



**HAL**  
open science

## Monospecific renaming within the cyanobacterial genus *Limnospira* (*Spirulina*) and consequences for food authorization

Théotime Roussel, Sébastien Halary, Charlotte Duval, Bérénice Piquet,  
Jean-Paul Cadoret, Léa Vernès, Cécile Bernard, Benjamin Marie

### ► To cite this version:

Théotime Roussel, Sébastien Halary, Charlotte Duval, Bérénice Piquet, Jean-Paul Cadoret, et al.. Monospecific renaming within the cyanobacterial genus *Limnospira* (*Spirulina*) and consequences for food authorization. *Journal of Applied Microbiology*, 2023, 134, 10.1093/jambio/lxad159 . mnhn-04288948

**HAL Id: mnhn-04288948**

**<https://mnhn.hal.science/mnhn-04288948>**

Submitted on 16 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Monospecific renaming within the cyanobacterial genus *Limnospira* (Spirulina) and consequences for food authorization

Théotime Roussel<sup>1,2</sup>, Sébastien Halary<sup>1</sup>, Charlotte Duval<sup>1</sup>, Bérénice Piquet<sup>3</sup>, Jean-Paul Cadoret<sup>2</sup>, Léa Vernès<sup>2</sup>, Cécile Bernard<sup>1,\*</sup>, Benjamin Marie<sup>1,\*</sup>

<sup>1</sup>UMR7245 MCAM MNHN-CNRS, Muséum National d'Histoire Naturelle, CP 39, 12 rue Buffon, F-75231 Paris Cedex 05, France

<sup>2</sup>Algama, 81 rue Réaumur, 75002 Paris, France

<sup>3</sup>Electron Microscopy Platform, Muséum National d'Histoire Naturelle, CP 39, 12 rue Buffon, F-75231 Paris Cedex 05, France,

\*Corresponding authors: UMR7245 MCAM MNHN-CNRS, Muséum National d'Histoire Naturelle, CP 39, 12 rue Buffon, F-75231 Paris Cedex 05, France. E-mail: cecile.bernard@mnhn.fr; benjamin.marie@mnhn.fr

## Abstract

**Aims:** The cyanobacterial genus, *Limnospira* (anc. *Arthrospira* Stizenberger ex Gomont 1892), commonly called “Spirulina”, is widely used for commercial purposes because of its high protein content and beneficial probiotic metabolites. Thus, the taxonomy of this genus is important because of its consequences for food applications.

**Methods and results:** We constructed a database with formation on all *Limnospira* strains plus new ones from 72 new French isolates. We used a polyphasic approach (phylogenetic, phylogenomic, presence or absence of coding DNA sequences, morphological, and ultrastructure analyses) to confirm that the species *A. platensis* belonged to the genus *Limnospira* (*L. platensis* Gomont comb. nov. Basionym. *Arthrospira platensis* Gomont 1892) and that the genus *Limnospira* was monospecific, only represented by *L. platensis*.

**Conclusions:** This study highlighted the large intra-specific diversity of *L. platensis*, independent of the affiliations of the phylogenetic clades or geographical location of the habitats and the subsequent physiological and metabolic plasticity.

## Significance and impact of study

The cyanobacterial genus, *Limnospira* (anc. *Arthrospira*), commonly called “Spirulina” is monospecific and the authorization for food applications should be extended to all *Limnospira* taxa, including the previously named *L. platensis*, *L. fusiformis*, *L. indica*, and *L. maxima*.

**Keywords:** cyanobacteria, *Limnospira*, spirulina, foods, genomics, identification, genotyping

## Introduction

The cyanobacteria are an ancient group of photoautotrophic microorganisms that have colonized diverse environments worldwide due to their huge genomic diversity and physiological plasticity (Whitton and Potts 2007). A total of 387 genera were listed in 2020, with nine different orders (Komárek 2020). As of 2013, 2698 species had been described, but model predictions estimated that up to 6000 were still undiscovered (Nabout et al. 2013). For the last five years, Google Scholar listed 58 publications that mentioned the description of new species of cyanobacteria, and PubMed had 46. Most of these studies used a polyphasic approach, combining molecular data, morphology, and ultrastructure (Duval et al. 2021, Rasouli-Dogaheh et al. 2022). This great species diversity was also manifested in the huge diversity of metabolites produced by the cyanobacteria, many of which showed biological activity and are being used in biotechnology for their potential in pharmaceutical, cosmetics, or food applications (Demay et al. 2019).

The genus *Limnospira* is one of the high-potential genera, well-known and commercialized as Spirulina. This cyanobacterium has been consumed by humans on several continents

for at least the past 500 years and is still traditionally eaten in Africa near Lake Chad, where it is known as “Dihé” (Abdulqader et al. 2000). The dry biomass of *Limnospira* is composed of up to 70% protein (Avila-Leon et al. 2012), primarily phycobiliproteins (up to 18% of total dry biomass) (Markou 2014) with antioxidant properties (Demay et al. 2019). *Limnospira* also contains antioxidant carotenoids, polyunsaturated fatty acids, and vitamins (Andrade et al. 2018, Edelmann et al. 2019). In addition to its nutritional properties, the popularity of *Limnospira* cultivation can be explained by its fast growth rate (more than one division per day) and its capacity to grow in alkaline medium (up to pH 12), which greatly reduces the risk of culture contamination by other photosynthetic microorganisms (Vonshak 1997).

Recent molecular and ultrastructure analyses split the genus *Arthrospira* into two different lineages: *Arthrospira jenniferi* described in 1892 by Stizenberger and Gomont (Gomont 1893), and the genus *Limnospira*, which contains the species previously named: *A. fusiformis*, *A. indica* and *A. maxima*, which become *L. fusiformis*, *L. indica*, and *L. maxima* (Nowicka-Krawczyk et al. 2019). However, the authors did not assign the species *A. platensis* to the new genus *Limnospira* because of the lack of information about the type strain. The

Received: January 7, 2023. Revised: June 27, 2023. Accepted: August 8, 2023

© The Author(s) 2023. Published by Oxford University Press on behalf of Applied Microbiology International. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

study of the taxonomy of cyanobacteria is important, especially for food applications. According to European regulations, only *A. platensis* is authorized for consumption in Europe (European Commission, Joint Research Centre 2021). For commercial purposes, it is therefore important to consider the changes of taxonomic affiliation within the *Limnospira* clade: *L. fusiformis*, *A. platensis*, *L. indica*, and *L. maxima*, *A. jeneri*, and *A. erdosensis* (Papapanagiotou and Gkelis 2019). New species have been proposed based on the unique metabolic and/or physiological capabilities of some strains: *A. nitrilium* (able to metabolize nitriles) and *A. sesilensis* (unable to glide) (Walter et al. 2017). However, these new species names have not been used in recent studies, and *Limnospira* taxonomy remains unresolved. Up to now, the determination of *Limnospira* (formerly *Arthrospira*) species was based on criteria used for cyanobacteria: single-locus phylogeny by 16S rRNA gene sequencing, internal transcribed spacers (ITS) between 16S and 23S rRNA genes, intergenic regions between *cpcB* and *cpcA* genes encoding for the C-phycoyanin beta and alpha chains (*cpcBA*-IGS) combined with morphology, secondary structure of the ITS, and geographic or habitat origin (Manen and Falquet 2022, Baurain et al. 2002, Dadhech et al. 2010, Cellamare et al. 2018, Papapanagiotou and Gkelis 2019). Other genes were once used for the phylogeny of these cyanobacteria like the *cpcHID* operon and the gas vesicle gene cluster *gvp* (Miklaszewska et al. 2012). In contrast, bacterial species delimitation currently involves whole-genome information, including the core genome (Chung et al. 2018) and average nucleotide identity (ANI) (Jain et al. 2018). Up to now, only one study used the genome to investigate the phylogenetic position of a *Limnospira* strain (Misztak et al. 2021).

Considering the cyanobacterial and bacterial criteria for species delimitation (16S rRNA and ITS identity and phylogenies, ITS secondary structure, phylogenomic, ANI, and core-genome identity, combined with morphology, biogeography, and habitat), we propose the hypothesis that the established phylogenetic clade *Limnospira* contains only one single species. All criteria and thresholds were applied for species delimitation (Kim et al. 2014, Chung et al. 2018, Jain et al. 2018, Berthold et al. 2022). To test this hypothesis, all data available in the literature and the 72 new French isolates of *Limnospira*, including 30 new sequenced genomes were used. The relevance of species delimitation within the genus *Limnospira* and the consequences to food industry authorization are discussed.

## Material and methods

### Database construction for reviewing *Limnospira* taxonomy

A database was constructed (Table S1) using different search engines, notably Google Scholar, PubMed, and NCBI. The keywords used were “*Limnospira*” or “*Arthrospira*” or “*Spirulina*” (for the oldest strains), “phylogeny” or “taxonomy” and “16S rRNA”, “internal spacer”, “*cpc*” or “genome”. The different entries from the databases were: the actual taxonomy of the strain (if the taxonomy of the strain changed, the most recent species name was used), the geographical location and the type of habitat of the sampling site (freshwater <5 psu, brackish water 5–35 psu, alkaline-saline water >35 psu, pH >8, alkaline-freshwater <5 psu with high alkalinity, marine, and soil), the accession numbers of the 16S

rRNA gene, the ITS, the *cpcBA*-IGS, and the whole sequenced genome. For each strain with at least one genetic sequence, a nucleotide BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) was performed to rapidly check the closest strain affiliation.

### Studied strains

A total of 92 *Limnospira* strains were obtained from the Paris Museum Collection: 72 new isolates of *Limnospira* were obtained from the Camargue area, a humid region located in the Rhône delta (South of France) mostly consisting of marshland and brine lagoons isolated as previously described (Cellamare et al. 2018), 13 from Lake Dziani Dzaha (Leboulanger et al. 2017, Cellamare et al. 2018), four from Lake Natron, two from a historical collection, and one strain provided by Spirulina Solutions (<https://spirulinasolutions.fr/>). Three additional strains were provided from other collections: *L. indica* PCC 8005 and *A. platensis* PCC 7345 (Pasteur Culture Collection of Cyanobacteria, France), and *L. fusiformis* SAG 85.79 (Culture Collection of Algae at Göttingen University, Germany). Those three strains are the reference strains of the species *L. indica*, *A. platensis*, and *L. fusiformis* that allow the identification of new isolated strains. Information about the strains is given in Table 1. In total, 95 strains were maintained as monospecific but non-axenic cultures, growing in 30 ml of Spirulina medium (Hamlaoui et al. 2022) at 24°C with a photon lux density of 30  $\mu\text{mol photon m}^{-2} \text{s}^{-1}$  under a 16:8-h light: dark photoperiod.

### Molecular and phylogenetic analyses

DNA extraction from *Limnospira* strains was carried out with a DNeasy® PowerLyser® PowerSoil® Kit (Cat N° 12855-100, Qiagen, Germany) following the manufacturer’s protocol. Mechanical lysis was carried out with a FastPrep-24™ 5 G BeadBeater (MP Biomedical, USA) for 5 cycles of 45 s ON, 45 s OFF at a frequency of six beats per second.

PCR amplification of the 16S rRNA gene was performed using the primers 8F (Lane 1991) and 920R (Gugger and Hoffmann 2004) with 35 cycles of hybridization at 58°C and elongation for 60 s; 861F (Gugger and Hoffmann 2004) and 1492R (Lane 1991) with 35 cycles of hybridization at 58°C and elongation for 60 s. Amplification of the ITS region was performed using the primers 322F and 340R (Ite-man et al. 2000) with 35 cycles of hybridization at 55°C and elongation for 75 s. Amplification of the *cpcBA*-IGS locus was performed using the primers *cpcB* and *cpcA* (Kumar et al. 2017) with 35 cycles of hybridization at 60°C and elongation for 45 s. All PCR reactions were carried out with the EconoTaq® Buffer with Mg and EconoTaq® DNA polymerase, using a Techne FTC-Plus/02 thermocycler (Bioer, China). The PCR products were separated by gel electrophoresis, extracted, and sequenced by Genoscreen (Lille, France). Sequences were assembled and corrected using MEGA-X software (Kumar et al. 2018) and aligned with MAFFT version 7 (<https://mafft.cbrc.jp/alignment/server/>) (Katoh et al. 2019).

