**An overview of diversity, occurrence, genetics and toxin production
of bloom-forming *Dolichospermum* (*Anabaena*) species**

Li Xiaochuang^{1,2}, Dreher W. Theo³, Li Renhui^{1*}

1. Key Laboratory of Algal Biology, Institute of Hydrobiology, Chinese Academy
of Sciences, Wuhan, 430072, P. R. China
- ²University of Chinese Academy of Sciences, Beijing 100049, China
3. Department of Microbiology, and Center for Genome Research and
Biocomputing, Oregon State University, Corvallis, OR, USA

* corresponding author: reli@ihb.ac.cn

ABSTRACT

Dolichospermum, the new genus name for most of the planktonic former members of the genus *Anabaena*, is one of the most ubiquitous bloom-forming cyanobacterial genera. Its dominance and persistence have increased in recent years, due to eutrophication from anthropogenic activities and global climate change. *Dolichospermum* blooms, with their production of secondary metabolites that commonly include toxins, present a worldwide threat to environmental and public health. In this review, we summarize recent advances of the genus *Dolichospermum*, including taxonomy, genetics, bloom occurrence, and production of toxin and taste-and-odor compounds. The recent and continuing acquisition of genome sequences is ushering new methods for monitoring and understanding the factors regulating bloom dynamics.

Key words: *Dolichospermum*, diversity, taxonomy, occurrence, *Dolichospermum* blooms, toxins

Introduction

Among water bloom forming cyanobacteria, the *Anabaena/Aphanizomenon* group, which have now been revised into several genera, notably *Dolichospermum*, possess high morphological diversity (Fig. 1). The original genus *Anabaena* Bory ex Bornet et Flahault 1888, belonging to the family Nostocaceae by traditional classification (Komárek and Anagnostidis, 1989; Komárek, 2010) and to subsection IV according to bacteriological classification (Rippka et al., 1979; Castenholz, 1989), has members with uniseriate trichomes with the capability for cell differentiation into heterocysts and akinetes (Castenholz, 1989; Komárek and Anagnostidis, 1989). The presence of gas vesicles in cells leads to buoyancy regulation in the water column, allowing access to light and available nutrients in competition with other phytoplankton (Walsby et al., 1995). At present, the taxonomy of *Dolichospermum* and related cyanobacteria remains in a state of transition, and the new taxonomy has not been adopted by all authors in the literature. In this review, we use the genus designation *Dolichospermum* to refer to cyanobacteria that are described in the literature as either *Dolichospermum* or planktonic *Anabaena*; when referring to specific isolates, we retain the *Anabaena* nomenclature as used by authors.

Water bloom forming *Dolichospermum* species, are free-living in

1 lentic water bodies, including freshwater lakes, ponds and reservoirs as
2 well as brackish waters (Oliver and Ganf, 2000; Whitton and Potts, 2000).
3 Their wide occurrence is a demonstration of their ecological success
4 (Kristiansen, 1996), and they have dispersed all over the world by natural
5 means (via migrating fishes, birds, winds, and river basins) and by
6 human activities (ship ballast waters, aquaculture, and etc.) (O’Neil et al.,
7 2012). Under favorable conditions, *Dolichospermum* blooms form
8 potential threats to the environment and human health due to their
9 production of various kinds of potent toxins. The reported toxins
10 produced by these species can be categorized by chemical structure into
11 three different groups: cyclic peptides (microcystin), alkaloids
12 (anatoxin-a, anatoxin-a(S), saxitoxin, and cylindrospermopsin)
13 (Carmichael, 1986; Chorus and Bartram, 1999; Stüken et al., 2009;
14 Al-Tebrineh et al., 2010; Žegura et al., 2011; Singh and Dhar, 2013;
15 Akcaalan et al., 2014; Corbel et al., 2014; Sanchez et al., 2014) and
16 lipopolysaccharide. In recent decades, eutrophication in waters from
17 anthropogenic activities, combined with global warming, have promoted
18 *Dolichospermum* blooms to expand both in size and duration (Singh et al.,
19 2010; Singh et al., 2013; Koreivienė and Kasperovičienė, 2011; Huber et
20 al., 2012; O’Neil et al., 2012; Cirés et al., 2013; Slim et al., 2014).
21 Occurrences of *Dolichospermum* species and their toxins in large
22 quantities have been documented in numerous water bodies (e.g.,

1 Tsujimura and Okubo, 2003; Cook et al., 2004; John and Kemp, 2006;
 2 Karjalainen et al., 2007; Kravchuk, 2009; Cirés et al., 2013; Wang et al.,
 3 2013; Callieri et al., 2014), and the incidences of public exposure through
 4 drinking water or recreational activities have become more serious (e.g.,
 5 Bowling and Baker, 1996; Al-Tebrineh et al., 2012), making a review on
 6 *Dolichospermum* cyanobacteria timely. Options for controlling or
 7 mitigating cyanobacterial blooms and their toxins have been previously
 8 summarized (Jia et al., 2010; Ho et al., 2012; Shao et al., 2013; Yang et
 9 al., 2013; Li et al., 2014; Lürling and Tolman, 2014; Steffen et al., 2014;
 10 Zhang et al., 2014; Zamyadi et al., 2015) and will not be discussed here.
 11 In this review, we summarize knowledge of water bloom forming
 12 *Dolichospermum*, focussing on taxonomy, the status of genomic
 13 information that is beginning to facilitate studies on the detailed controls
 14 of bloom dynamics, and the potential hazards caused by these blooms
 15 and their toxins.

16

17 **1. Diversity, taxonomy and genetics**

18 **1.1 Morphology and phenotypic variability**

19 The traditional filamentous cyanobacterial genus *Anabaena* has been
 20 classified as belonging to Nostocaceae, containing both benthic and
 21 planktonic morphospecies (Geitler, 1932; Komárek, 2013) (Table 1). The
 22 planktonic *Anabaena* species are distinguished by the presence of gas

1 vesicles in cells that allow regulation of buoyancy and vertical migration
 2 in the water (Wacklin et al., 2009). There have been 549 *Anabaena*
 3 morphospecies (and infraspecific) described in the AlgaeBase as of 22
 4 April, 2015, of which 151 have been taxonomically accepted (Guiry and
 5 Guiry, 2015). Wacklin et al. (2009) proposed the new genus
 6 *Dolichospermum* to represent most of the planktonic species of
 7 *Anabaena* based on the genetic separation of benthic *Anabaena* species
 8 from the planktonic cluster containing gas vesicles in cells (although this
 9 genetic separation is not universal: e.g., Gugger et al., 2002a; Rajaniemi
 10 et al., 2005a), and Table 2 lists more than fifty planktonic *Anabaena*
 11 species with the name change. With the establishment of the
 12 *Dolichospermum* genus, the *Anabaena* genus represents only benthic
 13 species, with *Anabaena oscillarioides* Bornet ex Bornet et Flahault 1888
 14 as the type species. According to a recent proposal (Komárek et al., 2014)
 15 the genus *Dolichospermum*, together with planktonic genera such as
 16 *Aphanizomenon*, *Anabaenopsis* and *Cylindrospermopsis*, could be
 17 classified in a new family Aphanizomenonaceae, with the *Anabaena*
 18 genus remaining with benthic genera in the family Nostocaceae.
 19 However, there is at present a general deficit of the genomic sequence
 20 information that will form the basis of an ultimately stable taxonomy. The
 21 recent creation of the *Dolichospermum* genus may have solved one
 22 problem, but has left some taxonomic difficulties among planktonic

1 *Anabaena*-like cyanobacteria, notably intermixture of the
 2 *Dolichospermum* and *Aphanizomenon* genera, and the disposition of
 3 currently unassigned *Anabaena*-like clusters I, II and III (Komarek and
 4 Mares, 2012). A concerted effort to generate genomic sequences from
 5 blooms and cultures would greatly clarify the taxonomy and relationships
 6 among the *Dolichospermum*-like cyanobacteria.

