). Unfortunately, some of these characters (e.g., sheaths) have been shown to be phenotypically plastic, and thus their use in phylogenetic reconstructions remains open to debate (e.g., Casamatta & Vis 2003). Over the last two decades, molecular markers, most notably the 16S rRNA gene sequence, have shown that some morphological features do not necessarily correspond to phylogenetic reconstructions elucidated by molecular methods. More recently, it has been suggested that many of the most common, traditional genera of cyanobacteria are in fact polyphyletic (reviewed in Komárek et al. 2014). For example, Robertson et al. (2001) found five polyphyletic lineages in the morphologically simple, coccoid genus *Synechococcus*. A more recent analysis by Dvořák et al. (2014a) found 12 lineages in this same genus and proposed that cyanobacteria undergo serially convergent events due to genome dynamics through horizontal gene transfer and homologous recombination. These and similar findings provide support for splitting polyphyletic genera into smaller, monophyletic lineages and subsequently describing new taxa. However, some cyanobacterial lineages are morphologically indistinguishable due to convergent evolution and their lack of identifiable character poor morphologies (Dvořák et al. 2014a; Komárek et al. 2014).

Since the majority of prokaryotes cannot be cultured using standard laboratory techniques (e.g. Amann et al. 1995), a complete description of their biodiversity

seems to be an impossible endeavor. Fortuitously, recent advances in metagenomics (genetic material recovered directly from uncultured organisms from environmental samples) have offered a direct approach to circumvent this limitation. However, this approach is used infrequently in taxonomy due to logistical constraints. The concept has recently been proposed under the International Code for Nomenclature of Prokaryotes (ICNP, http://icsp.org/) by expansion of the *Candidatus* species concept (Konstantinidis & Rosselló–Móra 2015). As a result, this approach allows for naming of taxa that cannot be grown and maintained in culture.

Conversely, the International Code for Algae, Fungi and Plants (ICN, http://www.iapt–taxon.org/nomen/main.php), which is the primary code employed with cyanobacterial taxa, makes provision for the description of new taxa based on type material stored as dried biomass. Thus, new approaches to elucidating cyanobacterial diversity such as the use of single filament PCR may be applied, although results employing this approach are still scarce in the literature. For example, Mareš et al. (2015) performed a revision of selected Stigonematales cyanobacteria and Hašler et al. (2014a) revised the culture–resistant genus *Komvophoron* utilizing this approach. In any case, it should be noted that the biodiversity of prokaryotes is so vast (with potentially many millions of species) that description of taxa using metagenomics tools would help to elucidate the uncultured (and undescribed) majority of prokaryotes (Konstantinidis & Rosselló–Móra 2015). Cyanobacterial systematic endeavors are further complicated by the fact that they have traditionally been named under the ICN, but may also be validly described under the ICNP.

Nabout et al. (2013) noted that the total number of described cyanobacterial species is 2698, as obtained from the continuously updated CyanoDB database (database version 2013, Komárek & Hauer 2011). The authors estimated that an expected number of species of cyanobacteria is 6280 (with a confidence interval of 4402 to 8159), the upper limit of which is close to an earlier estimation of 8000 species by Guiry (2012). It should be noted that these quantifications are mostly based on traditional (and often flawed due to convergent evolution) morphological approaches using a phenetic species concept under the ICN. The introduction of molecular markers for unraveling the systematics of cyanobacteria has allowed researchers to use species concepts derived from evolutionary principles, such as the phylogenetic or monophyletic species concepts, which are increasingly being employed by cyanobacterial taxonomists (reviewed in Johansen & Casamatta 2005).

Currently, there is no definitive indication of how many species of cyanobacteria may exist, how many have been described using phylogenetic based species concepts, or how many species may be anticipated among sequences gathered from metagenomics data without culturing. Thus, the purpose of this paper is to use the available sequences of cyanobacteria to elucidate some of the phylogenetic relationships. Further, we seek to quantify an overlap between described species and available DNA sequence–based species (e.g., taxa with names and 16S sequence data) and we will use these results to evaluate the reliability of database DNA sequence data for taxonomic and metagenomics purposes.