For each strain, the 16S rRNA gene and ITS sequences overlapped and were assembled in a 1882–1884 base pair (bp) sequence. For the *cpcBA*-IGS locus, a sequence of 511 bp was obtained. Phylogenetic analyses were performed according to three methods: maximum likelihood (ML), neighbor joining (NJ), and maximum parsimony (MP), using the MEGA-X software with 1000 iterations. The best-fit substitution model was found using the MEGA-X software and the general time-

**Table 1.** List of cyanobacteria strains studied, their collection numbers, habitats, and sampling dates.

Sampling date	Sampling site	Localization	Habitat	Strains
Unknown	Unknown	Unknown	Unknown	<i>L. platensis</i> (anc. <i>A. platensis</i> ) PMC 288.06
Unknown	Unknown	Turkey?	Unknown	<i>L. platensis</i> (anc. <i>A. maxima</i> ) PMC 289.06
October 2010	Water column, Lake Dziani Dzaha (12°46'15"S; 45°17'17"E)	Mayotte, France	Alkaline-saline	<i>L. platensis</i> (anc. <i>L. fusiformis</i> ) PMC 737.11, 738.11
April 2014	Stromatolite, Lake Dziani Dzaha (12°46'15"S; 45°17'17"E)	Mayotte, France	Alkaline-saline	<i>L. platensis</i> (anc. <i>L. fusiformis</i> ) PMC 851.14, 852.14, 853.14
April 2014	Water column (0 m), Lake Dziani Dzaha (12°46'15"S; 45°17'17"E)	Mayotte, France	Alkaline-saline	<i>L. platensis</i> (anc. <i>L. fusiformis</i> ) PMC 894.15
April 2014	Water column (0,25 m), Lake Dziani Dzaha (12°46'15"S; 45°17'17"E)	Mayotte, France	Alkaline-saline	<i>L. platensis</i> (anc. <i>L. fusiformis</i> ) PMC 910.15, 912.15, 913.15, 914.15, 916.15, 917.15, 918.15
Unknown	Natron Lake	Tanzania	Alkaline-saline	<i>L. platensis</i> (anc. <i>Limnospira</i> sp.) PMC 1041.18, 1042.18, 1044.18, 1045.18
July 2019	Water column (43°27'39.79"N; 4°27'35.76"E), Site n°1	Saintes Maries de la Mer, France	Brackish	<i>L. platensis</i> (anc. <i>Limnospira</i> sp.) PMC 1223.20, 1224.20, 1225.20, 1228.20, 1229.20, 1231.20 to 1237.20, 1253.20 to 1256.20
July 2019	Water column (43°27'39.31"N; 4°27'35.37"E), Site n°2	Saintes Maries de la Mer, France	Brackish	<i>L. platensis</i> (anc. <i>Limnospira</i> sp.) PMC 1238.20, 1239.20, 1240.20, 1241.20, 1242.20, 1257.20
July 2019	Water column (43°27'38.96"N; 4°27'36.42"E), Site n°3	Saintes Maries de la Mer, France	Brackish	<i>L. platensis</i> (anc. <i>Limnospira</i> sp.) PMC 1226.20, 1227.20, 1230.20, 1243.20 to 1252.20, 1258.20 to 1261.20
September 2020	Water column (43°27'39.79"N; 4°27'35.76"E), Site n°1	Saintes Maries de la Mer, France	Brackish	<i>L. platensis</i> (anc. <i>Limnospira</i> sp.) PMC 1277.21 to 1282.21, 1307.21
September 2020	Water column (43°27'39.31"N; 4°27'35.37"E), Site n°2	Saintes Maries de la Mer, France	Brackish	<i>L. platensis</i> (anc. <i>Limnospira</i> sp.) PMC 1283.21 to 1297.21, 1308.21, 1309.21
September 2020	Water column (43°27'38.96"N; 4°27'36.42"E), Site n°3	Saintes Maries de la Mer, France	Brackish	<i>L. platensis</i> (anc. <i>Limnospira</i> sp.) PMC 1298.21, 1299.21, 1300.21, 1301.21, 1302.21, 1303.21, 1304.21, 1305.21, 1306.21
Unknown	Unknown	Peru	Freshwater	<i>L. platensis</i> (anc. <i>A. platensis</i> ) Paracas R14
Unknown	Unknown	Unknown	Unknown	<i>L. platensis</i> (anc. <i>L. indica</i> ) PCC 8005 (T)
Unknown	Unknown	Unknown	Unknown	<i>L. platensis</i> (anc. <i>A. platensis</i> ) PCC 7345 (T)
Unknown	Natron Lake	Tanzania	Alkaline-saline	<i>L. platensis</i> (anc. <i>L. fusiformis</i> ) SAG 85.79 (T)

PMC: Paris Museum Collection, PCC: Pasteur Culture Collection of Cyanobacteria, SAG: Culture Collection of Algae at Göttingen University, (T): Reference strains.

reversible model (GTR), with non-uniformity of evolutionary rates among sites using a gamma distribution (+G) and by assuming that a certain fraction of sites is evolutionarily invariable (+I) (Kumar et al. 2018). Pairwise distances and substitution rates were estimated using MEGA-X with the same substitution model (Kumar et al. 2018). For 16S rRNA gene phylogeny, an overall alignment ( $n = 305$  sequences,  $n = 248$  *Limnospira*,  $n = 57$  other genera) was generated, including the 92 new sequences from *Limnospira* isolates, the sequences available in GenBank belonging to genus *Arthrospira* ( $n = 18$ ), and to closely related genera from *Limnospira* according to Berthold et al. (2022): representatives of the family Sirenicapillariaceae, such as *Affixifilum* ( $n = 5$ ), *Capilliphycus* ( $n = 7$ ), *Limnoraphis* ( $n = 6$ ), *Neolyngbya* ( $n = 9$ ), *Sirenicapillaria* ( $n = 6$ ), and *Tigriniifilum* ( $n = 6$ ). The selected sequences were all longer than 900 bp. *Arthrospira* strains were considered an outgroup. A cut-off value of 94.5% of 16S rRNA gene sequence identity was used for genus definition (Yarza et al. 2014).

The 16S rRNA gene, ITS (all containing tRNA<sup>le</sup> and tRNA<sup>Ala</sup>) and *cpcBA*-IGS phylogenies were also performed with a total of 248, 274, and 249 strains, respectively. The sequences were taken from a previously constructed database, and 12, 5, and 1 entries, respectively, were removed due to unacceptable quality. A concatenated 16S-ITS phylogeny was performed containing 163 strains with known geographical

and habitat origins. For the 16S rRNA gene, ITS, and concatenated 16S-ITS phylogenies, the *Limnoraphis* sp. RMC B02 (MT764792) was used as an outgroup. For the *cpcBA*-IGS phylogeny, *L. etoshii* KR2008/49 (KC014069) was used as an outgroup. All trees were created using iTOL (<https://itol.embl.de/>) (Letunic and Bork 2021).

Alignments have been used to determine the presence of barcode gaps using the automatic barcode gap discovery (ABGD) online service (<https://bioinfo.mnhn.fr/abi/public/abgd/>) (Puillandre et al. 2012).

The generated ITS sequences were used for secondary structure analysis. The conserved regions (D1–D1' and Box-B helices) were analyzed using the Mfold WebServer (<http://www.mfold.org/mfold/applications/rna-folding-form.php>) (Zuker 2003) with default settings. The nucleotide sequences reported in this study have been deposited in the National Center for Biotechnology Information (NCBI) database. GenBank accession numbers OP526651-OP526742 (for concatenated 16S-ITS) and OP877184-OP877275 (for *cpcBA*-IGS) are listed in Table S1.

### Genomic analyses

Genomic analyses were made on the 50 available genomes of *Limnospira* (in July 2022) from 30 strains in culture and by using 20 genomes available from the NCBI. A total of 30 *Limnospira* strains out of the 92 have been chosen for

genome sequencing. The strain *A. platensis* PCC 7345 was chosen because it is the reference strain of this species. The other strains were chosen to have representatives of all localities and sampling dates: *A. platensis* Paracas R14 from Peru and *A. maxima* PMC 289.06 from unknown origin. We selected the strains with the highest apparent growth for other localities: Lake Natron, Lake Dziani Dzaha, and the Camargue region. For the genomes of *Limnospira* cultures, DNA was extracted from 10 mg of freeze-dried cells (Table S2) using a Dneasy® PowerLyser® PowerSoil® Kit (Cat N° 12855-100, Qiagen, Germany) as previously described. Total DNA was sequenced as 150-bp paired-end reads using an Illumina MiSeq platform (Eurofins, Konstanz, Germany). Reads were corrected and assembled using MetaSpades v3.12 (Bankevich et al. 2012). For each strain, the resulting scaffolds were binned using MetaBat2 (Kang et al. 2019) with default parameters, required coverages being obtained using Bowtie2 (Langmead and Salzberg 2012), and taxonomically assigned with CAT (von Meijenfeldt et al. 2019) to extract sequences belonging to *Limnospira*. Genome completeness was assessed using CheckM (Parks et al. 2015). All information about the newly sequenced genomes is provided in Table S2.

Genomic analyses of the 47 genomes were done on genome of isolated strains, excluding strains *Limnospira* sp. PLM2.Bin9 (Zorz et al. 2019) and SH-MAG-29 (Castro-Severyn et al. 2021), and with a completeness >98%, excluding the strain *Arthrospira* sp. O9.13F (Misztak et al. 2021) (completeness = 83.7%). Open reading frames in the 47 *Limnospira* genomes were predicted using Prokka (Seemann 2014), and protein amino acid sequences were used with Roary (Page et al. 2015) to define the core genome (sequence similarity threshold >90%). Multiple alignments of 2672 concatenated core genes provided by Roary were trimmed by BMGE (Criscuolo and Grimaldo 2010). The phylogenomic analyses were carried out with RaxML v8.2.12 (GTR-gamma model, 100 bootstraps) (Stamatakis et al. 2012) and with MEGA-X for NJ and MP (GTR + G + I model) methods with 100 iterations (Kumar et al. 2018). The ANI was calculated using the ANI calculator (<https://www.ezbiocloud.net/tools/ani>) (Yoon et al. 2017). All trees were visualized using iTOL (<https://itol.embl.de/>) (Letunic and Bork 2021).

A sub-sample of 24 *Limnospira* genomes has been done, representing the genomes available with the clades. This sub-sample included the 17 available genomes on NCBI, plus seven new sequenced genomes arbitrarily chosen to have one representative of each location and sampling date: *Limnospira* sp. PMC 1223.20 (Camargue sampling in 2019), PMC 1293.21 (Camargue sampling in 2020), PMC 1042.18 (Lake Natron), *L. fusiformis* PMC 737.11 (Lake Dziani Dzaha sampling 2010), *L. maxima* PMC 289.06 (unknown location), *A. platensis* PCC 7345 (reference strain), and Paracas R14 (Peru). Functional annotations were obtained using the web-tool, EggNOG-Mapper (<http://eggno-mapper.embl.de/>) with default parameters (Cantalapiedra et al. 2021). The presence or absence of coding DNA sequences (CDS) has been reported for each genome. The enrichment of CDS and clusters of orthologous genes (COG) categories have been performed with Fisher's exact tests. The COG category correspondences are available on NCBI (<https://www.ncbi.nlm.nih.gov/research/COG/>). The EggNOG Mapper category "Unc" refers to uncharacterized CDS.

The genome sequences reported in this study have been deposited in the NCBI database. GenBank accession numbers

are listed in Table S1: BioProject PRJNA902545, accession numbers JARPLC00000000-JARPMF00000000.