7 **1.2 Phylogenetic intermixture between the planktonic** 8 ***Dolichospermum* and *Aphanizomenon***

9 The *Dolichospermum* species, described as planktonic species of
 10 *Anabaena* before or after 2009, were found to lack separation from the
 11 *Aphanizomenon* species in the phylogenies of 16S rRNA, *rpoB*, ITS1,
 12 *rpoc1* and *rbcLX* genes, supported by RFLP analyses of 16S rRNA
 13 genes (Lyra et al., 1997, 2001; Gugger et al., 2002a; Rajaniemi et al.,
 14 2005a; Halinen et al., 2007). Both genera *Dolichospermum* and
 15 *Aphanizomenon* are genetically heterogeneous and polyphyletic (Fig. 2)
 16 (Li et al., 2000, 2003; Rajaniemi et al., 2005a, b; Willame et al., 2006;).
 17 16S rRNA gene analysis showed that the similarity between some
 18 planktonic *Anabaena* and *Aphanizomenon* strains was higher than that
 19 observed among planktonic *Anabaena* species (Gugger et al., 2002a). It
 20 was suggested that the planktonic *Anabaena* (*Dolichospermum*) and
 21 *Aphanizomenon* isolates belonged to the same genus, since their 16S
 22 rDNA sequence similarity was higher than 97.5% (Gugger et al., 2002a),

1 above the limit of 95% 16S rDNA sequence similarity proposed as the
 2 definition of generic units (Stackebrandt and Goebel (1994)). On a local
 3 basis, phylogenetic analysis based on *cpcBA*-IGS sequences can
 4 distinguish *Dolichospermum flos-aquae* from *Aphanizomenon*
 5 *flos-aquae* Ralfs ex Bornet et Flahault 1888 and *Aphanizomenon gracile*
 6 (Lemmermann) Lemmermann 1907 that are present in the same
 7 environmental habitat (Wang et al., 2013). Examining more isolates of
 8 the genera *Dolichospermum* and *Aphanizomenon* from diverse
 9 geographical regions based on *cpc BA*-IGS and genome-wide
 10 comparisons (see later) is still needed to unravel the molecular overlap
 11 and divergence within these genera.

12 **1.3 Phylogenetic divergence within the genus *Dolichospermum***

13 Since the advent of the genus *Dolichospermum* in 2009, two new
 14 genera *Sphaerospermopsis* Zapomělová, Jezberová, Hrouzek, Hisem,
 15 Řeháková & Komárková 2010 and *Chrysosporum* Zapomělová,
 16 Skácelová, Pumann, Kopp & Janeček 2012 were further separated from
 17 *Dolichospermum* (Zapomělová et al., 2009, 2010a, 2012). Creation of
 18 these two most recent genera was supported by distinct autapomorphic
 19 features represented by the relatively spherical akinetes adjacent to one
 20 or both sides of heterocysts, and straight trichomes with narrowed ends,
 21 respectively, together with phylogeny based on 16S rRNA gene
 22 sequences (Table 1). Whether the remaining *Dolichospermum* members

1 will remain to be considered a monophyletic group, and precisely how
 2 they are phylogenetically positioned with respect to sister genera are
 3 questions that will be clarified by genome-wide analyses and recognition
 4 of past misidentifications. It has been estimated that about half of the
 5 cyanobacterial strains in culture collections have been misidentified
 6 (Komárek and Anagnostidis, 1989). Several studies have cautioned
 7 against the over-emphasis on morphological characters for identification
 8 and taxonomic classification (Beltran and Neilan, 2000; Rajaniemi et al.,
 9 2005a, b; Zapomělová et al., 2007; 2010b),

10 **1.4 Genotypes existing in *Dolichospermum* blooms in lakes**

11 A variety of genes have been used for studying the diversity and
 12 phylogenetic relationships among the *Dolichospermum* group and other
 13 cyanobacteria, including 16S rRNA, the internal transcribed spacer (ITS)
 14 between 16S and 23S rRNA genes, *rpoB* and *rpoC1* (RNA polymerase
 15 subunits), *rbcL* (ribulose-bisphosphate carboxylase-oxygenase) and
 16 *cpcBA*-IGS (phycocyanin intergenic spacer) (Lyra et al., 1997; Fergusson
 17 and Saint, 2000; Gugger et al., 2002a; Rajaniemi et al., 2005a; Willame
 18 et al., 2006; Halinen et al., 2008; Stüken et al., 2006; Ballot et al., 2010;
 19 Wang et al., 2013). ITS and *cpcBA*-IGS are the gene targets most
 20 commonly used for elucidating genotypic composition and dynamics in
 21 cyanobacterial communities, because they are more variable and
 22 therefore information-rich than 16S rRNA gene. Note, however, that ITS

1 amplifications for *Anabaena* typically produce more than one product of
2 different size due to multiple rRNA operons and tRNA gene insertions
3 between some 16S/23S rRNA gene pairs (e.g., *Anabaena* sp. 90).
4 Genotype compositions of natural populations have more widely been
5 studied during *Microcystis* blooms (Briand et al., 2009; Bozarth et al.,
6 2010; Xu et al., 2011; Zhu et al., 2012), and rarely studied in
7 *Dolichospermum* blooms (Liu et al., 2014). Miller and McMahon (2011)
8 studied the diversity and population ecology of bloom-forming
9 cyanobacteria in four lakes near Madison, Wisconsin, between May and
10 October 2008, by generating a clone library of *cpcBA*-IGS amplicons that
11 included many *Dolichospermum/Aphanizomenon* sequences among other
12 cyanobacteria. The gene sequences revealed a greater diversity than
13 indicated by morphological traits observable under the microscope, and
14 suggested that cyanobacterial diversity in these nearby lakes was greater
15 within lakes than between them. Liu et al. (2014) identified 118 different
16 *Dolichospermum* ITS genotypes, with two dominant genotypes (W-type
17 and L-type) existing during *Dolichospermum* blooms in Lake Taihu in
18 2009. The Wide-type dominated during both April and September, but a
19 succession from W- to L-type was observed in November, when much
20 less or no W-type was present. Correlation and stepwise regression
21 analyses showed that the succession of the two dominant genotypes was
22 significantly correlated with water temperature. These two dominant

Dolichospermum genotypes displayed striking seasonal pattern and temperature-associated features, implying that these are major ecotypes during *Dolichospermum* blooms in Lake Taihu.

In summary of the above taxonomic and phylogenetic aspects of *Dolichospermum*, there are large gaps in our knowledge of the genome sequences associated with bloom-forming lineages. Those genome sequences would allow progression from molecular phylogenies based on one (16S rRNA) or a few genes to genome-wide analyses. More studies focused on the genetic diversity of *Dolichospermum* species and populations in water blooms from various kinds of freshwater bodies (ponds, wiers, reservoirs, lakes, and etc.) will provide us a better understanding of the spatial and temporal dynamics of *Dolichospermum* populations during bloom seasons.