Material and Methods

We acquired a comprehensive dataset of 16S rRNA sequences of cyanobacteria from GenBank (database version 21st January 2015) using search query (((900:2500[Sequence Length]) AND cyanobacteria [organism]) AND 16S) to ensure we gathered sufficiently long sequences for analyses. Stored GenBank sequences were obtained mostly using Sanger dideoxy sequencing and 454 pyrosequencing. Multiple sequence alignment (MSA; Dataset S1 Supporting Information) was performed using MAFFT (Katoh et al. 2002) with automatic diagnostics of alignment parameters. Large gaps and uninformative regions were eliminated from the alignment using Gblocks (Castresana 2000) with the following settings: Maximum Number Of Contiguous Nonconserved Positions set to 50, Minimum Length Of A Block 2, and Allowed Gap Positions: half. Identical sequences were removed from alignment and their list has been stored for further analyses (Tab. S1 in the Supporting Information). Gblocks was employed in order to restrict sequence length and remove as much uninformative sequence as possible to shorten computing time. We chose our cutoff b.p. criteria for two reasons. First, many researchers only amplify a portion of the 16S rRNA gene and thus complete genes are not typically available. Second, we could not employ shorter reads (<899 b.p.) for MSA because these sequence fragments may be from the beginning or end of the 16S rRNA gene and might not overlap. The upper limit was set due to genome assemblies, which are very large and would make MSA impossible. The ITS region was not employed due to variability within and among strains, which would lead to erroneous MSA.

A phylogenetic reconstruction was performed using maximum likelihood criterion in RaxML 8.0.0 (Stamatakis 2014) under substitution model GTR+GAMMA (Dataset S2 in the Supporting Information). A phylogenetic reconstruction using neighbor joining optimality criterion was performed in MEGA 6 (Tamura et al. 2013) using Kimura 2–parameter model (Dataset S3 in the Supporting Information). Species were delimited in the Python programmed package PTP (Poisson Tree Process; Zhang et al., 2013), which uses the phylogenetic species concept *sensu* Eldredge & Cracraft (1980) and most recently modified by Nixon & Wheeler (1990). PTP uses substitution per site difference (not a genetic distance with a particular cutoff) for species identification. It assumes that there is a significantly (statistically speaking) higher evolutionary distance (measured in substitution per site) among species than within species. Since it uses substitution per site, it does not require an ultrametric tree. Thus, PTP also reflects different evolutionary rates within different species.

PTP produces a list of identified species or operational taxonomic units (OTUs). For consistency and simplicity, we will use the term "PTP–defined species" throughout the text. Our analysis employed units that include at least one cultured strain and PTP–defined species composed of sequences from uncultured cyanobacteria. We employed and assessed taxa with cultured strains if they were identified as an existing and validly

Fottea, Olomouc, 18(1): 130–136, 2018
DOI: 10.5507/fot.2017.023

described species under the ICN or the ICNP since species of cyanobacteria are considered under both codes. However, it should be noted that cyanobacterial species described under ICNP are valid under ICN, but ICNP does not accept species under ICN (Oren 2011). Also, species described under ICNP are much less frequent than under ICN. PTP outperforms other automatic delimitation algorithms such as GMYC (Fujisawa & Barraclough 2013; Pons et al. 2006), because it does not require an often error–prone ultrametric tree and it has significantly faster performance on large datasets (Zhang et al. 2013).

We used the following procedure to identify which sequences employed in the GenBank dataset corresponded to definitely named and reliably identified PTP–defined species (e.g., units consisting of both cultured and uncultured taxa which are definitively identified to species by a taxonomic expert; including species described based on single cell/filament PCR techniques). We also considered all identical sequences, which were subsequently assigned to PTP–defined species (Tab. S1 in the Supporting Information). First, all PTP–defined species with sequences deposited solely as sequences from "uncultured cyanobacteria" were removed from the list of definitively named and reliably identified species (excluding species described based on single cell/filament PCR techniques). Second, PTP–defined species containing only sequences without species epithets or with unsure species epithets were removed from the list of reliably identified species (GenBank accession numbers and available literature were checked to ensure missing epithets). Finally, we did an extensive GenBank and literature search to confirm whether a particular PTP–defined species contains any described species valid under either the ICN or the ICNP (including synonyms). Species with a *Candidatus* status under the ICPN were not employed in this analysis. If the PTP–defined species contained at least one definitively named and reliably identified species, we employed it in further analyses. The phylogeny reconstructed for this paper has been used as a template for taxonomic decisions. Multiple sequence alignment, template trees, and a list of identified cultured or uncultured species are available in Supporting Information.

Results and Discussion

How many "unculturable" cyanobacterial taxa actually exist?

The final alignment contained 10037 sequences with 4983 non–identical sequences. Employing the PTP species delimitation articulated earlier, we recovered 2741 PTP–defined species of cyanobacteria (Fig. 1, Table S2 in the Supporting Information). It has been suggested that only ca. 1% of prokaryotes may be cultured (Amann et al. 1995), but this figure is postulated mainly with heterotrophic bacteria in mind. This is in sharp contrast to our findings that only 51% of PTP–defined species were from uncultured environmental samples. However, this number may be biased since only sequences longer than 900 b.p. were used, while many metagenomics analyses employ far shorter sequences, often on the order of 300–500 b.p. Nevertheless, we excluded these shorter fragments since they would not be appropriate to utilize in our multiple sequence alignment.