## Morphological and ultrastructure analyses

Morphological and ultrastructure analyses were carried out as described in Cellamare et al. (2018) on a selection of nine representative strains, considering the three reference strains (T) and representative strains from each phylogenetic clade and each sampling site: *L. indica* PCC 8005 (T), *A. platensis* PCC 7345 (T), Paracas R14, *L. fusiformis* SAG 85.79 (T), PMC 851.14, *Limnospira* sp. PMC 1042.18, PMC 1242.20, PMC 1293.21, and *L. maxima* PMC 289.06. Several morphological characteristics were investigated under a photonic microscope (Axio ImagerM2 Zeiss), including length and width of cells, length of trichomes, width of apical cells, presence of calyptra, number of spirals, pitch and diameter, orientation and variability of the helix, attenuation of the trichome, and presence of granules at cross-walls. The length of trichome spiral filaments, the apical attenuation of trichomes, and the variability of the helix were obtained using calculus: length of spiral filaments =  $((\text{helix pitch} + \text{helix diameter})/2) * \pi * \text{number of spirals}$ , attenuation of trichomes = mean of widths middle cells of the trichome/mean of widths of terminal cells of the trichome, variability of the helix = shortest helix pitch/larger helix pitch. Differences between strains were tested for significance using Kruskal–Wallis's test and pairwise Wilcoxon with Holm adjustment test, and Fisher's exact test.

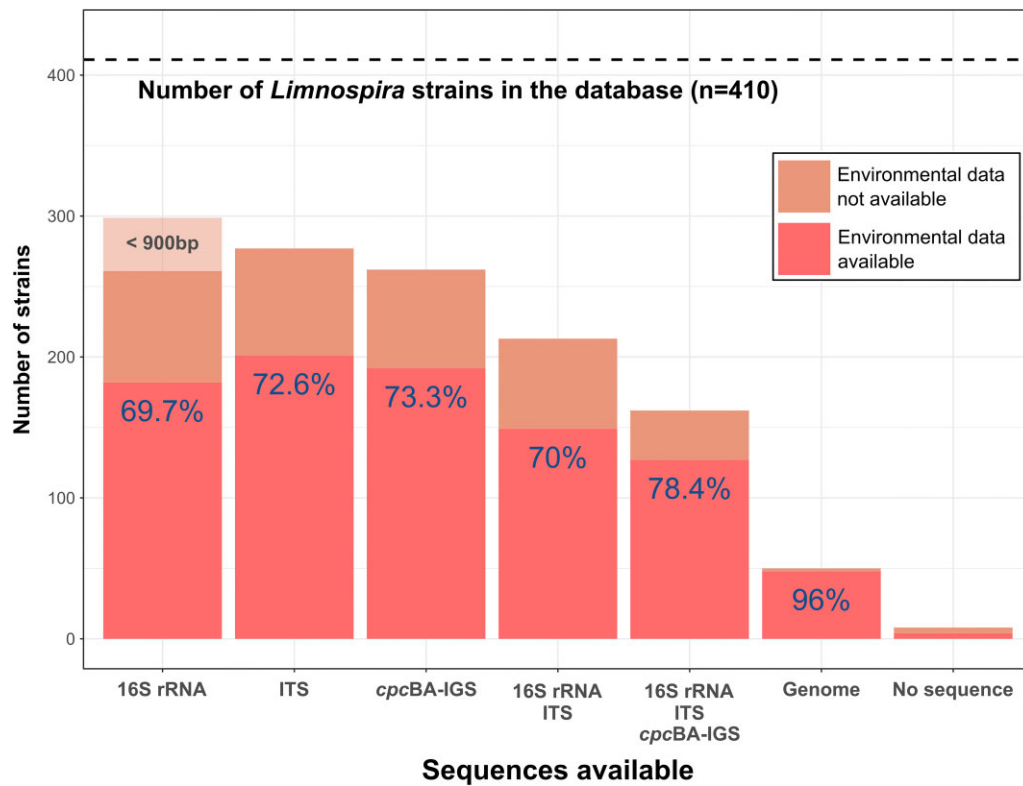
To study the disposition of the thylakoids and the presence or absence of aerotopes, strains were imaged by transmission electron microscopy (TEM) (Hitachi HT-7700, Japan) as previously described (Duval et al. 2021), with a slight modification of the fixation protocol. Filaments were harvested by centrifugation (1500 g, 5 min) from a growing culture and fixed with 2.5% (v/v) glutaraldehyde, 2% (v/v) formaldehyde, 0.18 M sucrose, and 0.1% picric acid in 0.1 M Sorensen's phosphate buffer (SPB; pH 7.4) for 1 h at room temperature.

## Results

### Review on *Limnospira* taxonomy

From 1994 to July 2022, 77 publications reported genomics data on isolated strains named *Arthrospira* or *Limnospira* (or *Spirulina* for the oldest strains). Including the sequences available on NCBI (October 2022) and those generated by the present study (including 16S rRNA, ITS, *cpcBA*-IGS, and whole sequenced genomes), a database of 494 sources ("strains") was used in this study (Table S1). Among the 494 strains, 410 (83%) were associated with the genus *Limnospira*. The others were *A. jenneri* ( $n = 20$ ), *Desertifilum* sp. ( $n = 3$ ), *Laspinema* sp. ( $n = 1$ ), *Planktothrix spiroides* ( $n = 3$ ), and *Tigrinihilum* sp. ( $n = 2$ ). Others were assigned to higher taxonomic levels (order or phylum) due to the absence or low quality of data ( $n = 36$ ). Eleven strains were not assigned to any known species, and eight had no genetic sequence because only their morphology was reported. Of the strains, 78.3% ( $n = 321$ ) had a reported geographic location, mostly Africa, America, Asia, and Europe, and 65.9% ( $n = 270$ ) had a reported type of habitat, including mostly freshwater, brackish water, saline–alkaline water, and alkaline freshwater.

Among the 410 *Limnospira* strains, 63.4% ( $n = 260$ ) had a 16S rRNA gene sequence >900 bp ( $n = 293$ , considering



**Figure 1.** Number of *Limnospira* strains of the database and studied sequences for phylogeny of 16S rRNA gene, ITS region, *cpcBA*-IGS region, and genome. Environmental information (geographical localization and habitat of origin) is reported. Percentages in blue represent the proportion of sequences associated to a strain with available environmental data.

all the sequences) (Fig. 1), 67.6% ( $n = 277$ ) had an ITS, and 61% ( $n = 250$ ) had a *cpcBA*-IGS sequence. Among the strains, 52.4% ( $n = 215$ ) had 16S rRNA gene and ITS sequences that had been concatenated for the phylogenetic studies, and 39% ( $n = 160$ ) of the *Limnospira* strains had three gene sequences and information about the localization and habitat. Among these, 12.2% ( $n = 50$ ) of the strains had an available genome (Fig. 1).

Among the 410 *Limnospira* strains, 25.4% ( $n = 104$ ) were isolated from Asia, 20.7% ( $n = 85$ ) from Africa, 20.2% ( $n = 83$ ) from Europe, and 11.7% ( $n = 48$ ) from America. Only one strain was isolated from Australia. The other 21.7% ( $n = 89$ ) strains had no information about their origin. Most of the strains were sampled from three types of habitats: 28% ( $n = 115$ ) from alkaline-saline water, 20.9% ( $n = 86$ ) from brackish water, 10.3% ( $n = 42$ ) from freshwater, and 1.5% ( $n = 6$ ) from alkaline-freshwater. Of these, 37.1% ( $n = 152$ ) had no information about the type of habitat from which they were sampled. A small number of strains ( $n = 9$ ) were isolated from other habitats: marine waters, geothermal ponds, and sediments.

### Phylogenetic and molecular analyses

The phylogenetic tree inferred from 305 16S rRNA gene sequences displayed a highly supported *Limnospira* clade (bootstrap, BS: 99), including the previously named *Arthrospira* ( $n = 146$ ), *Limnospira* ( $n = 99$ ), *Spirulina* ( $n = 2$ ), and *Phormidium* ( $n = 1$ ) strains (Fig. S1). The strains within this *Limnospira* clade shared >99.37% similarity of the 16S rRNA gene sequence and were distributed into three main clades with little bootstrap supports (BS: 62–82) I, II, and III

(Figs S2 and S3). Clade I contained 177 strains with >99.62% 16S rRNA gene similarity, including the reference strains of *L. fusiformis* SAG 85.79, *L. indica* PCC 8005, and *L. maxima* UTEX 2720, which are considered as living reference strains (Table 2). Clade II contained 57 strains sharing >99.86% 16S rRNA gene similarity and included the reference strain of *A. platensis*, PCC 7345. Clade III contained 14 *Limnospira* strains, sharing >99.85% 16S rRNA gene similarity. The ABGD method did not show any barcode gaps based on 16S rRNA gene sequences (Fig. S4). Among the 248 *Limnospira* strains with available 16S rRNA gene sequences, two (*Arthrospira* sp. Twintaung and Yekha) presented three different 16S rRNA gene sequences together sharing >99.73% gene similarity (Table 2). Within both strains, two different 16S rRNA copies supported relations to clade I and one copy to clade III.

The ITS-based phylogenetic tree revealed a pattern similar to the 16S rRNA tree, with three main phylogenetic clades I, II, and III (BS: 96, 95, 78) (Figs S5 and S6). The ITS sequence similarity of the 274 *Limnospira* strains was >89.3%. Within each clade, all strains shared >99% ITS sequence similarity (Table 3). The ITS similarity percentages were 89.9%, 93.3%, and 95.3% between clades I–II, I–III, and II–III, respectively. One strain (*A. platensis* Sp-11) was not affiliated with any clade and three strains (*Arthrospira* sp. Omsk O9.13F, Omsk O91.3H, and Yekha) presented two ITS copies. For each strain, the sequences were affiliated with clade I or clade III, sharing 93.57% ITS similarity. One strain (*Arthrospira* sp. Twintaung) had three ITS copies. Two of them were affiliated with clade I and the other with clade III, sharing 93.13–93.57% ITS similarity.

**Table 2.** Percent of similarity of the 16S rRNA gene sequences of representatives *Limnospira* strains of the 16S rRNA phylogenetic clades I (yellow), II (blue), and III (green).