2 Genomics and other omics of the *Dolichospermum* genus

2.1 Genomics

Table 3 lists the finished or draft genomes that are publicly available. The *Dolichospermum*/*Anabaena*/*Aphanizomenon* group is represented currently by four draft and two finished genomes. Three finished and one draft genome represent cultures that are named *Anabaena* (PCC 7108, PCC 7120, PCC 7122, and ATCC 29413), but that have low phylogenetic affinity to the clade containing core *Dolichospermum*/*Anabaena* /*Aphanizomenon* isolates (Fig. 3). Unlike finished genomes, draft

1 genomes retain some gaps but typically represent a near-complete set of
2 unique genes. The emerging evidence is that, as observed for *Microcystis*
3 (Humbert et al., 2013), these genomes are replete with mobile and
4 repetitive elements (Wang et al., 2012). This means that genome
5 sequencing using the short 100 nucleotide reads provided by the popular
6 Illumina platform results in hundreds of contigs that cannot be joined
7 during read assembly; the situation is not improved when using the Ion
8 Torrent or Roche 454 sequencing platforms capable of read lengths up to
9 400-600 nucleotides (see *Aphanizomenon* NIES-81 and 2012/KM1/D3,
10 Table 3). Previously completed genomes required some manual finishing
11 with Sanger sequencing of PCR products to span remaining gaps,
12 following either cloning and Sanger sequencing or Illumina
13 high-throughput sequencing (Table 3). We have recently been successful
14 in obtaining a finished assembly of the 5.7 kbp *Anabaena* sp. WA102
15 genome using a single sequencing run on two Pacific Biosciences SMRT
16 cells (N. Brown and T. Dreher, in prep; Table 3). The average read
17 length was 8556 nucleotides, achieving about 50x average coverage and
18 an estimated accuracy of 1 error in 76,913 nucleotides. By contrast,
19 Illumina sequencing that generated 37 million reads left the genome
20 fragmented in about 700 contigs. The increasing accessibility of
21 sequencing platforms capable of generating reads long enough to span
22 repeat elements such as ribosomal RNA operons and transposons (1 to 3

1 kilobases) will undoubtedly result in accelerated completion of genome
2 sequences. Whole genome phylogenetics (phylogenomics), as used in Fig.
3 3, will ultimately provide the clearest picture of the relationships within
4 the *Dolichospermum/Anabaena/Aphanizomenon* group of cyanobacteria.

5 The genomes of members of the *Dolichospermum/Anabaena/*
6 *Aphanizomenon* group are medium-sized, with 4.4 - 5.9 Mbp encoding
7 3800-5000 proteins and 4 or 5 copies of the 16S-23S rRNA operons.
8 *Anabaena* sp. 90 is currently the only fully annotated genome (Wang et
9 al., 2012), and thus serves as reference for the group. The genome is
10 comprised of two circular chromosomes (4.33 and 0.82 Mbp) that encode
11 all essential genes and have an average G+C content of 38.1%. Three
12 circular plasmids (80, 56 and 20 kbp) encoding elements that are agents
13 of genome plasticity (integrases, transposes, recombinases and
14 phage-related elements) or hypothetical proteins are also present.
15 *Anabaena* sp. 90 is a microcystin producer, with the biosynthetic genes
16 *mcyA-J* present in a 56 kbp cluster of non-ribosomal peptide synthetase
17 (NRPS), polyketide synthase (PKS) and tailoring enzymes (see also
18 below). Three other NRPS gene clusters are present, supporting the
19 synthesis of anabaenopeptilide, anabaenopeptin and the antifungal
20 compound hassilidin, as are gene clusters for the generation of the
21 ribosomally synthesized peptide anacyclamide and 7 putative
22 bacteriocins. Some 0.25 Mbp (5% of the genome) is thus dedicated to

1 the production of bioactive secondary metabolites. The production of a
2 rich variety of such compounds is a characteristic of many cyanobacteria
3 (Calteau et al., 2014).

4 One of the goals of genome sequencing is to reveal the genes relating
5 to the metabolic and physiological capabilities that determine the
6 ecological niche preferred by each isolate. The *Anabaena* sp. 90 genome
7 contains the *nif* operon responsible for nitrogen fixation, a *gvp* operon for
8 gas vesicles that confer buoyancy, and a large complement of transporter
9 genes for nutrient uptake (Wang et al., 2012). Remnants of three putative
10 prophages suggest bygone infections, but no current lysogeny. A
11 plasmid-borne 14-kbp array of clustered regularly interspaced
12 palindromic repeats (CRISPR) presumably contains spacers derived from
13 previously invading DNA elements. A large number of genes (88) related
14 to another mechanism that protects against externally-derived DNA
15 elements, restriction-modification systems, reflects considerable
16 investment in protecting the cell against infection and in protecting
17 genome integrity in the face of potentially gene-disrupting invaders.
18 Ironically, there appears to be ample potential for self-inflicted genome
19 disruption, as suggested by the presence of 81 insertion sequence
20 transposons (413-2038 bp in length) and 147 instances of 77-541 bp-long
21 miniature inverted-repeat transposable elements (MITEs) (Wang et al.,
22 2012).

Preliminary analysis of the finished *Anabaena* sp. WA102 genome (N. Brown and T. Dreher, in prep.) indicates the existence of pervasive rearrangements around repetitive sequence elements that have scrambled the overall synteny (conserved gene order) of operons relative to the *Anabaena* sp. 90 genome. It will be interesting to learn how such rearrangements influence gene expression, and about the role of mobile genetic elements in the short-term (e.g., during bloom development) and long-term evolution within the *Dolichospermum* group.

2.2 Transcriptomics and proteomics

Niche adaptation and responses to various environmental changes or stimuli can be explored with genome-wide transcriptomic or proteomic studies that utilize genomic reference data for mRNA or protein identification. Global transcription profiles have been studied for *Nostoc* (*Anabaena*) PCC 7120 to study transcription start sites (Mitschke et al., 2011) and reprogramming during the transition to diazotrophic growth (Ehira et al., 2003; Flaherty et al., 2011). Investigating *Anabaena circinalis* (*Dolichospermum circinale*), Pomati et al. (2006) used a DNA microarray to compare mRNA expression levels for a limited set of genes between two pairs of saxitoxin-producing and non-producing strains. The most striking result was the higher mRNA levels for gas vesicle and phycocyanin genes in the toxic strains, suggesting a preference for positioning high in the water column at high light intensities. Only one

1 *Dolichospermum* study has to date utilized the powerful genome-wide
2 RNA-Seq analysis, assessing changes in mRNAs in *Anabaena* sp. 90 in
3 response to phosphate starvation (Teikari et al., 2015). Many genes were
4 differentially expressed, with increased mRNA levels noted for one of
5 two phosphate transporter operons and for genes putatively involved in P
6 mobilization from organic sources.

7 An alternative or complementary approach to studying global gene
8 expression is to quantitatively assess the proteome with advanced mass
9 spectrometry. D'Agostino et al. (2014) used the iTRAQ shotgun
10 proteomic technique to further probe the differences between
11 saxitoxin producing and non-producing strains of *Anabaena circinalis*,
12 (AWQC131C, and AWQC310F, respectively) when grown under
13 identical culture conditions. *A. circinalis* has been responsible for major
14 toxic bloom events in the Murray-Darling basin of Australia (Bowling et
15 al., 1996; Al-Tebrineh et al., 2012). Using the draft genome sequences
16 (Table 2) as reference, 20% of the predicted encoded proteins were
17 detected, including the enzymes for saxitoxin synthesis (in strain 131C),
18 which had been thought to be expressed at low levels. Numerous proteins
19 were shown to be expressed from hypothetical genes of unknown
20 function, validating these as functional genes. Proteins of the
21 photosynthetic and electron transport apparatus as well as the
22 carbon-fixing dark reactions were more highly expressed in the toxic

1 strain, as were proteins able to neutralize reactive oxygen species.
 2 These observations seem to provide an explanation for the higher growth
 3 rate of strain 131C. Differential expression in the nontoxic strain 310F
 4 showed higher expression of proteins that respond to general stress and
 5 light stress, suggesting that 310F is a low-light adapted strain that was
 6 under stress in the laboratory culture conditions (D'Agostino et al, 2014).
 7 This study showed that the two strains of *A. circinalis* are distinct
 8 ecotypes that prefer distinct environmental conditions. Such information
 9 is a major step towards building predictive models that can forecast
 10 blooms dominated by particular (e.g., toxic) cyanobacterial strains.