This finding may imply either a high success rate of culturing the vast majority of cyanobacteria or that habitats with high uncultured cyanobacterial diversity are lacking in metagenomic data. For instance, fine, freshwater sediments (epipelon) are known for very low sampling efforts (Poulíčková et al. 2014), even though they are inhabited by a complex and diverse community of cyanobacteria and other algae (Mann et al. 2008; Poulíčková et al. 2008). Moreover, the sampling efforts of cyanobacteriologists are traditionally concentrated in temperate zones, although tropical latitudes may serve as hot spots of biodiversity (see Hohner–Divine et al. 2004 for review). Describing the taxonomy of tropical cyanobacteria is an active research endeavor, especially in aerophytic habitats (e.g. Fiore et al. 2007; Neustupa & Škaloud 2008). However, a dearth of tropical papers is evident when compared to temperate zones, as seen in the number of Web of Knowledge indexed papers (Dvořák et al. 2015a). Together with the rapid pace of new species descriptions in recent years (Komárek et al. 2014), we conclude that most of the cyanobacteria biodiversity remains undescribed and that a sizable portion of uncultured biodiversity may remain unnamed.

On the incompatibility of molecular and traditional cyanobacterial systematics

We found that only 571 PTP–defined species (20.9%) may be assigned to definitively named and reliably identified species under either the ICN or ICNP (Fig. 1, see Supporting Information for details). Furthermore, the 571 PTP–defined species represents 12.7–21.2% of the total described species under the ICN included in either AlgaeBase (Guiry & Guiry 2015; 4484) or the CyanoDB (information from Nabout et al. 2013; 2698 species), respectively. It should be noted that species under the ICN and ICNP are described using various species concepts, mostly phenetic with some form of a phylogenetic species concept, but a majority of recently published descriptions and revisions do not cite any particular species concept (for a review, see Johansen & Casamatta 2005). Species identified using the PTP are purely phylogenetic species as advocated by Nixon & Wheeler (1990). Therefore, the total numbers of species may differ based on the species concepts employed. Since modern molecular methods are able to provide greater taxonomic resolution due to a bounty of additional data (e.g. Johansen & Casamatta 2005; Erwin & Thacker 2008), we consider the discrepancy between the number of phylogenetic species and number of described species to reflect that the total diversity is underestimated.

The vast increase in described cyanobacterial diversity over the last 20 years is a direct result of molecular methods, which has in turn significantly influenced the taxonomic reasoning of cyanobacteriologists. Both 16S rRNA gene sequence data and metagenomic or community level assessments appear to have taken different directions. 16S rRNA sequences are deposited in DNA databases at near exponential levels yet are themselves flooded by ambiguously identified sequences resulting from previously, unequivocally identified sequences.