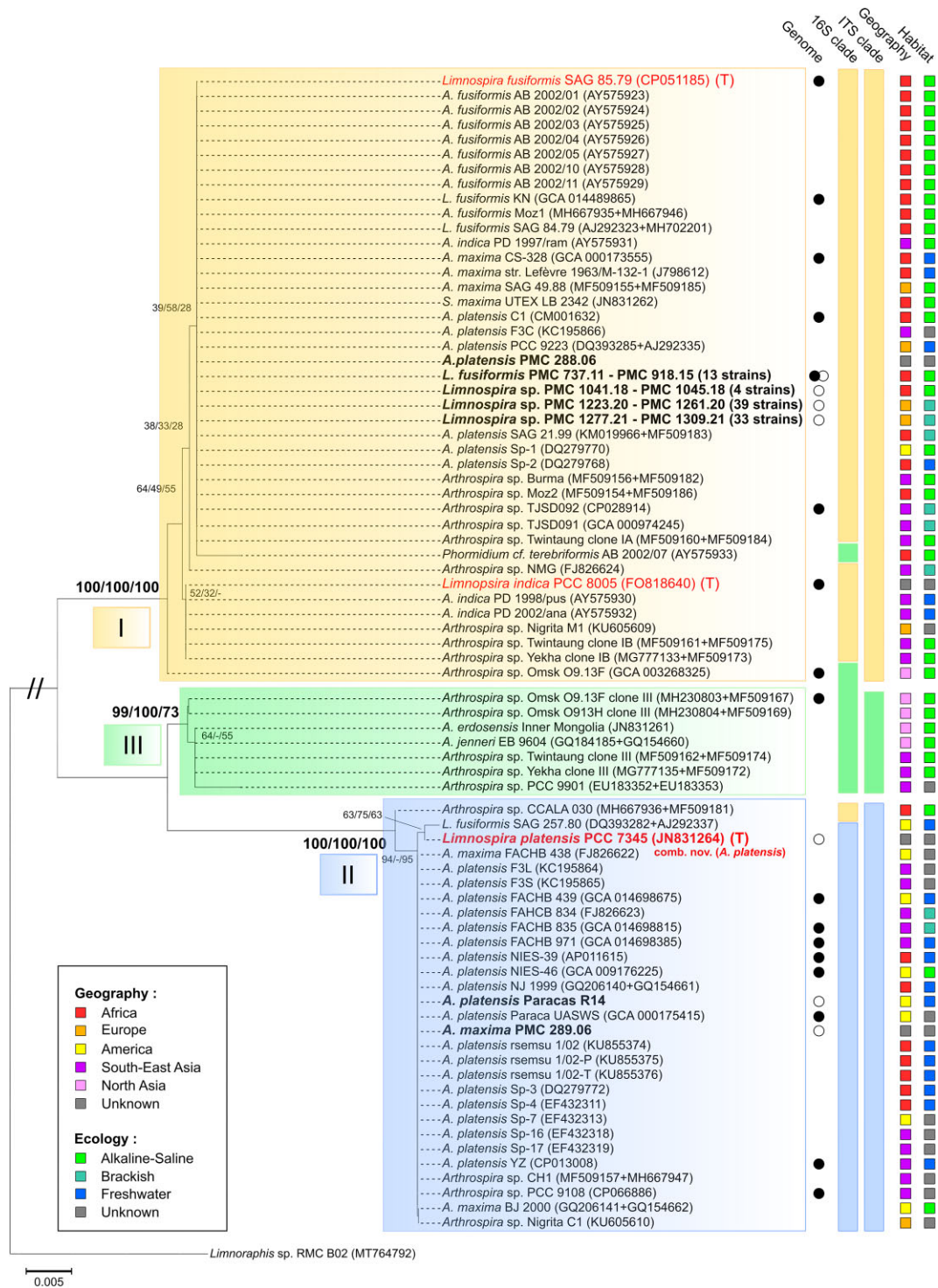
Clade	Strain	I									III				II				
		SAG 85.79 <sup>T</sup>	PCC 8005 <sup>T</sup>	UTEX 2720 <sup>T</sup>	Twintaung IA*	Twintaung IB*	Yekha IA*	Yekha IB*	Clone LAI-18	DICP-F	AUC	EB 9602	PCC 9901	Twintaung III*	Yekha III*	AB 2002/07	MV86	PCC 7345 <sup>T</sup>	IAM M-135
I	SAG 85.79 <sup>T</sup>	100																	
	PCC 8005 <sup>T</sup>	100	100																
	UTEX 2720 <sup>T</sup>	100	100	100															
	Twintaung IA*	100	100	100	100														
	Twintaung IB*	100	100	100	100	100													
	Yekha IA*	100	100	100	100	100	100												
	Yekha IB*	100	100	100	100	100	100	100											
	Clone LAI-18	99.9	99.9	99.9	99.9	99.9	99.9	99.9	100										
	DICP-F	99.85	99.85	99.84	99.9	99.9	99.9	99.9	99.79	100									
AUC	99.79	99.79	99.77	99.91	99.91	99.91	99.91	99.79	99.62	100									
III	EB 9602	99.77	99.77	99.77	99.81	99.81	99.81	99.81	99.68	99.62	99.7	100							
	PCC 9901	99.7	99.7	99.69	99.72	99.72	99.72	99.72	99.58	99.55	99.63	99.92	100						
	Twintaung III*	99.73	99.73	99.72	99.73	99.73	99.73	99.73	99.58	99.62	99.64	99.91	100	100					
	Yekha III*	99.73	99.73	99.72	99.73	99.73	99.73	99.73	99.58	99.62	99.64	99.91	100	100	100				
	AB 2002/07	99.64	99.64	99.61	99.64	99.64	99.64	99.64	99.48	99.47	99.57	99.85	99.93	99.91	99.91	100			
II	MV86	99.86	99.86	99.85	99.82	99.82	99.82	99.82	99.69	99.7	99.64	99.62	99.55	99.55	99.55	99.5	100		
	PCC 7345 <sup>T</sup>	99.79	99.79	99.77	99.73	99.73	99.73	99.73	99.58	99.62	99.57	99.55	99.48	99.47	99.47	99.57	99.93	100	
	IAM M-135	99.71	99.71	99.69	99.64	99.64	99.64	99.64	99.48	99.55	99.49	99.47	99.41	99.37	99.37	99.49	99.86	99.93	100

T: reference strains, \*: strains with multicopy. The 16S rRNA sequences of the strains PMC 737.11-917.15, PMC 1041.18-1045.18, PMC 1223.20-1261.21, PMC 1277.21-1309.21, and PMC 288.06 have 100% sequence similarity with strain SAG 85.79. The 16S rRNA sequences of the strains PMC 289.06 and Paracas R14 have 100% sequence similarity with strain PCC 7345.

**Table 3.** Percent of similarity of the 16S-23S ITS region sequences of representative *Limnospira* strains of the 16S-23S ITS phylogenetic clades I (yellow), II (blue), and III (green).

Clade	Strain	I										III					II				
		SAG 85.79 <sup>T</sup>	Twintaung IA*	AB 2002/07	Sp-14	NMG	PCC 8005 <sup>T</sup>	Omsk 09.13F IB*	Omsk 09.13H IB*	Twintaung IB*	Yekha IB*	Sp-11	Omsk 09.13F III*	Omsk 09.13H III*	Twintaung III*	Yekha III*	PCC 9901	KR 2004/12	PCC 7345 <sup>T</sup>	SAG 257.80	NIES 39
I	SAG 85.79 <sup>T</sup>	100																			
	Twintaung IA*	100	100																		
	AB 2002/07	99.79	99.78	100																	
	Sp-14	99.37	99.34	99.15	100																
	NMG	99.79	99.78	99.57	99.16	100															
	PCC 8005 <sup>T</sup>	99.58	99.56	99.36	98.95	99.79	100														
	Omsk 09.13F IB*	99.56	99.56	99.34	98.9	99.78	100	100													
	Omsk 09.13H IB*	99.56	99.56	99.34	98.9	99.78	100	100	100												
	Twintaung IB*	99.56	99.56	99.34	98.9	99.78	100	100	100	100											
	Yekha IB*	99.56	99.56	99.34	98.9	99.78	100	100	100	100	100										
III	Sp-11	97.91	97.81	97.66	97.28	98.12	98.33	98.25	98.25	98.25	100										
	Omsk 09.13F III*	93.13	93.13	92.9	92.46	93.35	93.57	93.57	93.57	93.57	95.34	100									
	Omsk 09.13H III*	93.13	93.13	92.9	92.46	93.35	93.57	93.57	93.57	93.57	95.34	100	100								
	Twintaung III*	93.13	93.13	92.9	92.46	93.35	93.57	93.57	93.57	93.57	95.34	100	100	100							
	Yekha III*	93.13	93.13	92.9	92.46	93.35	93.57	93.57	93.57	93.57	95.34	100	100	100	100						
	PCC 9901	92.55	92.24	92.26	91.91	92.77	92.98	92.68	92.68	92.68	94.68	99.12	99.12	99.12	99.12	100					
	II	KR 2004/12	89.85	89.8	89.58	89.62	90.07	90.29	90.24	90.24	90.24	92.05	95.36	95.36	95.36	95.36	94.95	100			
		PCC 7345 <sup>T</sup>	90.17	89.69	89.79	89.96	90.38	90.59	90.13	90.13	90.13	92.26	95.58	95.58	95.58	95.58	95.33	99.78	100		
		SAG 257.80	89.68	89.25	89.29	89.47	89.89	90.11	89.69	89.69	89.69	91.79	95.14	95.14	95.14	95.14	94.88	99.12	99.58	100	
		NIES 39	90.17	89.69	89.79	89.96	90.38	90.59	90.13	90.13	90.13	92.26	95.36	95.36	95.36	95.36	95.13	100	99.79	99.16	100

T: reference strains, \*: strains with multicopy. The 16S rRNA sequences of the strains PMC 737.11-917.15, PMC 1041.18-1045.18, PMC 1223.20-1261.21, PMC 1277.21-1309.21, and PMC 288.06 have 100% sequence similarity with strain SAG 85.79. The 16S rRNA sequences of the strains PMC 289.06 and Paracas R14 have 100% sequence similarity with strain NIES-39.



Downloaded from https://academic.oup.com/jambio/article/134/8/ixad159/7240153 by guest on 13 August 2023

**Figure 2.** Consensus phylogenetic tree based on 16S-ITS concatenated sequences of 163 *Limnospira* strains (*Limnorphis* sp. RMC B02 as outgroup). Strains newly sequenced in this study are shown in bold. Reference strains (T) are shown in red. Clades I, II, and III are, respectively, boxed in yellow, blue, and green. Bootstrap values are given following this order: ML/NJ/MP. Only support values >50 are shown. Available genomes are shown as black dots, and newly sequenced genomes are shown as white dots. The appearance of 16S rDNA gene and ITS region phylogenetic clades are shown as yellow (I), blue (II), and green (III) boxes. Geographical origins and habitats are highlighted in color squares.

The concatenated 16S-ITS phylogenetic tree showed a similar pattern to 16S rRNA and ITS phylogenies with bootstrap scores supporting the monophyly of the three clades at 100, 100, and 99, respectively (Fig. 2). All strains (except for four) had 16S rDNA gene and ITS sequences from the same clade affiliation. Two topologies were found: (1)

the 16S rDNA gene showed a relationship to clade III and ITS to clade I for the strains *Phormidium* cf. *terebriiformis* AB2002/07, *Arthrospira* sp. Omsk O9.13F and Omsk O913H; and (2) the 16S rDNA gene showed a relationship to clade I and ITS to sub-clade II for the strain *Arthrospira* sp. CCALA 030.



The phylogenetic tree based on the *cpcBA*-IGS sequences showed a slightly different pattern from the previous ones since the clade III (BS:86) containing 21 strains was further divided in two sub-clades (IIIa and IIIb) (Figs S7 and S8). The phylogenetic tree typology is detailed in supplementary Figs 7 and 8. Within each clade, the strains shared >99.5% *cpcBA*-IGS sequence similarity. Between each clade, the strains shared ~97.3% *cpcBA*-IGS sequence similarity (Table S3).

The secondary structure of the D1–D1' and Box-B helices showed some difference between strains. Three structural typologies of the D1–D1' helix were recovered within the genus *Limnospira* (Fig. S9). Two of them belonged to strains from clade I of the ITS phylogenetic tree, while the last one was shared by all the strains within clades II and III. Five structural typologies of Box-B have been recovered within the genus *Limnospira* (Fig. S10). Two of them were very similar and were shared by all strains of clade I of the ITS phylogenetic tree, and almost all strains of clade III. The three other structural typologies were very similar to each other. One was unique and belonged to the strain *Arthrospira* sp. PCC 9901 (clade III). The last two structural typologies belonged to all strains within clade II. The details about the D1–D1' and Box-B helices are depicted, respectively, in supplementary Figs 9 and 10.

Regardless of the phylogenetic analysis, the distribution of strains in the three clades I, II, and III appeared not to be specific to one geographic area (Africa, America, Asia, or Europe) or habitat (freshwater, brackish, or alkaline-saline waters) (Fig. 2).

## Genomic analyses

The genomes of 47 strains representative of the two clades I and II (based on the concatenated 16S-ITS phylogenetic tree, no genome was available for clade III) ranged from 5.5 to 6.79 Mbp with a  $44.6 \pm 0.2\%$  GC. A mean of  $4975 \pm 520$  CDS was detected, ranging from 4625 to 6673 CDS for *Limnospira* sp. PMC 1295.21 and *Arthrospira* sp. PCC 9108, respectively. The core-genome of the 47 *Limnospira* strains was composed of 2672 CDS, which represented  $45.8\% \pm 2.6\%$  of the genome length, and  $54.2\% \pm 4.9\%$  of the coding region of each strain.

The phylogenomic analysis based on the sequence of 2672 proteins belonging to the core-genome discriminated two main clades I and II (BS: 100, 100) (Fig. 3), which were also supported by ANI analysis (Fig. S11). This pattern was homogeneous with those observed in both 16S rRNA, ITS and concatenated 16S-ITS phylogenetic trees. Among the 47 sequenced *Limnospira* genomes, 36 belonged to clade I and 11 to clade II. Within each clade, strains shared >99.5% core-genome similarity and >98.2% of global ANI. Between each clade, strains shared  $95.3\% \pm 0.2\%$  core-genome similarity and  $93.4\% \pm 0.1\%$  of global ANI.

The distribution of strains in clades I, II, and III was not specific to one geographic area (Figs 3 and 4). Two strains originating from Lake Chad (*A. maxima* CS-328 and *A. platensis* NIES-39) belonged to different clades I and II (Fig. 4), showing individual diversity within the *Limnospira* population.