11

12 **3. Occurrence and distribution of *Dolichospermum* blooms**

13 The *Dolichospermum* (*Anabaena*) group represents diverse and
 14 ubiquitous freshwater cyanobacteria with worldwide distribution.
 15 Reynold et al. (2012) classified the diazotrophic *Dolichospermum* as
 16 tolerant of low nitrogen conditions but sensitive to mixing and low light,
 17 using buoyancy regulation to optimize opportunities for growth.

18 **3.1 Internal factors favouring *Dolichospermum* blooms**

19 A competitive advantage is provided by the ability of cells to adjust
 20 their vertical position in the water column via gas vesicles to obtain
 21 available nutrients and light. A further advantageous trait of
 22 *Dolichospermum* is the capacity for cell differentiation into heterocysts

1 and akinetes (Komárek and Anagnostidis, 1989; Komárek, 2010).
 2 Heterocysts are able to fix molecular nitrogen (Stewart, 1973), while
 3 akinetes sustain long-term dormancy and enable *Dolichospermum*
 4 populations to survive under harsh conditions (such as overwintering, dry
 5 seasons, flooding, and etc.) (Stüken, 2006).

6 It has been reported that cyanobacteria could maintain their
 7 dominance through allelochemical interactions, inhibiting the growth of
 8 other phytoplankton (Schlegel et al., 1999; Burja et al., 2001; Fujii et al.,
 9 2002; Suikkanen et al., 2004, 2005; Gorokhova and Engström-Öst, 2009;
 10 Engström-Öst et al., 2011). The cell-free filtrates of one *D.*
 11 *lemmermannii* strain could decrease cell numbers of *Rhodomonas* sp. and
 12 *Thalassiosira weissflogii* (Suikkanen et al., 2004). Previous studies
 13 indicated that toxins might be the compounds that were responsible for
 14 allelopathic properties (Rice, 1984; DeMott and Moxter, 1991), but the
 15 results of Suikkanen et al. (2004) indicated no correlation between toxin
 16 production and allelopathy, as the non-toxic *D. lemmermannii* strains
 17 could also cause allelopathic effects.

18 **3.2 External conditions promoting *Dolichospermum* blooms**

19 Despite the advantages of buoyancy and nitrogen fixation,
 20 *Dolichospermum* blooms principally form under conditions of
 21 eutrophication, with excess nitrogen (N) and phosphorus (P) inputs in
 22 freshwater bodies (Heiskanen and Leppänen, 1995; Anderson et al., 2002;

1 Paerl et al., 2011a). However, individual *Dolichospermum* blooms
2 occurring in different water bodies are associated with specific and
3 different external conditions promoting their blooms. The influences that
4 determine which species will dominate a bloom are at present
5 incompletely understood.

6 **3.2.1 Eutrophication associated with *Dolichospermum* blooms**

7 Phosphorus (P) has been traditionally considered as the principal
8 nutrient that spurs cyanobacterial bloom formation (Schindler, 1975).
9 However, various *Dolichospermum* species have been observed to
10 produce blooms under a wide range of nutrient (and temperature)
11 conditions. Rapala et al. (1997) found that growth rates of two
12 *Dolichospermum* (*Anabaena*) isolates increased with increasing P
13 concentrations, while no significant increases were shown with
14 increasing DIN (dissolved inorganic nitrogen, e.g., nitrate). Development
15 of blooms of *D. planktonicum* (*Anabaena planctonica*) in Lower Karori
16 Reservoir, New Zealand, a temperate eutrophic reservoir, was associated
17 with the establishment of stratification and increase in surface water
18 temperature. An annual peak in heterocyst frequency preceded the bloom
19 and occurred at a time of lower dissolved inorganic nitrogen
20 concentration (Wood et al 2010). A severe *D. spiroides* bloom was
21 observed in Erhai Lake, a plateau subtropical lake in southwest China
22 during June –August 1996 at water temperature (WT) of 22.2 °C, TN

1 (total nitrogen) of 0.559 mg.L^{-1} and TN/TP (TP, total phosphorous) of 20
 2 (Wang 2008). Similarly, another large *D. spiroides* bloom with 70,000
 3 cells mL^{-1} occurred in eutrophic Yanghe Reservoir, North China, during
 4 the summer of 2007 with WT 24.4°C , and a TN/TP of 16 (Li et al. 2010,
 5 Hui et al. 2015). *D. mendotae*, commonly found in meso to eutrophic
 6 lakes, was found to dominate in Lake Iznik, a deep lake of Turkey, on 22
 7 May 2013 with abundance of $4.9 \times 10^7 \text{ cell.L}^{-1}$ at WT of 20.6°C and TN of
 8 1.889 mg.L^{-1} (Akcaalan et al., 2014). *D. flos-aquae* was abundant in
 9 Taihu Lake, the third largest freshwater lake in China, during the period
 10 from late fall to spring, with the peak abundance of $4.0 \times 10^7 \text{ cpc}$ gene
 11 copies. L^{-1} (based on real-time PCR detection) in April 2009 at a time of
 12 TN $>4.0 \text{ mg L}^{-1}$ (Wang et al. 2013, Liu et al. 2014). A severe
 13 cyanobacterial bloom dominated by *D. circinalis* was present in the
 14 Barwon-Darling River, Australia, in November and December 1991 with
 15 concentrations of around $5.0 \times 10^8 \text{ cells L}^{-1}$ in many localities (Bowling
 16 and Baker 1996), and this bloom was triggered by the combination of
 17 many factors including lower flow in the river, higher temperatures
 18 between $25\text{-}30^{\circ}\text{C}$, high TP, high TN with low Nox-N and lower TN/TP
 19 values of 5-10. Transcriptomic and proteomic gene expression analyses
 20 of the type described earlier will in the future enable us to identify the
 21 genetic attributes that determine the distinct and varied niche adaptations
 22 exhibited by these various *Dolichospermum* species.

3.2.2 Climate change associated with *Dolichospermum* blooms

Recently, climate change dominated by global warming has been linked to the extension and intensification of cyanobacterial bloom events (Peeters et al., 2007; Ye et al., 2011; Huber et al., 2012; Paerl and Paul, 2012; Slim et al., 2014). A direct influence of global warming is rising water temperature. In 2008–2011, with temperatures exceeding 40°C in Lake Karaoun, dominance in the phytoplankton community of diatoms and green algae shifted towards cyanobacterial blooms (Slim et al., 2014). The cyanobacteria are better adapted to higher temperatures than other phytoplankton, such as diatoms and green algae (Robarts and Zohary, 1987; Paerl and Huisman, 2008). Climate change could also result in modified patterns, intensities, and duration of precipitation and drought (Paerl and Huisman, 2009). For example, extreme events, such as dry periods followed by heavy rains, have increased water level fluctuations in oligotrophic Lake Maggiore since 2005 (Callieri et al., 2014). A synchrony between the peaks of *D. lemmermannii* blooms and large water level fluctuations was found, with bloom formation attributed to nutrient releases from biofilms formed in the littoral zone when the lake water level was high, made available during a period of low water and dessication, followed by release when water levels rose again (Callieri et al., 2014). Global climate change is likely to induce salinity fluctuations in freshwaters due to an increase in drought frequency and

1 duration, or in precipitation. Unlike many eukaryotic phytoplankton that
2 are unable to tolerate changes in salinity, a number of cyanobacterial
3 species are capable of tolerating even very euryhaline conditions
4 (Laamanen et al., 2001; Orr et al., 2004). *Dolichospermum* is a major
5 contributor to blooms in brackish regions of the Baltic Sea (Halinen et al.,
6 2008); some *Dolichospermum* species can withstand high salinities up to
7 15 g/L NaCl (Moisander et al., 2002).