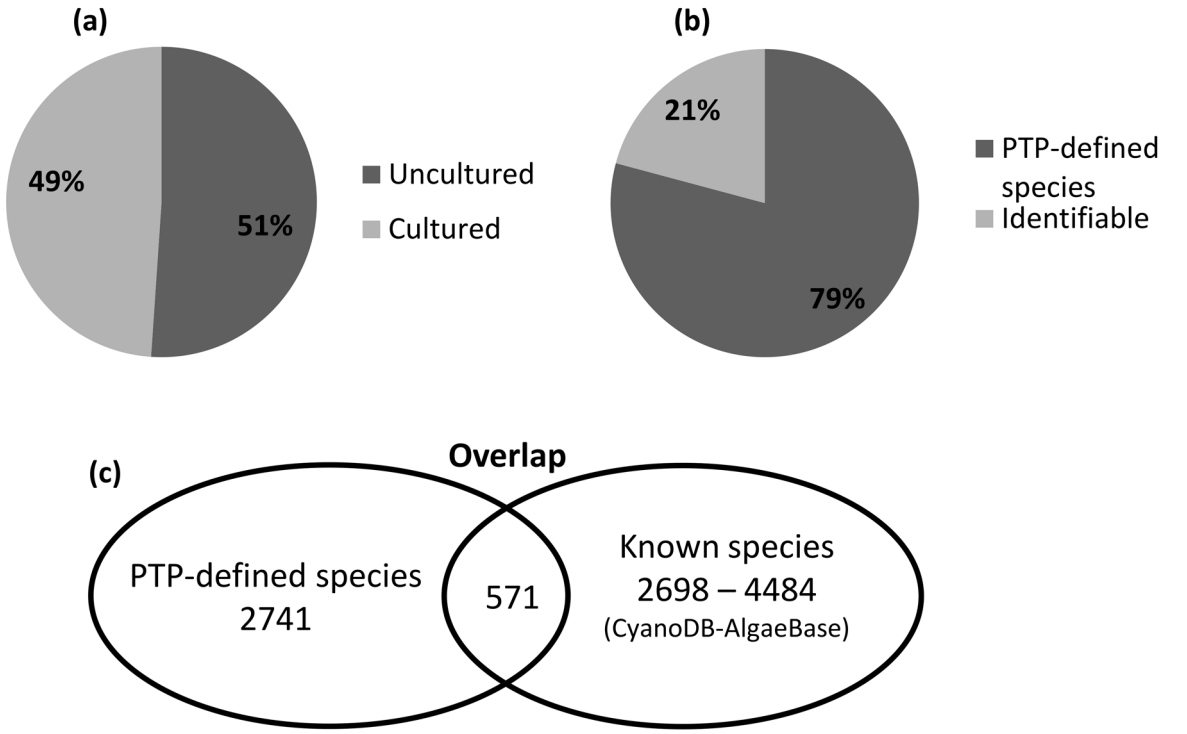

Fig. 1. Graphical representation of PTP analysis results. a) The ratio of uncultured and cultured PTP–defined species. b) The portion of PTP–defined species, which may be assigned to existing described species based on CyanoDB. c) A Venn diagram representing an overlap of PTP–defined species and species present in databases CyanoDB and AlgaeBase.

Thus, the difference between both systems is expanding daily due to erroneously identified sequences and species.

One possible reason for such a striking difference may rise from the entangled evolutionary relationships among cyanobacteria. As previously noted, a majority of described genera are currently considered polyphyletic (Komárek et al. 2014; Fig. S1 in the Supporting Information). The ultimate goal in reconstructing evolutionary relationships is that separate lineages should be classified and named based on rules of cladistics and phylogeny. Our results show that not much effort has been given to resolve this conundrum until now. The most striking example is the genus *Synechococcus*, which contains 12 polyphyletic lineages (Dvořák et al. 2014a), of which only one has been properly revised under the ICN (Dvořák et al. 2014b) and none under the ICNP. For example, the marine picoplanktic *Synechococcus* are one of the most abundant organisms on earth (Flombaum et al. 2013), but this lineage lacks proper taxonomic treatment. There are multiple species within this group, as previously recognized by Robertson et al. (2001) and our analysis in this paper. Thus, we may anticipate further disagreement of molecular data with traditional systems. Similarly, extensive evidence of polyphyly has been found in other taxa–rich genera such as *Leptolyngbya* (Osorio–Santos et al. 2014; Jahodářová et al. 2017), *Phormidium* (Hašler et al. 2012), or *Cylindrospermum* (Johansen et al. 2014).

Another source of taxonomic inconsistency may emerge from practical use of different taxonomic classifications by researchers. First, there is no consensus among researchers whether to use the ICN or ICNP. While the ICN accepts all names generated under the ICNP, the reciprocal is not so. Recently, however, Pinevich (2015) proposed changes to some principals of the ICNP, which would allow acceptance of valid species published under the ICN to also be valid under the ICNP. Second, the traditional ICN has many versions of the nomenclatural schemes, for example those proposed by Geitler (1932), Desikachary (1959), or Komárek et al. (2014). Unfortunately, there is typically no corresponding change to the appropriate cyanobacterial databases, leading to taxonomic confusion. Third, and perhaps most importantly, carefully articulated taxonomic revisions are undertaken and published slowly, while ambiguously identified sequences are still accumulating very rapidly in GenBank. Coupled with the enigmatic evolutionary histories of most cyanobacteria, this can lead to much confusion and uncertainty. One possible remedy would be to establish an approved repository for all cyanobacterial taxonomic and systematic endeavors, easily accessible to all interested researchers.

Another potential source of inconsistency may be variability in the ease of DNA recovery resulting in over–representation of some groups and under–representation of others. While some taxa with thin or lacking sheathes are relatively easy to amplify, other lineages, especially those with firm, copious sheaths (e.g., *Nostoc*, *Petalonema*, *Gloeocapsa*) might be recalcitrant to DNA analyses (Mareš et al. 2015).

Fottea, Olomouc, 18(1): 130–136, 2018
DOI: 10.5507/fot.2017.023

A capability of PTP to recognize cyanobacterial species

Automatic species delimitation approaches are often at odds with manually curated delimitations. This paper represents a novel approach to apply automatic delimitation of species to cyanobacteria, barring barcoding efforts which have met with limited success (Eckert et al. 2015). Thus, we endeavor to evaluate the performance of PTP on several well–defined and revised cyanobacterial genera.