## Clustering of orthologous groups of proteins

The hierarchical clustering based on the presence or absence of CDS in the *Limnospira* genomes showed that CDS segregated into two clades, I and II, the same as those for the phy-

logenetic approach (Fig. S12). Among these, 956 and 1104 CDS were significantly (Fisher's exact test,  $P < .001$ ) more represented in clade I than clade II (Tables S4 and S5), with a mean of  $17.3\% \pm 0.8\%$  and  $19.7\% \pm 1.7\%$  of the coding genes of the genomes, respectively (Fig. S12). However, no specific COG function was significantly associated with clade I or II (Fisher's exact test,  $P > .01$ ), when 37.2% and 19.7% presented uncharacterized or unknown functions (COG categories Unc and S, respectively) (Fig. 5).

The 2672 CDS of the core-genome common to all *Limnospira* strains (Table S6) were assigned to seven COG categories, significantly highly represented (Fisher's exact test,  $P < .01$ ) compared to the flexible part of the genomes (Fig. 5). The functions of these COGs were assigned to primary metabolism (amino acids, carbohydrates, nucleotides, coenzymes, inorganic ion metabolisms, and energy production). In the flexible genome (Table S8), three COG categories were significantly over-represented (Fisher's exact test,  $P < .01$ ): those associated with DNA replication, recombination, and repair (L), defense mechanisms (V), and unknown functions (S).

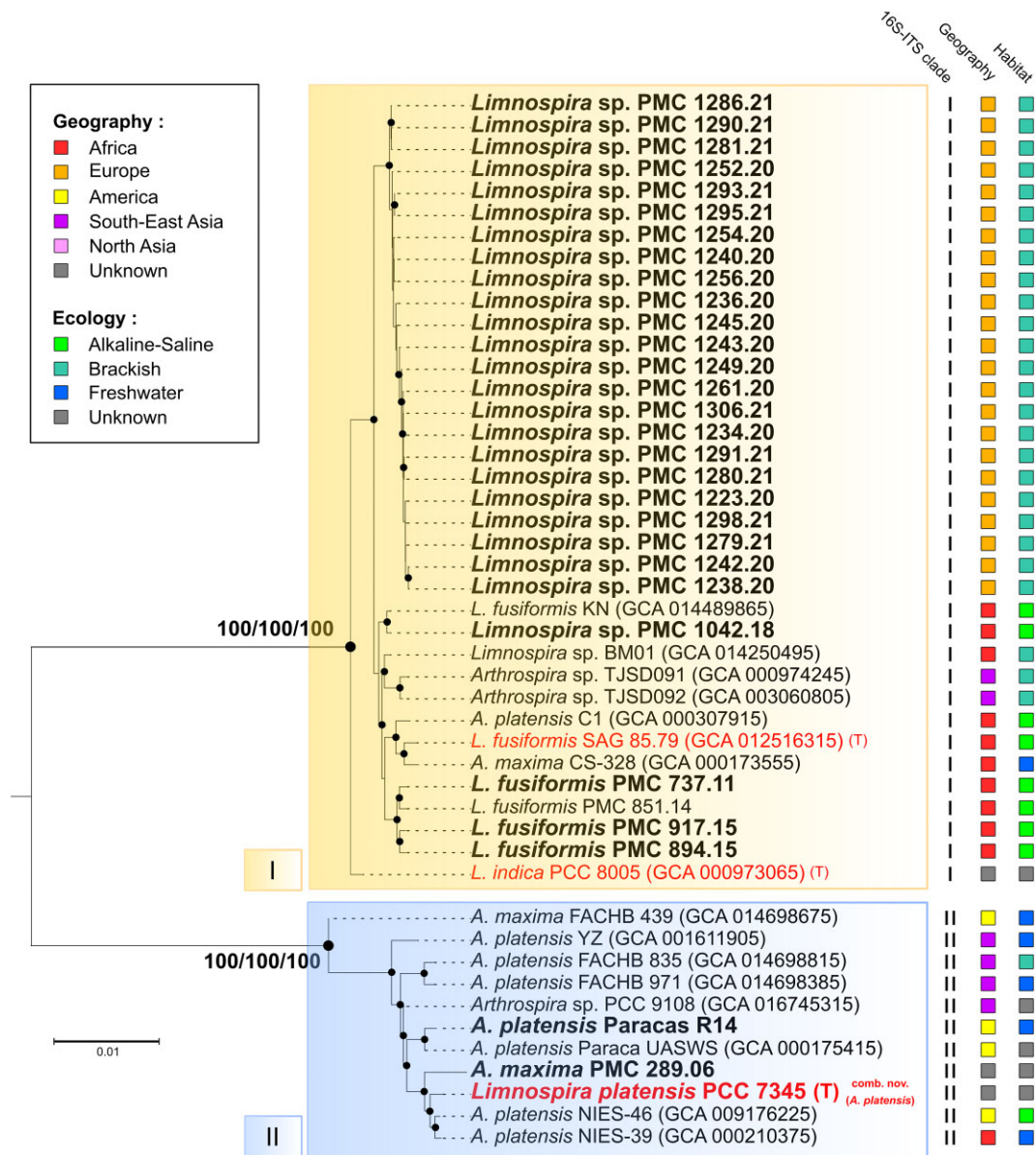
## Morphology and ultrastructure of representative *Limnospira*

Among the nine *Limnospira* strains for which morphological and ultrastructural analyses were done, seven showed coiled trichomes: *L. fusiformis* SAG 85.79, *Limnospira* sp. PMC 1042.18, PMC 1242.20, PMC 1293.21, *L. maxima* PMC 289.06, *A. platensis* PCC 7345, and Paracas R14. *Limnospira fusiformis* PMC 851.14 had straight trichomes and *L. indica* PCC 8005 had both coiled and straight trichomes in similar proportions, with shorter straight filaments (Table S8, Fig. S13). *Limnospira fusiformis* PMC 851.14 had very long filaments in culture (up to 4500  $\mu\text{m}$ ), which was between three and 13 times longer than all other strains. Cell dimensions were 4.7–10.7  $\mu\text{m}$  wide and 2.5–4.2  $\mu\text{m}$  long and showed some significant differences between strains (Fisher's exact test,  $P < .05$ ), with apical cells mostly presenting a calyptra. A slight trichome attenuation was observed, with no difference between strains (Fisher's exact test,  $P > .05$ ). Significant differences were observed between *Limnospira* strains regarding the helix dimensions (Fisher's exact test,  $P < .05$ ). Coiled trichomes had 3.9–7.1 helices per trichome, with diameters of 17.4–41.4  $\mu\text{m}$ , and a pitch of 40–76.6  $\mu\text{m}$  (Table S8). All strains showed granules at cross-walls. TEM observations showed that the nine studied strains had irregular thylakoid arrangements and aerotopes (Table S8, Fig. S13).

## Discussion

### *Limnospira* taxonomy, phylogeny, and species discrimination

Between 1892 and 2019, five species assigned to the genus *Limnospira* were described (Gomont 1893, Gardner 1917, Voronichin 1934, Desikachary and Jeeji Bai 1992, Tian et al. 2002). The genus affiliation changed three times: from *Arthrospira* to *Spirulina* (Geitler 1925), back to *Arthrospira* (Castenholz 1989), and now from *Arthrospira* to *Limnospira* (Nowicka-Krawczyk et al. 2019). The strains associated with the first described strain of *A. jenneri* Stizenberger and Gomont in 1892 (Gomont 1893) kept the genus name *Arthrospira*, and the strains presenting morphological (pres-

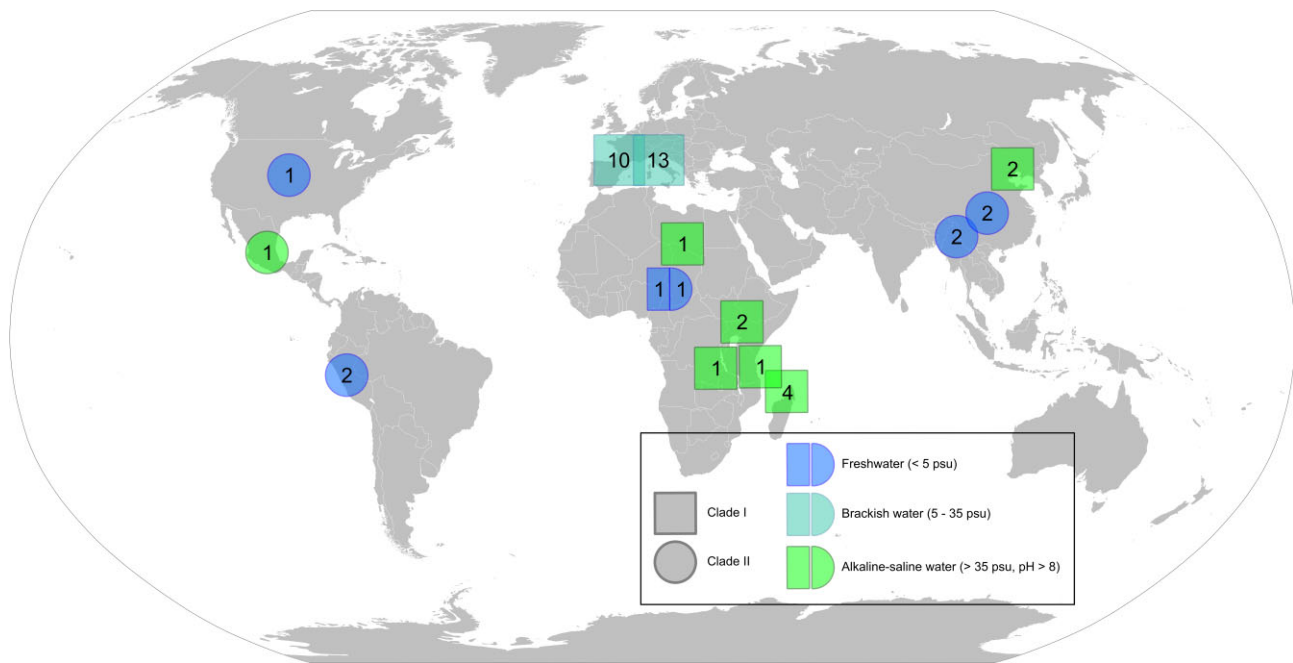


**Figure 3.** Consensus phylogenetic tree based on the 2672 most conserved genes of 47 *Limnospira* strains. Strains newly sequenced in this study are shown in bold. Reference strains (T) are shown in red. Clades I and II are respectively boxed in yellow and blue. Bootstrap values are given following this order: ML/NJ/MP. Only support values >90 are shown as black dots. The appurtenance of 16S-ITS concatenated phylogenetic clade is shown as I or II. Geographical origins and habitats are highlighted in color squares.

ence of calyptra), ultrastructural (irregular thylakoids), and molecular differences, and thus were assigned to the new genus, *Limnospira* (Nowicka-Krawczyk et al. 2019), including *L. fusiformis*, *L. maxima*, and *L. indica*. However, the authors did not assign the species *A. platensis* to the genus *Limnospira* because of its ecological characteristics. They were more like those of *A. jeneri*, benthic in freshwater habitat (Nowicka-Krawczyk et al. 2019). No biological or molecular material was available for the first-described *A. platensis* species (Gomont 1893). The strain *A. platensis* PCC 7345, which is part of the clade *Limnospira*, should be considered as the reference strain, as proposed earlier (Castenholz 1989).

Previous phylogenetic studies described clades I, II, and III and sub-clades IA, IB, IIA, and IIB within the *Limnospira/Arthrospira* genus based on mono-locus phylogenies (Scheldeman et al. 1999, Baurain et al. 2002, Comte et al.

2013, Papapanagiotou and Gkelis 2019). We confirmed the existence of these three clades based on the 16S-ITS concatenated phylogeny. However, the highest number of sequences used for our phylogenies showed that the different sub-clades were not supported. Based on a phylogenomic approach, we confirmed the presence of only two previously described clades (Misztak et al. 2021). The taxonomic discrimination of *Limnospira* species was based on these phylogenetic clades, the secondary structure of the ITS helices, the morphology and ultrastructure of the filaments, the geographical localization, and the respective habitat of the strains (Dadheech et al. 2010, Cellamare et al. 2018). Recently, however, based on three gene phylogenies and morphological characteristics, Papapanagiotou and Gkelis (2019) suggested that the genus *Limnospira* was monospecific, and the different phylogenetic clades corresponded to sub-species.