8 Cyanobacteria are known to benefit from the intensified and
9 prolonged thermal stratification in the water column that is associated
10 with increased temperatures due to global warming. Firstly, stratification
11 provides anoxic conditions in the hypolimnion that promote the release
12 of nutrients and iron and manganese from the sediments (Ismail et al.,
13 2002) that can be accessed by cyanobacteria through buoyancy control.
14 Secondly, high thermal stratification and low turbulence facilitate
15 buoyant cyanobacteria, enabling them to float to the water surface and
16 outcompete other algae for available light and nutrients. Westwood and
17 Ganf (2004) found that *D. circinale* (*Anabaena circinalis*) populations
18 grew faster when exposed to persistent stratification than under diurnally
19 stratified or mixed conditions in the lower Murray River, Australia. A
20 flow rate of only 0.05 m/s was observed to be sufficient to prevent the
21 persistent stratification needed for cyanobacterial growth (Mitrovic *et al.*,
22 2003). Several enclosure experiments showed that the growth of

1 *Dolichospermum* populations reduced with artificial mixing under a
 2 non-stratified regime (Reynolds et al., 1983; Nakano et al., 2001).

3 **3.2.3 Seasonal pattern and life cycle of *Dolichospermum* blooms**

4 From the combined effects of eutrophication and climate change,
 5 *Dolichospermum* blooms have experienced wider occurrence and
 6 distribution, prolonged duration and stronger intensity (Singh et al., 2010,
 7 2013; Koreivienė and Kasperovičienė, 2011; Huber et al., 2012; O’Neil
 8 et al., 2012; Cirés et al., 2013; Slim et al., 2014). *Dolichospermum*
 9 blooms commonly occur between late spring and middle autumn, in
 10 some systems followed by *Microcystis* blooms (Cook et al., 2004; Wang
 11 et al., 2013). Akinetes form and sink to the sediment when environmental
 12 conditions are favourable (Kaplan-Levy et al., 2010).

13 Akinete deposition and recruitment from sediment into the water
 14 column can be the key factors for bloom formation the following year
 15 (Preston et al., 1980; Hansson, 1996; Baker, 1999; Tsujimura and Okubo,
 16 2003; Kravchuk, 2009). A range of environmental factors have been
 17 reported to affect akinete formation, and these factors can be summarized
 18 as three main types: nutrients, light and temperature (Wolk 1965, Fay et
 19 al. 1984, van Dok and Hart 1996, Li et al. 1997, Moore et al. 2005,
 20 Thompson et al. 2009). Tsujimura and Okubo (2003) found that a high
 21 density of *Anabaena ucrainica* akinetes in the sediment increased the
 22 probability of bloom formation in the next year.

1 Akinete recruitment in the littoral zone is generally higher than that
2 from deep water and contributes more to facilitating the initiation of the
3 blooms; several factors could explain this phenomenon (Head et al., 1999;
4 Faithfull and Burns, 2006; Kravchuk, 2009). Firstly, the shallowness of
5 the littoral zone facilitates akinete germination, as the adequate light and
6 temperature requirements for germination are more readily met for
7 shallow rather than deep sediments (Ståhl-Delbanco and Hansson, 2002).
8 Secondly, wind-generated turbulence can easily bring the akinetes into
9 the water column in the shallow zone, thus making them accessible to
10 available light and nutrients (Rengefors et al., 2004). Thirdly, sediments
11 obtained from shallow waters can have higher nutrient concentrations
12 than those sourced from deep water, thereby promoting cyanobacterial
13 growth and recruitment. The higher nutrient inputs can be from farms in
14 the catchment, or in the form of organic nutrients resulting from
15 microbial decomposition of plants, such as *Typha* beds, that can be
16 located in the shallow water (Wetzel, 2001).

17 However, bloom recruitment is unlikely to follow the same pattern in
18 all systems. Thus, Head et al. (1999) indicated that, in the Esthwaite
19 Water, northern England, the major source of inocula for *D. flos-aquae*
20 blooms in summer were overwintering populations in the water column.
21 In the Murray River, Australia, Baker (1999) found that a small
22 proportion of filaments of *D. circinalis* in the winter water column

1 coexisted with a large number of akinetes in the sediment.

2

3 **4 Cyanotoxins from *Dolichospermum* species and blooms**

4 The heightened recent intensity and duration of *Dolichospermum*
 5 blooms (Singh et al., 2010, 2013; Koreivienė and Kasperovičienė, 2011;
 6 Huber et al., 2012; O’Neil et al., 2012; Cirés et al., 2013; Slim et al.,
 7 2014) **have** been accompanied by surface scums, reduced water
 8 transparency, anoxic conditions due to bacterial decomposition of the
 9 bloom biomass, and reduced diversity of natural phytoplankton (O’Neil
 10 et al., 2012). These effects have increased the financial burden of water
 11 treatment when associated with drinking water sources (Lee et al., 2000).
 12 *Dolichospermum* species are able to produce various kinds of toxic
 13 secondary metabolites, notably toxins; about 25% to 75% of
 14 cyanobacterial blooms are toxic in natural waters (Chorus and Bartram,
 15 1999). The toxins can be classified into three different categories by
 16 chemical structure: cyclic peptides (microcystin), alkaloids (anatoxin-a,
 17 anatoxin-a (S), saxitoxin, and cylindrospermopsin) and
 18 lipopolysaccharides (LPS); and four groups according to the target
 19 organs: neurotoxins (anatoxin-a, anatoxin-a (S) and saxitoxin),
 20 hepatotoxins (microcystin), cytotoxins (cylindrospermopsin), and
 21 dermatotoxins (LPS) (Krishnamurthy et al., 1986; Chorus and Bartram,
 22 1999; Stüken et al., 2009; Al-Tebrineh et al., 2010; Žegura et al., 2011;

1 Singh and Dhar, 2013; Akcaalan et al., 2014; Corbel et al., 2014; Sanchez
 2 et al., 2014). Toxins are synthesized at all stages of cyanobacterial
 3 growth and remain mostly in the cell until their release into waters after
 4 cell lysis (Sivonen and Jones, 1999). The toxins produced by
 5 *Dolichospermum* blooms represent a potential threat to environmental
 6 and human health (Wiegand and Pflugmacher, 2005 Torokne et al., 2011;
 7 Žegura et al., 2011; Singh and Dhar, 2013; Corbel et al., 2014; Freitas et
 8 al., 2014; Smutná et al., 2014; Weirich and Miller, 2014). Cyanotoxins
 9 can be bioaccumulated by aquatic invertebrates and have been found in
 10 plants and aquatic vertebrates (Codd et al., 1999; Pflugmacher et al.,
 11 2001; Chen et al., 2004; Yin et al., 2005; Ibelings and Chorus, 2007;
 12 Saqrane et al., 2007; Ettoumi et al., 2011). This can pose a potential
 13 health risk to animals and humans through the food chain, with children
 14 more susceptible to toxins because of their lower body weight, behavior,
 15 and toxic effects on development (Weirich and Miller, 2014).