For example, PTP recognized 6 of 7 PTP–defined species within *Oculatella* (except *O. mojaviensis*; Osorio–Santos et al. 2014), while *O. subterranea* was divided into three PTP–defined species. In the genus *Nodosilinea*, only one species (*Nodosilinea* sp. NB1a–A5) has not been recognized by PTP. On the other hand, new sequences have appeared in GenBank since Perkerson et al. (2001) established *Nodosilinea*; we found three PTP–defined species of *Nodosilinea* sp. CENA 183, CENA 144, and CENA 137. Less persuasive results were shown by analysis of *Cylindrospermum*, which belongs to the heterocystous cyanobacteria. Recently revised by Johansen et al. (2014), PTP recognized *C. badium*, *C. moravicum*, and *C. marchicum*. However, *C. catenatum*, *C. pellucidum*, *C. licheniforme*, and *C. muscicola* were collapsed into one PTP–defined species and the same appeared also in cases of *C. allatosporum* and *C. maius*. Thus, we may conclude that PTP species delimitation provides a hint for species identification and enumeration, but results should be cautiously interpreted because there are certainly gaps in this method. In any case, the purpose of this paper is to show how significant the gaps are between molecular and phenotype based species delimitations. It seems that PTP adequately serves this purpose.

Conclusions: limitations of available DNA sequence data in systematics of cyanobacteria.

Cyanobacteria are certainly not unique, but they are challenging in the employment of molecular data for identification and taxonomy. The ambiguity of cyanobacterial systematics creates a potential pitfall for metagenomic research and DNA barcoding. Based on our results, reliable species identifications based solely on sequence data from DNA databases may be unlikely. Moreover, recent work has indicated that different species within a genus cannot be differentiated solely by 16S rRNA sequences alone. Thus, additional data, such as provided by the ITS region *rbc*L gene sequences, are increasingly being employed (e.g. Osorio–Santos et al. 2014). While traditional employment of the 16S rRNA gene might be sufficient for generic level assignments, we see many misidentified strains due to the polyphyletic nature of many currently circumscribed genera (Komárek et al. 2014). Thus, it is likely that cyanobacteria with entirely different evolutionary histories are conflagrated, as evidenced with *Synechococcus*. On another front, practical identifications of cyanobacteria (for applied phycology, ecology, and genetics) are still obtained mainly using morphological observations which may be increasingly problematic as more polyphyletic lineages are elucidated and rarely revised. Further, there is a lack of morphological apomorphies for some newly described lineages (Komárek et al. 2014; Dvořák et al. 2015b). The potential use of cyanobacteria for environmental bioassessment and biomonitoring efforts is largely limited by available sequence data (for those employing genetic markers) and proper species descriptions using phenotypic data (e.g. morphology, Manoylov 2014). For example, *Phormidium retzii*, considered one of the most commonly encountered lotic taxa in North America, has been shown to be a collection of cryptic taxa, and thus, further taxonomic revisions are warranted (Casamatta et al. 2003).

Taken together, we have shown that genetic and taxonomic databases (e.g. AlgaeBase, CyanoDB, GenBank, etc.) may not clearly articulate the diversity of cyanobacteria. Many species are ambiguously identified or come from uncultured specimens with a concurrent low certainty of proper identification. To avoid further confusion, we propose the following recommendations: Curation of taxonomic revisions and descriptions should be more widely linked with DNA databases and authors should be more actively involved with the curation of databases. For example, authors should update their sequence identifiers even after publication. Furthermore, cyanobacterial taxonomists using the ICN have an opportunity to make descriptions of new taxa without cultures. If they took advantage of this opportunity, it would allow for faster and more effective taxonomic revisions and naming of new taxa from unculturable species (e.g., *Johanseninema*; Hašler et al. 2014a, b). Taxonomists using ICNP may use the *Candidatus* concept to provide a putative name to uncultured taxa. However, *Candidatus* names are not valid under ICN or ICNP. Therefore, two names for the same species may be proposed, one for each code and eventually both validated, leading to heterotypic synonyms.

Revisions of polyphyletic genera are essential for proper identification of cyanobacteria. Without this effort, we would not be able to recognize a majority of species within polyphyletic clusters (e.g., *Leptolyngbya*, *Phormidium*, and *Synechococcus*).

Species identifications based on metagenomic data should be assessed more carefully. Sequence data should be named only based on properly described or bar–coded species. DNA database accessions should contain a statement containing a level of confidence and elucidation of the method of identification.

All revisions should be performed only with robust taxon sampling with numerous, abundant, phylogenetically relevant outgroups and sister taxa. For example, if only "*Leptolyngbya*" sequences are employed in an analysis this might mask polyphyletic relationships as one might recover a single, monophyletic clade. This might be illustrated with *Leptolyngbya nodulosa* which was originally described and phylogenetically analyzed with rather depauperate sister and outgroup taxa (all that

was available at the time) before subsequently being transferred to a new genus. Revisions using only phenotypic data are discouraged, due to serial convergence of morphotypes (Dvořák et al. 2014a, 2015b). It should be noted that cyanobacteria have a relatively rich herbaria presence in museums, and thus an investigation of type material is often possible and recommended (Komárek et al. 2014; Palinska & Surosz 2014). If type species are unavailable, they may be retypified based on recent samples from type locality or close to type locality.