**Figure 4.** Geographical repartition of the *Limnospira* strains with sequenced genomes. Shape indicates to phylogenetic clade inferred from the phylogenomic analysis: square (I), circle (II). The color indicates the habitat: blue (freshwater), turquoise blue (brackish water), and green (alkaline-saline water). The number indicates the number of strains sampled in each site.

Our 16S rRNA gene analysis did not reveal a genuine barcode gap between genotypes, unlike the results of a previous study (Eckert et al. 2014). The *Limnospira* 16S rRNA inter-clade gene similarity was >99.3%. The threshold used for bacterial species was fixed at 98.65% (Kim et al. 2014) and those closest to *Limnospira* were 98.1%, 98.9%, and 99.3% for the genera *Capilliphycus*, *Sirenicapillaria*, and *Tigrinifilum*, respectively (Berthold et al. 2022). Based on the species diversity among the cyanobacterium *Desertifilum* (Oscillatoriales), an ITS similarity >97% could be considered as intra-specific diversity (González-Resendiz et al. 2019). If ITS similarity ranged from 93% to 97%, other criteria such as morphology, habitat, biogeography, and secondary structure of the ITS could be used for species delimitation (González-Resendiz et al. 2019), or description of a new species (e.g. in the Sirenicapillariaceae family) (Berthold et al. 2022) (*Limnospira*: ITS inter-clade similarity 90%–95.3%, no inter-clade difference of geographical repartition, habitats, or secondary structure of ITS).

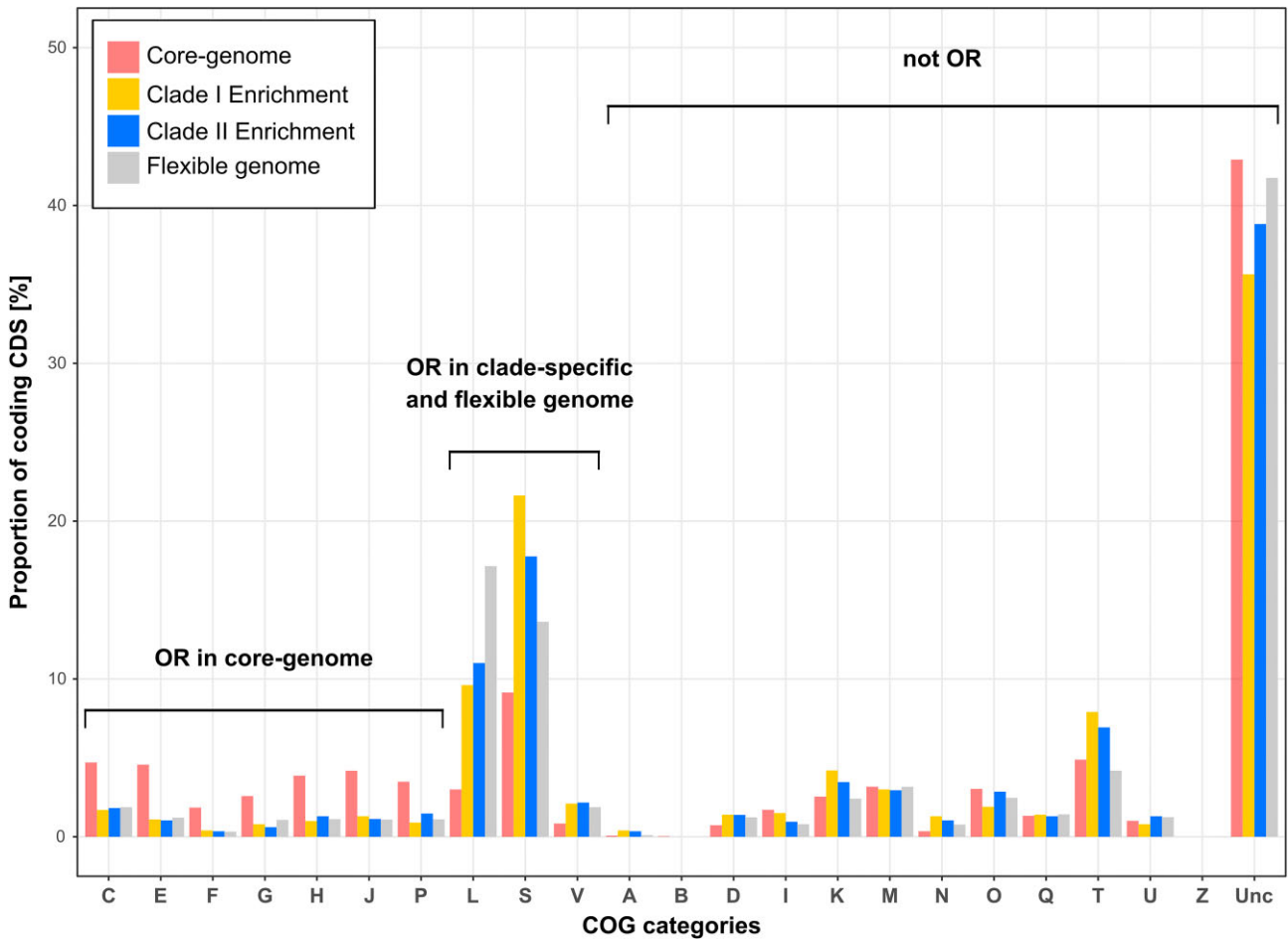
Previous investigators have proposed that the bacterial species delimitation based on ANI analysis should rely on two threshold values, >95%, which reflected intra-specific diversity, and <83%, which reflected inter-specific diversity (Jain et al. 2018), with a huge gap between those two discriminative values. ANI analysis is starting to be used more and more for species/genus delimitation of cyanobacteria. For example, within the genera *Microcystis* and *Raphidiopsis*, 15 and 2 species could be retrieved, respectively, based on a 96% ANI threshold (Willis and Woodhouse 2020, Willis et al. 2022). Within the genus *Microcoleus*, three species have been delimited with an inter-clade ANI between 86% and 92% (Bouma-Gregson et al. 2022). However, the clades revealed by ANI analyses were, in these different cases, related to the morphological differences for the genus *Microcystis*, and geographical separation for the genera *Raphidiopsis* and *Microcoleus* (*Limnospira*: inter-clade ANI = 93.4%).

Based on a study of the core-genome alignment similarity identity (CGASI) of 12 genera in the order Rickettsiales (Alphaproteobacteria), a primary threshold for bacterial species determination was proposed (Chung et al. 2018). For each genus, the core-genome comprised from 40% to 50% of each genome. Two organisms could be considered as belonging to the same species if the CGASI was around 96%. However, between two species, the CGASI ranged from 80% to 93%. As Rickettsiales and Cyanobacteria are considered bacteria, we used the same thresholds to predict the number of species within the genus *Limnospira* (*Limnospira*: CGASI > 95%).

Considering all the data from this study, the taxonomical approaches, and thresholds used for bacterial species delimitation, we confirmed that the genus *Limnospira* was monospecific, as proposed by Papapanagiotou and Gkelis (2019). These authors also suggested that the different phylogenetic clades corresponded to sub-species rather than to different *Limnospira* species. As the first described species, *A. platensis* (Gomont 1893) belongs to the *Limnospira* clade, and the monospecific genus *Limnospira* should be represented by *L. platensis*, with PCC 7345 renamed *L. platensis* as the reference strain. The description of *L. platensis* should be as follows: *L. platensis* (Gomont) comb. nov. Basionym *A. platensis* Gomont 1892; Order: Oscillatoriales (J.H. Schnaffer); Family: Sirenicapillariaceae (D.E. Berthold, Lefler, and Laughinghouse); Genus: *Limnospira* (Nowicka-Kraczyk, Mühlsteinova, and Hauer); and Reference strain: PCC 7345.

#### *Limnospira platensis*, plastic, and cosmopolite species

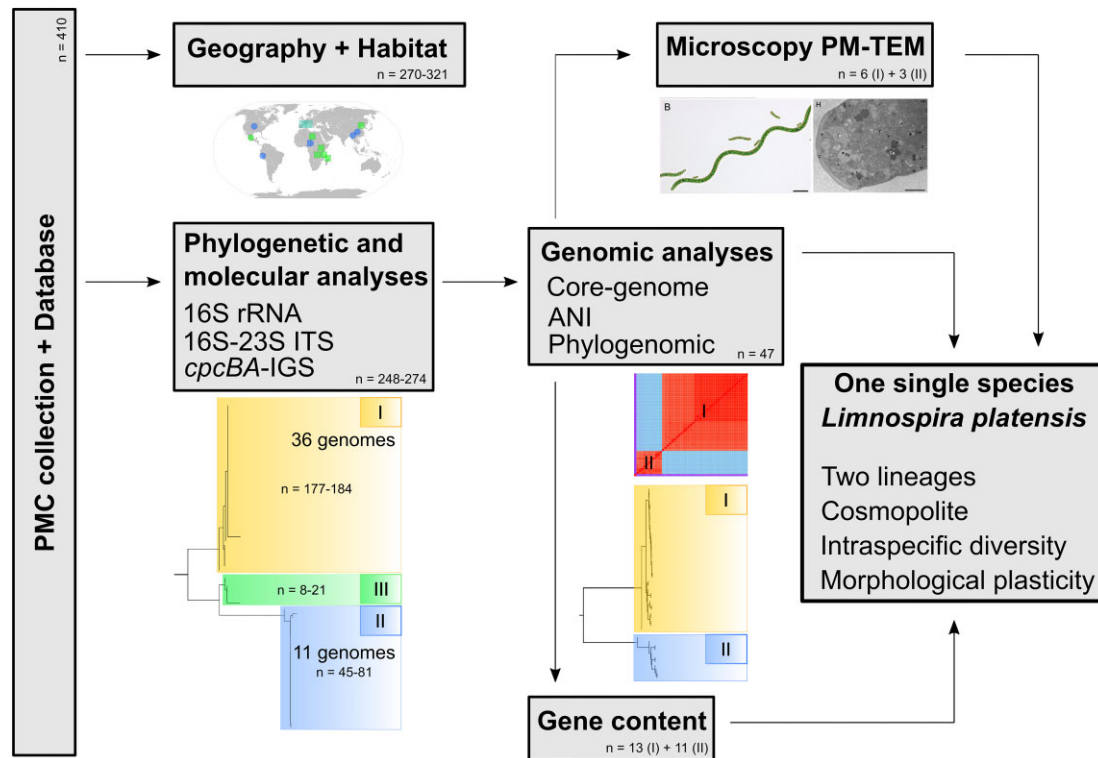
A huge infra-specific diversity exists within the species *L. platensis* (Baurain et al. 2002, Ballot et al. 2004, Dadheech et al. 2010), and comparison of genomic composition and gene function allowed us to describe this diversity. Regarding the mandatory house-keeping gene functions of *Limnospira* in the



**Figure 5.** Proportion of CDS of core-genome (black), clade I enrichment genome part (yellow), clade II enrichment genome part (blue), and whole flexible genome (gray) depending on the COG category. Brackets indicate where the COG category is over-represented in the core genome, the flexible genome, and clade-enrichment genome (Fisher’s exact test  $P < .01$ ), or not over-represented. OR = overrepresented. COG categories correspondence (available on NCBI) for those where differences are observed: C = Energy production and conversion/E = Amino acid transport and metabolism/F = Nucleotide transport and metabolism/G = Carbohydrate transport and metabolism/H = Coenzyme transport and metabolism/J = Translation, ribosomal structure, and biogenesis/P = Inorganic ion transport and metabolism/L = Replication, recombination, and repair of DNA/S = Function unknown/V = Defense mechanisms.

core-genome, most of them were associated with amino acid, carbohydrate, and nucleotide metabolism, coenzyme synthesis, inorganic ion regulation, and energy production. However, some COG categories, such as the prokaryotic defense system and transposases, appeared to be over-represented in the flexible *Limnospira* genomes belonging to clade I or clade II. Interestingly, such COG distributions on a flexible genome have also been described for *Aphanizomenon* (Halary et al. 2021). These COG distribution results are preliminary, however, because many cyanobacterial genes and COG functions remain unknown (Galperin et al. 2021). Among the enriched COGs with accurate annotation, ~70% were redundant between clades I and II and were not specific for a phylogenetic clade. Considering the plasticity among *Limnospira* taxa, this could also be explained by the high number of transposases and transposable elements. This phenomenon was previously described in a large number of freshwater cyanobacteria (Mazel et al. 1990) and proposed as a strategy for adapting to environmental changes. These genetic elements mediated the exchange of genetic materials and the generation of mutants allowing for rapid adaptation (Mazel et al. 1991, Lin et al. 2010).