16 **4.1 Nonribosomal peptide toxin**

17 **4.1.1 Microcystin**

18 Microcystins are the most ubiquitous toxins detected in cyanobacterial
 19 blooms (Singh and Dhar, 2013), and have been detected in a number of
 20 *Dolichospermum* species, including *D. flos-aquae*, *D. lemmermannii* and
 21 *D. circinale* (Krishnamurthy et al. 1986, Sivonen et al., 1990, 1992). The
 22 cyclic heptapeptides made by non-ribosomal peptide synthetases and

1 associated trimming enzymes exist in over 80 congeners with varying
 2 toxicity. The gene cluster responsible for microcystin synthesis (genes
 3 *mcyA-J* spanning 56 kbp) has been characterized in *Dolichospermum*
 4 (*Anabaena*) sp. 90 (Fig. 4), facilitating assays for distinguishing
 5 microcystin producing and non-producing strains (Rouhiainen et al.,
 6 2004). Vaitomaa et al. (2003) have described an *mcyE*-specific QPCR
 7 assay capable of distinguishing *mcy* genes from *Microcystis* and
 8 *Anabaena*.

9 **4.2 Alkaloid toxins**

10 **4.2.1 Cyindrospermopsin**

11 Cyindrospermopsin, an alkaloid cytotoxin that possesses a tricyclic
 12 guanidine moiety combined with hydroxymethyluracil, has a molecular
 13 weight of 415 (Falconer, 1999; Wiegand and Pflugmacher, 2005). CYN
 14 was chemically characterized from the species *Cylindrospermopsis*
 15 *raciborskii* (Svrcek and Smith, 2004), and two derivatives,
 16 7-epi-cylindrospermopsin and deoxy-cylindrospermopsin, have been
 17 identified (Li et al., 2001). Cyindrospermopsin has been detected in
 18 *Anabaena bergii* (now *Chrysosporum bergii*) (Schembri et al., 2001) and
 19 *Anabaena lapponica* (Spoof et al. 2006). The gene cluster supporting
 20 cylindrospermopsin synthesis (genes *aoaA-C* and *cyrD-JK* spanning 42
 21 kbp) has been sequenced from *Aphanizomenon* sp. 10E6 (Stüken and
 22 Jakobsen, 2010) (Fig. 4).

4.2.2 Saxitoxin

The saxitoxins are best known as paralytic shellfish poisons (PSPs) associated with their bioaccumulation in shellfish such as *Saxidomus giganteus*, from which the first identification was made (Wiegand and Pflugmacher, 2005). Saxitoxins are tricyclic purine derivatives (Wiegand and Pflugmacher, 2005), with a molecular weight around 299 (Svrcek and Smith, 2004). So far, twenty-seven variants of saxitoxins have been found (Ho et al., 2012). Saxitoxin is the most potent cyanobacterial toxin, with the LD₅₀ of 5 µg/kg in mice (administered intraperitoneally) (Carmichael, 1992).

Within the *Dolichospermum* genus, saxitoxin production has been only reported in *D. circinalis* (*Anabaena circinalis*) blooms in Australian freshwater systems, such as dams, weirs on rivers, and reservoirs (Beltran and Neilan, 2000). The gene cluster responsible for saxitoxin biosynthesis (genes *sxtA-E* and *sxtG-X* spanning 28 kbp) has been described for *D. circinalis* (Mihali et al. 2009) (Fig. 4). A specific quantitative PCR (qPCR) method based on SYBR green chemistry was developed to quantify saxitoxin-producing *D. circinalis* (Al-Tebrineh et al. 2010), and saxitoxin concentrations were positively correlated with *sxtA* gene copy numbers, indicating that the latter can be used as a measure of potential toxigenicity in *D. circinalis* and possibly other cyanobacterial blooms. This assay was expanded and combined into a

1 multiplex qPCR assay to detect *sxtA*, *cyrA*, *mcyE-ndaF* and 16S rRNA
 2 from most cyanobacterial species known to produce the saxitoxin,
 3 cylindrospermopsin, microcystin and nodularin toxins (Al-Tebrineh et al.,
 4 2012). This assay was applied to assess cyanobacterial composition and
 5 toxigenicity of concurrent blooms along 1,100 km of the Murray River in
 6 Australia during 2009.

7 **4.2.3 Anatoxin-a**

8 Anatoxin-a is a bicyclic secondary amine (van Apeldoorn et al., 2007)
 9 with a molecular weight of 165 (Svrcek and Smith, 2004). It is unique to
 10 cyanobacteria (Carmichael, 1994). Anatoxin-a degrades rapidly, with a
 11 half-life of 1 to 2 h (Sivonen and Jones, 1999), but is more stable in the
 12 protonated form than the free base (van Apeldoorn et al., 2007). The
 13 LD₅₀ (i.e. subscript) is 200 µg/kg in mice (i.p.) (Carmichael et al., 1979).
 14 Anatoxin-a was first found in *Anabaena. flos-aquae* isolated from
 15 Canada (Carmichael et al., 1979), and later in blooms of
 16 *Dolichospermum* spp. (Sivonen et al., 1989). The anatoxin-a synthetase
 17 gene cluster has been elucidated from *Dolichospermum (Anabaena)*
 18 *flos-aquae* NRC-44 (Devlin et al., 1977) and *Anabaena* sp. strain 37
 19 (genes *anaA-G* spanning 29 kbp) (Rantala-Yilnen et al. 2011) (Fig. 4)
 20 and molecular methods to detect potential anatoxin-a producers of the
 21 genera *Dolichospermum*, *Aphanizomenon*, and *Oscillatoria* were
 22 developed by designing primers to recognize the *anaC* gene in the cluster

(Rantala-Yilnen et al. 2011).

4.2.4 Anatoxin-a(S)

Anatoxin-a(S) (S associated with the symptoms of salivation) (Svrcek and Smith, 2004), is different both in structure and toxicity mechanism from anatoxin-a (Corbel et al., 2014). No structural variants have yet been found. Anatoxin-a(S) is a phosphate ester of a cyclic N-hydroxy guanine with a molecular weight of 252 (Henriksen et al., 1997; van Apeldoorn et al., 2007). It is also unique to cyanobacteria (Carmichael, 1994). Its rare detection may be due to inherent chemical instability (Matsunaga et al., 1989); anatoxin-a(S) decomposes rapidly in basic solutions but is more stable under neutral and acidic conditions.

Anatoxin-a(s) is a potent acetylcholinesterase inhibitor at the nerve synapse, which results in paralysis, asphyxiation, decreased movement, exaggerated abdominal breathing, cyanosis, convulsion, and ultimately death (Mahmood and Carmichael, 198, 1987; Chorus and Bartram, 1999). The LD50 is 20 µg/kg in mouse (i.p.) (Carmichael, 2001b). To date, anatoxin-a(S) has been detected only in *Dolichospermum* species, including *D. flos-aquae*, *D. lemmermannii*, and *D. spiroides* (Mahmood & Carmichael, 1987; Henriksen et al., 1997; Onodera et al., 1997). The natural occurrence of anatoxin-a(S) remains poorly documented (Devic et al. 2002). The primary analytical method for detection is high-performance liquid chromatography plus mass spectroscopy. An

alternative is to use the strong inhibition of acetylcholinesterase (AChE) by the toxin as a sensor (Mahmood & Carmichael, 1987). The genetic basis for biosynthesis of anatoxin-a(S) is not known, and no molecular methods for detecting anatoxin-a(S) genes are available.

4.3 Lipopolysaccharides (LPS)

Lipopolysaccharides are present in all cyanobacteria as an obligate part of the outer cell wall (Codd et al., 2005). LPS is composed of lipid A (core polysaccharides) and an outer polysaccharide chain (Wiegand and Pflugmacher, 2005), with structural features conserved within a genus and varying between genera (Kerr et al., 1995).