Acknowledgements
This work was supported by European Social Fund Post–UP II CZ.1.07/2.3.00/30.0041 and Internal Grant Agency of Palacký University Prf–2017–001.

References

Amann, R.I.; Ludwig, W. & Schleifer, K.H. (1995): Phylogenetic identification and in situ detection of individual microbial cells without cultivation. – Microbioogical Reviews 59: 143–169.

Casamatta, D.A. & Vis, M.L. (2003): Ecological role of *Phormidium retzii* (Cyanobacteria) in a small Ohio woodland stream. – Journal of Freshwater Ecology 18: 395–404.

Casamatta, D.A.; Vis, M.L. & Sheath, R.G. (2003): Cryptic species in cyanobacterial systematics: a case study of *Phormidium retzii* (Oscillatoriales) using 16S rDNA and RAPD analyses. – Aquatic Botany 77: 295–309.

Castenholz, R.W. (1992): Species usage, concept, and evolution in the cyanobacteria (blue–green algae). – Journal of Phycology 28: 737–745.

Castresana, J. (2000): Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. – Molecular Biology and Evolution 17: 540–552.

Desikachary, T.V. (1959): Cyanophyta. – In: Randhawa, M.S. (ed.): I.C.A.R. monographs on algae. – 686 pp., Indian Council of Agriculture Research, New Dehli.

Dvořák, P.; Casamatta, D.A.; Hašler, P.; Ondřej, V.; Poulíčková, A. & Sanges, R. (2014a): *Synechococcus*: 3 billion years of global dominance. – Molecular Ecology 23: 5538–5551.

Dvořák, P.; Hindák, F.; Hašler, P.; Hindáková, A. & Poulíčková, A. (2014b): Morphological and molecular studies of *Neosynechococcus sphagnicola*, gen. et sp. nov. (Cyanobacteria, Chroococcales). – Phytotaxa 170: 24–34.

Dvořák, P.; Jahodářová, E.; Hašler, P.; Gusev, E. & Pulíčková, A. (2015a): A new tropical cyanobacterium *Pinocchia polymorpha* gen. et sp. nov. derived from genus *Pseudanabaena*. –Fottea 15: 113–120.

Dvořák, P.; Poulíčková, A.; Hašler, P.; Belli, M.; Casamatta, D.A. & Papini, A. (2015b): Species concepts and speciation factors in cyanobacteria, with connection to the problems of diversity and classification. – Biodiversity and Conservation 24: 739–757.

Eckert, E.M.; Fontaneto, D.; Coci, M. & Callieri, C. (2015): Does a barcoding gap exists in prokaryotes? Evidences from species delimitation in cyanobacteria. – Life 5: 50–64.

Eldredge, N. & Cracraft, J. (1980): Phylogenetic Patterns and the Evolutionary Process: Method and Theory in Comparative Biology. – 329 pp., Columbia University Press, New York, USA.

Erwin, P.M. & Thacker, R.W. (2008): Cryptic diversity of the symbiotic cyanobacterium *Synechococcus spongiarum* among sponge hosts. – Molecular Ecology 17: 2937–2947.

Fiore, M.F.; Sant'Anna, C.L.; Azevedo, M.T.P.; Komárek, J.; Kaštovský, J.; Sulek, J. & Lorenzi, A.S. (2007): The cyanobacterial genus *Brasilonema*, gen. nov., a molecular and phenotypic evaluation. – Journal of Phycology 43: 789–98.

Flombaum, P.; Gallegos, J.L.; Gordillo, R.A.; Rincon, J.; Zabala, L.L.; Jiao, N.; Karl, D.M.; Li, W.K.W.; Lomas, M. W.; Venezioano, D.; Vera, C.S.; Vrught, J.A. & Martiny A.C. (2013): Present and future global distributions of the marine cyanobacteria *Prochlorococcus* and *Synechococcus*. – Proceedings of the National Academy of Science U.S.A. 110: 9824–9829.

Fujisawa, T. & Barraclough, T.G. (2013): Delimiting species using single–locus data and the generalized mixed yule coalescent (GMYC) approach: a revised method and evaluation on simulated datasets. – Systematic Biology 62: 707–724.

Geitler, L. (1932): Cyanophyceae. – In: Rabenhorst L. (ed.): Kryptogamenflora. – pp. 673–1196, Verlagsgesellschaft Akademische, Leipzig.

Guiry, M.D. (2012): How many species of algae are there? – Journal of Phycology 48: 1057–1063.

Guiry, M.D. & Guiry, G.M. (2015): AlgaeBase. World–wide electronic publication, National University of Ireland, Galway. http://www.algaebase.org; searched on 02 September 2015.