The worldwide *L. platensis* distribution in many habitats, could be accounted for by dissemination through the air (Moustaka-Gouni 2011) and by specific animals (Curren and Leong 2020). It is well known that flamingos feed on *Limnospira* and live around numerous lakes where *Limnospira* have been described, such as Lake Bogoria or Lake Nakuru in Kenya (Ballot et al. 2004, Dadheech et al. 2010), Lake Natron in Tanzania (this study), Lake Chad in Chad (Carrieri et al. 2011), or the Camargue in France (this study). Flamingos could have played a significant role in the dispersion of *L. platensis* between reservoirs within Africa, Southern Europe, and Southeast Asia. Furthermore, we cannot exclude human populations as transport vectors. Both Aztec peoples and African tribes were known to have consumed *Limnospira* for centuries (Ciferri 1983). Humans could then have introduced *L. platensis* into new environments to cultivate and harvest it for their consumption, and thus have contributed to the geographic spread of the species. We noted that within the same lake (Lake Chad), both *L. platensis* genetic lineages from clade I and clade II coexist, constituting one more example of the genotypic diversity within *Limnospira*, at the intra-specific level. One wonders if this co-occurrence is restricted to this



**Figure 6.** Summary of the experimental approach and main results of the study. PMC: Paris Museum Collection, PM: Photonic microscopy, TEM: Transmission electron microscopy.

habitat or if it could be found in other habitats and other lakes. A large temporal sampling of ecological niches within a single lake could help to answer this question. The mechanism of speciation is still an evolving concept for bacteria, and even moreso for the cyanobacteria. Geographical location has been proposed to constitute the main vector of the clade separation within the *Raphidiopsis* genetic lineages (Gugger et al. 2005). The two *L. platensis* genetic lineages appear in at least one habitat (as Lake Chad), and the mechanism of speciation of these two genetic lineages may imply temporal or functional separation processes in distinct ecological niches. These two clades could be adapted to different ecological micro-niches or to different temporalities. However, up to now, no data have been available to test this hypothesis.

### Consequences for the food industry and some recommendations

In the European Union, the commercialization of strains of cyanobacteria and microalgae for human consumption is regulated by the European Commission, which regularly publishes the Novel Food Catalogue (NFC). The authorization to sell a species is established according to two criteria. If it can be proven that the species was consumed by the European population before 25 May 1997, the date of the first published NFC, then it is considered “food grade” and safe. Otherwise, it is mandatory to follow a lengthy procedure dictated by the European Commission to obtain novel food status. The NFC of authorized species includes genus modifications of the species *L. fusiformis*, *L. indica*, and *L. maxima* (European Commission Joint Research Centre 2021). However, only *A. platensis* can be sold in the food industry, although the extension of the authorization to the genus *Limnospira* has been proposed for

the next NFC. The authorization for the genus *Limnospira* (“Spirulina”) was based on its consumption since at least the 15th century in America and Africa (Ciferri 1983), but it is not clear how the commercial Spirulina food products correspond to the species *A. platensis* proposed in 1997 for food authorization. The criteria used for the identification and taxonomic affiliation of *A. platensis* were not specified in the NFC. Based on our study and others on *Arthrospira/Limnospira* taxonomy (Nowicka-Krawczyk et al. 2019, Papapanagiotou and Gkelis 2019), we demonstrated that within *Limnospira* only one species must be considered: *L. platensis*. Thus, we propose that all current and future strains of *Limnospira* should be authorized for human consumption. Verification of the correct genus for potential commercial use can be rapidly confirmed by the sequencing of the 16S rRNA gene.

As a conclusion and based on multiple approaches (Fig. 6), we demonstrated that the genus *Limnospira* was monospecific and represented by the species *L. platensis*. The genomic analyses used for the first time on *Limnospira* strains highlighted the large intra-specific diversity of this species and the subsequent physiological and metabolic plasticity, accounting for the worldwide distribution of the species and its adaptation to a variety of environments. Lastly, this study opens a discussion about the authorization of species for food use and the extension of the “Novel Food” status to more strains now grouped under the genus name *Limnospira* once all strains are renamed *L. platensis*.

### Acknowledgments

We would like to thank the UMR 7245 MCAM, Muséum National d’Histoire Naturelle, Paris, France for laboratories facilities and Algama for

funds. The authors thank the Pasteur Collection of Cyanobacteria, the Culture Collection of Algae at Göttingen University, and Gilles Planchon for sending strains, Églantine Soubrand for assisting in morphological and ultrastructural analyses, and Genoscreen and Eurofins for PCR products and genomes sequencing. The authors would like to thank the reviewers who helped to improve the manuscript.

## Supplementary data

Supplementary data is available at *JAMBIO Journal* online.

## Conflict of interest

The authors have no conflict of interest to declare.

## Funding

This research was funded by ANRT, through a convention Cifre n° 2020/058 awarded to the Company Algama employing T. Roussel for a PhD thesis and by the Museum National of Natural History. This work was also granted by the French National Research Agency (project DZIANI, ANR-13-BS06-0001 and project SUBSILAKE, ANR-21-CE02-0027-01).

## Author contributions

Théotime Roussel (Conceptualization, Data curation, Investigation, Methodology, Writing – original draft, Writing – review & editing), Sébastien Halary (Formal analysis, Methodology, Writing – review & editing), Charlotte Duval (Data curation, Formal analysis, Methodology, Validation), Bérénice Piquet (Data curation, Methodology, Validation), Jean-Paul Cadoret (Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review & editing), Léa Vernès (Conceptualization, Supervision, Validation, Writing – review & editing), Cécile Bernard (Conceptualization, Formal analysis, Funding acquisition, Investigation, Project administration, Supervision, Validation, Writing – original draft, Writing – review & editing), and Benjamin Marie (Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing)

## Data availability

The data underlying this article are available in the article and in its online supplementary material. GenBank accession numbers OP526651-OP526742 (for concatenated 16S-ITS) and OP877184-OP877275 (for *cpcBA*-IGS) are listed in Table S1.

## References

- Abdulqader G, Barsanti L, Tredici MR. Harvest of *Arthrospira platensis* from Lake Kossorom (Chad) and its household usage among the Kanembu. *J Appl Phycol* 2000;12:493–98. <https://doi.org/10.1023/A:1008177925799>.
- Andrade LM, Andrade CJ, Dias M *et al.* Chlorella and spirulina microalgae as sources of functional foods, nutraceuticals, and food supplements; an overview. *MOJ Food Process Tech* 2018;6:45–58. <https://doi.org/10.15406/mojfpt.2018.06.00144>.
- Avila-Leon I, Chuei Matsudo M, Sato S *et al.* *Arthrospira platensis* biomass with high protein content cultivated in continuous process using urea as nitrogen source: high protein biomass by continuous process with urea. *J Appl Microbiol* 2012;112:1086–94. <https://doi.org/10.1111/j.1365-2672.2012.05303.x>.
- Ballot A, Dadheech PK, Krienitz L. Phylogenetic relationship of *Arthrospira*, *phormidium* and *Spirulina* strains from Kenyan and Indian waterbodies. *Algol Stud* 2004;113:37–56. <https://doi.org/10.1127/1864-1318/2004/0113-0037>.
- Bankevich A, Nurk S, Antipov D *et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012;19:455–77. <https://doi.org/10.1089/cmb.2012.0021>.
- Baurain D, Renquin L, Grubisic S *et al.* Remarkable conservation of internally transcribed Spacer sequences of *Arthrospira* ('Spirulina') (Cyanophyceae, Cyanobacteria) strains from four continents and of recent and 30-year-old dried samples from Africa. *J Phycol* 2002;38:384–93. <https://doi.org/10.1046/j.1529-8817.2002.01010.x>.
- Berthold DE, Lefler FW, Laughinghouse HD. Recognizing novel cyanobacterial diversity in marine benthic mats, with the description of *Sirenicapillaria* fam. nov., two new genera, *Sirenicapillaria* gen. nov. and *Tigrinifilum* gen. nov., and seven new species. *Phycologia* 2022;61:146–65. <https://doi.org/10.1080/00318884.2021.2006589>.
- Bouma-Gregson K, Crits-Christoph A, Olm MR *et al.* *Microcoleus* (Cyanobacteria) form watershed-wide populations without strong gradients in population structure. *Mol Ecol* 2022;31:86–103. <https://doi.org/10.1111/mec.16208>.
- Cantalapiedra CP, Hernández-Plaza A, Letunic I *et al.* eggNOG-mapper V2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol Biol Evol* 2021;38:5825–9. <https://doi.org/10.1093/molbev/msab293>.
- Carrieri D, Ananyev G, Lenz O *et al.* Contribution of a sodium ion gradient to energy conservation during fermentation in the cyanobacterium *Arthrospira (Spirulina) maxima* CS-328. *Appl Environ Microbiol* 2011;77:7185–94. <https://doi.org/10.1128/AEM.00612-11>.
- Castenholz RW. Subsection III, order oscillatoriales. In: Stanley JT, Bryant MP, Pfennig N, Holt JG (eds), *Bergey's Manual Syst Bacteriol*. Baltimore: William and Wilkins, 1989, 1771–80.
- Castro-Severyn J, Pardo-Esté C, Mendez KN *et al.* Living in the High Extreme: Unraveling the Composition, Structure, and Functional Insights of Bacterial Communities Thriving in the Arsenic-Rich Salar de Huasco Altiplanic Ecosystem. *Am Soc Microbiol* 2021;9:e00444–21. <https://doi.org/10.1128/Spectrum.00444-21>
- Cellamare M, Duval C, Drelin Y *et al.* Characterization of phototrophic microorganisms and description of new cyanobacteria isolated from the saline-alkaline crater-lake Dziani Dzaha (Mayotte, Indian Ocean). *FEMS Microbiol Ecol* 2018;94:fiy108. <https://doi.org/10.1093/femsec/fiy108>.
- Chung M, Munro JB, Tettelin H *et al.* Using core genome alignments to assign bacterial species. *Am Soc Microbiol* 2018;3:21.
- Ciferri O. *Spirulina*, the edible microorganism. *Microbiol Rev* 1983;47:28. <https://doi.org/10.1128/mr.47.4.551-578.1983>.
- Comte K, Coursin T, Carre-Mlouka A. A new genotype in the genus *Arthrospira* (Oscillatoriales, Cyanobacteria) revealed by a mosaic-like structure of the 16S-23S rRNA intergenic spacer region in strain PCC 9901. *Phycologia* 2013;52:333–7. <https://doi.org/10.2216/12-063.1>.
- Criscuolo A, Gribaldo S. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol Biol* 2010;10:210. <https://doi.org/10.1186/1471-2148-10-210>.
- Curren E, Leong SCY. Natural and anthropogenic dispersal of cyanobacteria: a review. *Hydrobiol* 2020;847:2801–22. <https://doi.org/10.1007/s10750-020-04286-y>.
- Dadheech PK, Ballot A, Casper P *et al.* Phylogenetic relationship and divergence among planktonic strains of *Arthrospira* (Oscillatoriales, Cyanobacteria) of African, Asian and American origin deduced by 16S-23S ITS and phycocyanin operon sequences. *Phycologia* 2010;49:361–72. <https://doi.org/10.2216/09-71.1>.