The mammalian inflammatory response to LPS occurs through toll-like receptor TLR4, a component of the innate immune system. Exposure experiments with zebra fish have demonstrated inhibition of glutathione S-transferase, an enzyme involved in detoxifying bioactive compounds such as microcystin (Best et al., 2002; Jaja-Chimedza et al., 2012 missing). LPS may thus exacerbate the toxic effects of cyanotoxins.

5. Taste and odor substances

Taste and odor (T & O) problems, mainly caused by secondary metabolites of microorganisms in aquatic ecosystems, have been frequently reported worldwide (Jiang et al., 2007 Giglio et al., 2008; Wang et al. 2011; Kutovaya and Watson 2014). Geosmin and

1 2-methylisoborneol (2-MIB) are sesquiterpene and monoterpene
2 molecules, respectively, that are responsible for the most unpleasant and
3 widely distributed earthy-musty off-flavors in freshwater bodies.
4 Cyanobacteria are often the main biological source of geosmin and
5 2-MIB in aquatic eutrophic systems, and earthy-musty odor problems
6 caused by cyanobacteria have been frequently reported. These problems
7 have a significant economic impact on drinking water production and
8 have become a research priority. Various species of *Dolichospermum*
9 (*Anabaena*) such as *D. flos-aquae*, *D. laxum*, *D. lemmermannii*, *D.*
10 *planktonicum* and *D. ucrainicum* have been reported to produce
11 geosmin (Giglio et al., 2008; Kutovaya and Watson 2014). Chemical
12 analytic methods using GC and GC-MS have long been used to detect
13 geosmin and 2-MIB in natural waters or in cyanobacterial cultures, but
14 there is a strong need for a sensitive, reliable and rapid diagnostic method
15 for detecting, monitoring and managing geosmin events.

16 Geosmin is synthesized from the universal sesquiterpene precursor,
17 farnesyl diphosphate (FPP), in a two-step process catalyzed by a
18 bifunctional sesquiterpene cyclase known as geosmin synthase (*geoA*) in
19 the presence of Mg^{2+} (Giglio et al., 2008). Su et al. (2013) developed a
20 quantitative PCR method for assaying *Anabaena* (*Dolichospermum*) sp.
21 geosmin synthase gene copy numbers. Based on sequences from the
22 *geoA* gene in the cyanobacteria *Phormidium* sp. and *Nostoc punctiforme*

1 PCC 73102, Kutovaya and Watson (2014) developed an assay for the
2 detection of *geoA* in a wide range of geosmin-producing taxa in both
3 laboratory cultures and environmental samples. This study indicated that
4 potential geosmin producers in Laurentian Great Lakes water and
5 sediment samples are predominantly represented by members of the
6 Nostocaceae, in particular *Dolichospermum* spp. Recently, Wang et al.
7 (2015) identified and cloned the geosmin synthase gene (*geo*) of two *D.*
8 *ucrainicum* strains isolated from Erhai Lake in China and their flanking
9 regions by polymerase chain reaction (PCR) and genome walking. The
10 *geo* gene was found to be located in a transcription unit with two cyclic
11 nucleotide-binding protein genes (*cnb*) (Fig 4). The two *cnb* genes were
12 highly similar and were predicted to be members of the cyclic adenosine
13 monophosphate (cAMP) receptor protein/fumarate nitrate reductase
14 regulator (Crp–Fnr) family (Wang et al. 2015). Identification of the *geo*
15 gene cluster in additional *Dolichospermum* species will allow better
16 molecular detection of *Dolichospermum* geosmin producers.

17 The gene cluster for 2-MIB synthesis has been identified in
18 oscillatorean cyanobacteria such as *Pseudanabaena* sp. and
19 *Planktothricoides raciborskii* (Giglio et al., 2011; Wang et al., 2011). So
20 far, no 2-MIB producing *Dolichospermum* species have been found in
21 natural and cultured strains.

22 **6. Conclusion**

1 In conclusion, the heterocystous *Dolichospermum* group of
2 cyanobacteria is characterized by high morphological diversity and
3 diverse cyanotoxin production. Taxonomic uncertainties affecting this
4 group, especially the relationship between *Dolichospermum* and
5 *Aphanizomenon* species, remain to be clarified and will benefit from
6 genome sequencing of additional isolates. Genome sequences will also
7 facilitate studies on the precise environmental preferences of particular
8 species or isolates, which ultimately should provide an explanation of the
9 wide variation of conditions in which members of the *Dolichospermum*
10 genus produce blooms. The long-term goal of much current research is to
11 develop the understanding that would support models capable of
12 predicting *Dolichospermum* population dynamics, and mitigating or at
13 least anticipating the detrimental effects of their blooms.

14

15 **Acknowledgements**

16 Work in the Li Renhui laboratory is funded by National Science
17 Foundation China (No. 31470310). Research in the Dreher laboratory
18 was supported by the US Geological Survey (G12AP20157) and
19 Washington State Department of Ecology (G1300115), with additional
20 support from the Oregon State University Agricultural Experiment
21 Station and the Mabel E. Pernot Trust. We thank Nathan Brown for
22 provision of preliminary results and for assistance with Fig. 3.

Legends for Figures

Fig.1

Morphology of *Dolichospermum* species, *Chrysosporum* species, *Sphaerospermopsis* species and *Aphanizomenon* species in pure cultures or in environmental samples. (a) *Dolichospermum compactum*; (b) *Dolichospermum planctonicum*; (c, d) *Dolichospermum affinis*; (e) *Dolichospermum circinalis*; (f, g) *Dolichospermum spiroides*; (h) *Dolichospermum mendotae*; (i) *Dolichospermum ucranicum*; (j) *Dolichospermum smithii*; (k) *Dolichospermum flos-aquae*; (l, m) *Chrysosporum bergii*; (n) *Sphaerospermopsis eucompacta*; (o) *Sphaerospermopsis oumiana*; (p) *Sphaerospermopsis kisseleviana*; (q) *Sphaerospermopsis aphanizomenoides*; (r) *Aphanizomenon gracil*; (s, t, u, v) *Aphanizomenon flos-aquae*. Scale bars, 10 μ m (scale bars for f, n, t, 20 μ m).

Fig.2. Neighbor joining phylogenetic tree of heterocytous cyanobacteria (within the order Nostocales) inferred from the 16S rRNA gene. Bootstrap values below 50% are not shown at the nodes.

Fig.3 Genome-wide average nucleotide identity (ANI) among *Dolichospermum*/*Anabaena*/*Aphanizomenon* isolates listed in Table 2. Two-way pairwise ANI was calculated according to Goris et al. (2009) (<http://enveomics.ce.gatech.edu/ani/>)

Fig. 4

Gene clusters of toxins produced by various *Dolichospermum* or *Aphanizomenon* morphospecies. Microcystin in *Dolichospermum* sp. 90 (54.858 kb), anatoxin-a in *Dolichospermum* sp. 37 (28.026 kb), saxitoxin in *Dolichospermum circinalis* AWQC131C (28.397 kb), and cylindrospermopsin in *Aphanizomenon* sp. 10E6 (39.612 kb). Genebank numbers for the upper four toxin gene clusters are CP003284, JF803645, DQ787201, and GQ385961.

Fig. 5.