Hašler, P.; Dvořák, P.; Johansen, J.R.; Kitner, M.; Ondřej, V. & Poulíčková, A. (2012): Morphological and molecular study of epipelic filamentous genera *Phormidium*, *Microcoleus* and *Geitlerinema* (Oscillatoriales, Cyanophyta/Cyanobacteria). – Fottea 12: 341–356.

Hašler, P.; Dvořák, P. & Poulíčková, P. 2014a. A new genus of filamentous epipelic cyanobacteria, *Johansenia*. – Preslia 86:81–94.

Hašler, P.; Dvořák, P. & Poulíčková, A. (2014b): *Johanseninema*, a corrected name for a recently described genus of filamentous epipelic cyanobacteria. – Preslia 86: 293–294.

Horner–Divine, M.C.; Carney, K.M. & Bohannan, B.J.M. (2004): An ecological perspective on bacterial biodiversity. – Proceedings of Royal Society of London B 271: 113–122.

Jahodářová, E.; Dvořák, P.; Hašler, P. & Poulíčková, A. (2017): Revealing the hidden tropical diversity among cyanobacteria using polyphasic approach, new genus *Onodrimia* (Synechococcales, Cyanobacteria). – Phytotaxa 326: 1–13.

Johansen, J.R.; Bohunická, M.; Lukešová, A.; Hrčková, K.; Vaccarino, M.A. & Chesarino N.M. (2014): Morphological and molecular characterization within 26 strains of the genus *Cylindrospermum* (Nostocaceae, Cyanobacteria), with description of three new species. – Journal of Phycology 50: 187–202.

Johansen, J.R. & Casamatta, D.A. (2005): Recognizing cyanobacterial diversity through adoption of a new species paradigm. – Algological Studies 117: 71–93.

Katoh, K.; Misawa, K.; Kuma, K. & Miyata, T. (2002): MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. – Nucleic Acids Research 30: 3059–3066.

Komárek, J. (2003): Problem of the taxonomic category "species" in cyanobacteria. – Algological Studies 109: 281–297.

Komárek, J. & Hauer, T. (2011): CyanoDB.cz—On–line database of cyanobacterial genera. Word–wide electronic publication, Univ. of South Bohemia & Inst. of Botany AS CR. http://www.cyanodb.cz. Accessed June 2011

Komárek, J.; Kaštovský, J.; Mareš, J. & Johansen, J.R. (2014): Taxonomic classification of cyanoprokaryotes (cyanobacterial genera) 2014 using a polyphasic approach. – Preslia 86: 295–335.

Komárek, J. (2013): Cyanoprokaryota. 3. Heterocytous genera. – In: Büdel, B.; Gärtner, G.; Krienitz, L. & Schagerl, M. (eds): Süsswasserflora von Mitteleuropa/Freshwater

Fottea, Olomouc, 18(1): 130–136, 2018
DOI: 10.5507/fot.2017.023

flora of Central Europe, Vol. 19/3. – 1130 pp., Springer Spektrum, Berlin, Heidelberg.

Komárek, J. & Anagnostidis, K. (1998): Cyanoprokaryota 1. Teil: Chroococcales. – In: Ettl, H.; Gärtner, G.; Heynig H. & Mollenhauer, D. (eds): Süsswasserflora von Mitteleuropa, Vol. 19/1. – 548 pp., Gustav Fischer, Jena–Stuttgart–Lübeck–Ulm.

Komárek, J. & Anagnostidis, K. (2005): Cyanoprokaryota 2. Teil Oscillatoriales. – In: Büdel, B.; Gärtner, G.; Krienitz, L. & Schagerl, M. (eds): Süßwasserflora von Mitteleuropa, Vol. 19/2. – 759 pp., Elsevier, München.

Konstantinidis, K.T. & Rosselló–Móra, R. (2015): Classifying the uncultivated microbial majority: a place for metagenomics data in the Candidatus proposal. – Systematic and Appied Microbiology 38: 223–230.

Mann, D.G.; Thomas, S.J. & Evans, K.M. (2008): Revision of the diatom genus *Sellaphora*: a first account of the larger species in the British Isles. – Fottea 8: 15–78.

Manoylov, K.M. (2014): Taxonomic identification of algae (morphological and molecular): species concepts, methodologies, and their implication for ecological bioassessment. – Journal of Phycology 50: 409–424.

Mareš, J.; Lara, Y.; Dadáková, I.; Hauer, T.; Uher, B.; Wilmotte, A. & Kaštovský, J. (2015): Phylogenetic analysis of cultivation–resistant terrestrial cyanobacteria with massive sheaths (*Stigonema* spp. and *Petalonema alatum*, Nostocales, Cyanobacteria) using single–cell and filament sequencing of environmental samples. – Journal of Phycology 51: 288–297.