- Demay J, Bernard C, Reinhardt A et al. Natural products from cyanobacteria: focus on beneficial activities. *Mar Drugs* 2019;17:320. <https://doi.org/10.3390/md17060320>.
- Desikachary TV, Jeeji Bai N. Taxonomic studies in Spirulina. In: Seshrathi CV, Jeeji Bai N (eds), *Spirulina, ETTA National Symposium*. Madras: MCRC, 1992. 12–21.
- Duval C, Hamlaoui S, Piquet B et al. Diversity of cyanobacteria from thermal muds (Balaruc-Les-Bains, France) with the description of *Pseudochroococcus coutei* gen. nov., sp. nov. *FEMS Microbes* 2021;2:xtab006. <https://doi.org/10.1093/femsmc/xtab006>.
- Eckert E, Fontaneto D, Coci M et al. Does a barcoding gap exist in prokaryotes? Evidences from species delimitation in cyanobacteria. *Life* 2014;5:50–64. <https://doi.org/10.3390/life5010050>.
- Edelmann M, Aalto S, Chamlagain B et al. Riboflavin, niacin, folate and vitamin B12 in commercial microalgae powders. *J Food Comp Anal* 2019;82:103226. <https://doi.org/10.1016/j.jfca.2019.05.009>.
- European Commission. Joint Research Centre. *Algae as Food and Food Supplements in Europe*. LU: Publications Office. 2021. <https://data.europa.eu/doi/10.2760/049515> (20 September 2022, date last accessed).
- Galperin MY, Wolf YI, Makarova KS et al. COG database update: focus on microbial diversity, model organisms, and widespread pathogens. *Nucleic Acids Res* 2021;49:D274–81. <https://doi.org/10.1093/nar/gkaa1018>.
- Gardner NL. *New Pacific Coast Marine Algae I*. 6, University of California Publications in Botany, 1917, 377–416.
- Geitler L. Cyanophyceae. In: Pascher A (ed.), *Süßwasserflora Deutschlands, Österreichs Und Der Schweiz*. Jena: Fischer, G. Verlag, 1925, 450.
- Gomont MA. *Monographie Des Oscillariées: (Nostacacées Homocystées)*. 716. In: Annales des Sciences Naturelles, Botanique, Série. Masson, 1893. 91–264.
- González-Resendiz L, Johansen JR, León-Tejera H et al. A bridge too far in naming species: a total evidence approach does not support recognition of four species in *Desertifilum* (Cyanobacteria). *J Phycol* 2019;55:898–911. <https://doi.org/10.1111/jpy.12867>.
- Gugger M, Molica R, Le Berre B et al. Genetic diversity of *Cylindrospermopsis* strains (Cyanobacteria) isolated from four continents. *Appl Environ Microbiol* 2005;71:1097–100. <https://doi.org/10.1128/AEM.71.2.1097-1100.2005>.
- Gugger MF, Hoffmann L. Polyphyly of true branching cyanobacteria (Stigonematales). *Int J Syst Evol Microbiol* 2004;54:349–57. <https://doi.org/10.1099/ijs.0.02744-0>.
- Halary S, Duperron S, Kim Tiam S et al. Intra-population genomic diversity of the bloom-forming cyanobacterium, *Aphanizomenon gracile*, at low spatial scale. *ISME Commun* 2023;3:1–4. <https://doi.org/10.21203/rs.3.rs-617160/v1>.
- Hamlaoui S, Yéprémian C, Duval C et al. The culture collection of cyanobacteria and microalgae at the French National Museum of Natural History: a century old but still alive and kicking! including in memoriam: professor Alain Couté. *Cryptogamie, Algologie* 2022;43:41–83. <https://doi.org/10.5252/cryptogamie-algologie2022v43a3>.
- Iteman I, Rippka R, de Marsac NT et al. Comparison of conserved structural and regulatory domains within divergent 16S rRNA-23S rRNA spacer sequences of cyanobacteria. *Microbiol* 2000;146:12. <https://doi.org/10.1099/00221287-146-6-1275>.
- Jain C, Rodriguez-R LM, Phillippy AM et al. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 2018;9:5114. <https://doi.org/10.1038/s41467-018-07641-9>.
- Kang DD, Li F, Kirton E et al. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* 2019;7:e7359. <https://doi.org/10.7717/peerj.7359>.
- Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform* 2019;20:1160–6. <https://doi.org/10.1093/bib/bbx108>.
- Kim M, Oh H-S, Park S-C et al. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int J Syst Evol Microbiol* 2014;64:346–51. <https://doi.org/10.1099/ijs.0.059774-0>.
- Komárek J. Quo vadis, taxonomy of cyanobacteria (2019). *Fottea* 2020;20:104–10. <https://doi.org/10.5507/fof.2019.020>.
- Kumar D, Dhar DW, Nehra P et al. Similarity analysis of *Spirulina Arthrospira* strains on the basis of phycocyanin operon locus (*cpcB-IGS-cpcA*) and 16S rRNA gene sequences. *Indian J Biotechnol* 2017;7:84–90.
- Kumar S, Stecher G, Li M et al. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* 2018;35:1547–9. <https://doi.org/10.1093/molbev/msy096>.
- Lane DJ. 16S/23S sequencing. In: Stackebrandt E., Goodfellow M (eds), *Nucleic Acid Tech Bacterial Syst*. Chichester; New York: John Wiley & Sons Ltd, 1991,115–175 .
- Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;9:357–9. <https://doi.org/10.1038/nmeth.1923>.
- Leboulanger C, Agogue H, Bernard C et al. Microbial diversity and cyanobacterial production in Dziani Dzaha Crater Lake, a unique tropical thalassohaline environment. *PLoS One* 2017;12:e0168879. <https://doi.org/10.1371/journal.pone.0168879>.
- Leticia I, Bork P. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* 2021;49:W293–6. <https://doi.org/10.1093/nar/gkab301>.
- Lin S, Haas S, Zemojtel T et al. Genome-wide comparison of cyanobacterial transposable elements, 2 potential genetic diversity indicators. *Nat Proc* 2010;33.
- Manen J-F, Falquet J. The *cpcB-cpcA* locus as a tool for the genetic characterization of the genus *Arthrospira* (cyanobacteria): evidence for horizontal transfer. *Int J Syst Evol Microbiol* 2002;52:861–7. <https://doi.org/10.1099/00207713-52-3-861>.
- Markou G. Effect of various colors of light-emitting diodes (LEDs) on the biomass composition of *Arthrospira platensis* cultivated in semi-continuous mode. *Appl Biochem Biotechnol* 2014;172:2758–68. <https://doi.org/10.1007/s12010-014-0727-3>.
- Mazel D, Bernard C, Schwarz R et al. Characterization of two insertion sequences, IS701 and IS702, from the cyanobacterium *Calothrix* species PCC 7601. *Mol Microbiol* 1991;5:2165–70. <https://doi.org/10.1111/j.1365-2958.1991.tb02146.x>.
- Mazel D, Houmard J, Castets AM et al. Highly repetitive DNA sequences in cyanobacterial genomes. *J Bacteriol* 1990;172:2755–61. <https://doi.org/10.1128/jb.172.5.2755-2761.1990>.
- Miklaszewska M, Waleron M, Morin N et al. Elucidation of the gas vesicule gene clusters in cyanobacteria of the genus *Arthrospira* (Oscillatoriales, Cyanophyta) and correlation with ITS phylogeny. *Eur J Phycol* 2012;47:233–44. <https://doi.org/10.1080/09670262.2012.692817>.
- Misztak AE, Waleron M, Furmaniak M et al. Comparative genomics and physiological investigation of a new *Arthrospira/Limnospira* strain O9.13F isolated from an Alkaline, Winter Freezing, Siberian Lake. *Cells* 2021;10:3411. <https://doi.org/10.3390/cells10123411>.
- Moustaka-Gouni M. Airborne algae and cyanobacteria occurrence and related health effects. *Front Biosc* 2011;E3:772–87. <https://doi.org/10.2741/e285>.
- Nabout JC, da Silva Rocha B, Carneiro FM et al. How many species of cyanobacteria are there? Using a discovery curve to predict the species number. *Biodiv Conserv* 2013;22:2907–18. <https://doi.org/10.1007/s10531-013-0561-x>.
- Nowicka-Krawczyk P, Mühlsteinová R, Hauer T. Detailed characterization of the arthrospira type species separating commercially grown taxa into the new genus *Limnospira* (Cyanobacteria). *Sci Rep* 2019;9:694. <https://doi.org/10.1038/s41598-018-36831-0>.
- Page AJ, Cummins CA, Hunt M et al. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 2015;31:3691–3. <https://doi.org/10.1093/bioinformatics/btv421>.
- Papapanagiotou G, Gkelis S. Taxonomic revision of commercially used *Arthrospira* (Cyanobacteria) strains: a polyphasic approach. *Eur J Phycol* 2019;54:595–608. <https://doi.org/10.1080/09670262.2019.1624832>.

- Parks DH, Imelfort M, Skennerton CT *et al.* CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 2015;25:1043–55. <https://doi.org/10.1101/gr.186072.114>.
- Puillandre N, Lambert A, Brouillet S *et al.* ABGD, automatic barcode gap discovery for primary species delimitation. *Mol Ecol* 2012;21:1864–77. <https://doi.org/10.1111/j.1365-294X.2011.05239.x>.
- Rasouli-Dogaheh S, Komárek J, Chatchawan T *et al.* *Thainema* gen. nov. (Leptolyngbyaceae, Synechococcales): a new genus of simple trichal cyanobacteria isolated from a solar saltern environment in Thailand. *PLoS One* 2022;17:e0261682. <https://doi.org/10.1371/journal.pone.0261682>.
- Scheldeman P, Baurain D, Bouhy R *et al.* *Arthrospira* ('Spirulina') strains from four continents are resolved into only two clusters, based on amplified ribosomal DNA restriction analysis of the internally transcribed spacer. *FEMS Microbiol Lett* 1999;172:213–22. <https://doi.org/10.1111/j.1574-6968.1999.tb13471.x>.
- Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014;30:2068–9. <https://doi.org/10.1093/bioinformatics/btu153>.
- Stamatakis A, Aberer AJ, Goll C *et al.* RAxML-Light: a tool for computing terabyte phylogenies. *Bioinformatics* 2012;28:2064–6. <https://doi.org/10.1093/bioinformatics/bts309>.
- Tian X, Qiao C, Gao L *et al.* Studies on the microstructure and ultrastructure of *Spirulina erdosensis*. *J Inner Mongolia Agric Univ* 2002;23:74–8.
- von Meijenfeldt FAB, Arkhipova K, Cambuy DD *et al.* Robust taxonomic classification of uncharted microbial sequences and bins with CAT and BAT. *Genome Biol* 2019;20:217. <https://doi.org/10.1186/s13059-019-1817-x>.
- Vonshak A. (ed.) *Spirulina Platensis (Arthrospira): Physiology, Cell-biology, and Biotechnology*. London ; Bristol, PA: Taylor & Francis, 1997.
- Voronichin NN. K biologii mineralizovanykh vodojemov Kulundinskoj stepi. *Rudy Soveta Izucheniu Prirodných Resursov, Ser Sibirsk* 1934;8:177–183.
- Walter JM, Coutinho FH, Dutilh BE *et al.* Ecogenomics and taxonomy of Cyanobacteria phylum. *Front Microbiol* 2017;8:2132. <https://doi.org/10.3389/fmicb.2017.02132>.
- Whitton BA, Potts M. *The Ecology of Cyanobacteria: Their Diversity in Time and Space*. Dordrecht, The Netherlands: Springer Science & Business Media, 2007.
- Willis A, Woodhouse JN, Neilan BA *et al.* Comparative genomics for understanding intraspecific diversity: a case study of the cyanobacterium raphidiopsis raciborskii. In *Advances in Phytoplankton Ecology*. Amsterdam, The Netherlands: Elsevier, 2022, 415–34. <https://doi.org/10.1016/B978-0-12-822861-6.00021-2>.
- Willis A, Woodhouse JN. Defining cyanobacterial species: diversity and description through genomics. *Crit Rev Plant Sci* 2020;39:101–24. <https://doi.org/10.1080/07352689.2020.1763541>.
- Yarza P, Yilmaz P, Pruesse E *et al.* Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol* 2014;12:635–45. <https://doi.org/10.1038/nrmicro3330>.
- Yoon S-H, Ha S, Lim J *et al.* A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek* 2017;110:1281–6. <https://doi.org/10.1007/s10482-017-0844-4>.
- Zorz JK, Sharp C, Kleiner M *et al.* A shared core microbiome in soda lakes separated by large distances. *Nat Commun* 2019;10:1–10. <https://doi.org/10.1038/s41467-019-12195-5>.
- Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 2003;31:3406–15. <https://doi.org/10.1093/nar/gkg595>.