Nabout, J.C.; da Silva Rocha, B.; Carneiro, F.M. & Sant'Anna, C.L. (2013): How many species of Cyanobacteria are there? Using a discovery curve to predict the species number. –Biodiversity and Conservation 22: 2907–2918

Neustupa, J. & Škaloud, P. (2008): Diversity of subaerial algae and cyanobacteria on tree bark in tropical mountain habitats. – Biologia 63: 806–812.

Nixon, K.C. & Wheeler, Q.D. (1990): An amplification of the phylogenetic species concept. –Cladistics 6: 211–223.

Oren, A. (2011): Cyanobacterial systematics and nomenclature as featured in the International Bulletin of Bacteriological Nomenclature and Taxonomy/International Journal of Systematic Bacteriology/International Journal of Systematic and Evolutionary Microbiology. – International Journal of Systematic and Evolutionary Microbiology 61: 10–15.

Osorio–Santos, K.; Pietrasiak, N.; Bohunická, M.; Miscoe, L.H.; Kovácik, L.; Martin, M.P. & Johansen, J.R. (2014): Seven new species of *Oculatella* (Pseudanabaenales, Cyanobacteria). – European Journal of Phycology 49: 450–470.

Palinska, K.A. & Surosz, W. (2014): Taxonomy of cyanobacteria: a contribution to consensus approach. – Hydrobiologia 740: 1–11.

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: 1397–1412.

Pinevich, A.V. (2015): Proposal to consistently apply the International Code of Nomenclature of Prokaryotes (ICNP) to names of the oxygenic photosynthetic bacteria (cyanobacteria), including those validly published under the International Code of Botanical Nomenclature (ICBN)/International Code of Nomenclature for algae, fungi and plants (ICN), and proposal to change Principle 2 of the ICNP. – International Journal of Systematic and Evolutionary Microbiology 65: 1070–1074.

Pons, J.; Barraclough, T.G.; Gomez–Zurita, J.; Carduso, A.; Duran, D.P.; Hazell, S.; Kamoun, S.; Sumlin, W.D. & Vogler, A.P. (2006): Sequence–based species delimitation for the DNA taxonomy of undescribed insects. – Systematic Biology 55: 595–609.

Poulíčková, A.; Dvořák, P.; Mazalová, P. & Hašler, P. (2014): Epipelic microphototrophs: an overlooked assemblage in lake ecosystems. – Freshwater Science 33: 513–523.

Poulíčková, A.; Hašler, P.; Lysáková, M. & Spears, B. (2008): The ecology of freshwater epipelic algae: an update. – Phycologia 47: 437–450.

Robertson, B.R.; Tezuka, N. & Watanabe, M. (2001): Phylogenetic analyses of *Synechococcus* strains (cyanobacteria) using sequences of 16S rDNA and part of the phycocyanin operon reveal multiple evolutionary lines and reflect phycobilin content. –International Journal of Systematic and Evolutionary Microbiology 51: 861–871.

Stamatakis, A. (2014): RAxML Version 8: A tool for Phylogenetic Analysis and Post–Analysis of Large Phylogenies. – Bioinformatics 30: 1312–1313.

Tamura, K.; Stecher, G.; Peterson, D.; Filipski, A. & Kumar, S. (2013): MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. – Molecular Biology and Evolution 30: 2725–2729.

Whitton, B.A. & Potts, M. (2000): The ecology of cyanobacteria. Their diversity in time and space. – 632 pp., Springer, Berlin, Germany.

Zhang, J.; Kapli, P.; Pavlidis, P. & Stamatakis, A. (2013): A general species delimitation method with applications to phylogenetic placements. – Bioinformatics 29: 2869–2876.

Supplementary material

the following supplementary material is available for this article:

Table S1. A list of identical sequences removed from multiple sequence alignment.
Table S2. All cultured cyanobacterial species analyzed in this paper. XLS table with all cultivable cyanobacterial species used for this paper. Each line represents one PTP–defined species. Identifiable (green; see methods section for definition) and unidentifiable (red) species are labeled.
Dataset S1. Multiple sequence alignment of cyanobacteria. A PHYLIP formatted 16S rRNA
multiple sequence alignment of cyanobacteria with removed duplicated sequences.
Dataset S2. Phylogenetic tree of cyanobacterial 16S rRNA in a Newick format reconstructed in
RaxML based on Dataset S1.
Dataset S3. Phylogenetic tree of cyanobacterial 16S rRNA in a Nexus format reconstructed in
MEGA (neighbor joining optimality criterion) based on Dataset S1.
Fig. S1. Phylogenetic relationships among cyanobacteria based on 16S rRNA. Parallel lines represent clades, which are on in a real scale. They were shortened due to their excessive length. Each collapsed clades is annotated by one strain name from each genus in particular clade.

This material is available as part of the online article (http://fottea.czechphycology.cz/contents)

Received September 24, 2017
Accepted November 7, 2017