Gene cluster responsible for synthesis of geosmin in two strains of *D. ucranicum*. Upper: strain CHAB 1432 ; Lower: strain CHAB 2155

1

2 Fig.1

3

4

5

6

7

8

9

10

11

12

13

14

Fig. 2

	Ana. ATCC 29413	Nos. PCC 7120	Ana. PCC 7122	Ana. PCC 7108	Apha. UKL13	Dol. 131F	Dol. 310F	Ana. 90	Apha. NIES 81	Apha. KM1-D3	Ana. WA102
Ana. ATCC 29413	100	93.2	78	78.1	76.5	76.5	76.5	77.3	76.7	77.4	77.2
Nos. PCC 7120	93.2	100	78.5	78	76.4	76.6	76.4	77.5	76.7	77.3	77.5
Ana. PCC 7122	78	78.5	100	84.4	81	80.8	81	81.7	81.4	81.6	82
Ana. PCC 7108	78.1	78	84.4	100	81.9	81.5	81.5	82.7	82.4	82.5	82.6
Apha. UKL13	76.5	76.4	81	81.5	100	87.4	87.6	87.7	87	87.2	87.4
Dol. 131F	76.5	76.6	80.8	81.5	87.4	100	96.8	85.9	85.3	85.4	85.6
Dol. 310F	76.5	76.4	81	81.9	87.4	87.6	100	86	85.3	85.5	85.8
Ana. 90	77.3	77.5	81.7	82.7	87.7	85.9	86	100	90.4	91	91.6
Apha. NIES 81	76.7	76.7	81.4	82.4	87	85.3	85.3	90.4	100	97	97.4
Apha. KM1-D3	77.4	77.3	81.6	82.5	87.2	85.4	85.5	91	97	100	97.4
Ana. WA102	77.2	77.5	82	82.6	87.4	85.6	85.8	91.6	96.6	97.4	100

Fig.4

1

2 Fig. 5

3

4

Table 1. Comparison of phenotypic characters among the genus *Anabaena*, *Dolichospermum* and morphologically similar genera, such as *Aphanizomenon*, *Sphaerospermopsis* and *Chrysosporum*.

Genus	Type species	Distinct phenotypic characters	The relative position of akinete to heterocyst	Aerotopes	Habitat	Reference
<i>Anabaena</i>	<i>Anabaena oscillarioides</i>	Benthic, aerotopes absent in cells				
Bory de Saint-Vincent ex Bornet et Flahault 1888	Bornet ex Bornet et Flahault 1888	Trichomes in mats, straight or coiled	Remote or adjacent	Absent	Benthic	Komárek, 2013
		Terminal cells with rounded ends				
<i>Dolichospermum</i>	<i>Dolichospermum flos-aquae</i>	Planktic, aerotopes present in cells				
(Ralfs ex Bornet et Flahault)	(Brébisson ex Bornet et Flahault)	Trichomes straight or coiled	Remote or adjacent	Present	Planktic	Wacklin <i>et al.</i> , 2009
Wacklin, Hoffmann et Komárek 2009	Wacklin <i>et al.</i> , 2009	Terminal cells with rounded ends				
<i>Aphanizomenon</i>	<i>Aphanizomenon flos-aquae</i>	Planktic, aerotopes present in cells				Rajaniemi <i>et al.</i> , 2005a
Morren ex Bornet et Flahault 1888	Ralfs ex Bornet et Flahault 1888	Trichomes straight	Remote or adjacent	Present	Planktic	Rajaniemi <i>et al.</i> , 2005b
		Terminal cells cylindrical, elongated				
<i>Sphaerospermopsis</i>	<i>Sphaerospermopsis reniformis</i>	Planktic, aerotopes present in cells				Zapomělová <i>et al.</i> , 2009
Zapomělová <i>et al.</i> , 2010	Zapomělová <i>et al.</i> , 2010	Trichomes straight or coiled	Adjacent	Present	Planktic	Zapomělová <i>et al.</i> , 2010a
		Spherical akinetes adjacent to heterocystes				
<i>Chrysosporum</i>	<i>Chrysosporum bergii</i>	Planktic, aerotopes present in cells				
Zapomělová <i>et al.</i> , 2012	(Stenfeldt) Zapomělová <i>et al.</i> , 2012	Trichomes straight or slightly curved	Remote or adjacent	Present	Planktic	Zapomělová <i>et al.</i> , 2012
		Terminal cells slightly narrowed, with conical ends				

Table 2. The name change of previously planktonic species of *Anabaena*

Previous

<i>Anabaena compacta</i> (Nygaard) Hickel 1985
<i>Anabaena flos-aquae</i> (Lyngbye) Brébisson ex Bornet et Flahault 1888
<i>Anabaena circinalis</i> Rabenhorst ex Bornet et Flahault 1888
<i>Anabaena perturbata</i> Hill 1976
<i>Anabaena spiroides</i> Klebahn 1895
<i>Anabaena crassa</i> (Lemmermann) Komárková-Legnerová et Cronberg 1992
<i>Anabaena ellipsoides</i> Bolochoncev 1909
<i>Anabaena curva</i> Hill 1976
<i>Anabaena berezowskii</i> Usačev 1928
<i>Anabaena mucosa</i> Komárková-Legnerová et Eloranta 1992
<i>Anabaena sigmoidea</i> Nygaard 1949
<i>Anabaena mendotae</i> Trelease 1898
<i>Anabaena farciminiiformis</i> Cronberg et Komárková-Legnerová 1988
<i>Anabaena longicellularis</i> (Pankow) Komárková-Legnerová et Eloranta 1992
<i>Anabaena lemmermannii</i> Richter 1903
<i>Anabaena shujae-laxa</i> Komárek et Zapomělová 2007
<i>Anabaena delicatula</i> Lemmermann 1898
<i>Anabaena halbfassii</i> Bachmann 1913
<i>Anabaena heterospora</i> Nygaard 1949
<i>Anabaena viguieri</i> Denis et Frémy 1924
<i>Anabaena danica</i> (Nygaard) Komárková-Legnerová et Eloranta 1992
<i>Anabaena solitaria</i> Klebahn 1895
<i>Anabaena macrospora</i> Klebahn 1895
<i>Anabaena zinserlingii</i> Kosinskaja 1929
<i>Anabaena planctonica</i> Brunnthaler 1903
<i>Anabaena smithii</i> (Komárek) M. Watanabe 1992
<i>Anabaena affinis</i> Lemmermann 1897
<i>Anabaena akankoensis</i> M. Watanabe 2003
<i>Anabaena bituri</i> Cronberg et Komárek 2004
<i>Anabaena caspica</i> Ostefeld 1902
<i>Anabaena fallax</i> Komárek et Komárková 2002
<i>Anabaena fusca</i> Hill 1976
<i>Anabaena helicoidea</i> Bernard 1908
<i>Anabaena nygaardii</i> Cronberg et Komárek 2004
<i>Anabaena pseudocompacta</i> M. Watanabe 1996
<i>Anabaena arctica</i> Kiselev 1932
<i>Anabaena bothai</i> Welsh 1961
<i>Anabaena citrispora</i> M. Watanabe 1992
<i>Anabaena jacutica</i> Kiselev 1932
<i>Anabaena maxima</i> Cronberg et Komárek 2004
<i>Anabaena nathii</i> Vasishta 1960
<i>Anabaena wernerii</i> Brunnthaler 1903
<i>Anabaena reniformis</i> Lemmermann 1898
<i>Anabaena aphanizomenoides</i> Forti 1912
<i>Anabaena kisseleviana</i> Elenkin 1938

Anabaena oumiana M. Watanabe 1996
Anabaena torques-reginae Komárek 1984
Anabaena eucompacta R. Li et M. Watanabe 1999

Aphanizomenon ovalisporum Forti 1911
Anabaena bergii Ostenfeld 1908
Anabaena bergii var. *minor* Kiselev 1927

Present

- Dolichospermum compactum* (Nygaard) Wacklin et al. 2009
Dolichospermum flos-aquae (Brébisson ex Bornet et Flahault) Wacklin et al. 2009
Dolichospermum circinale (Rabenhorst ex Bornet et Flahault) Wacklin et al. 2009
Dolichospermum perturbatum (Hill) Wacklin et al. 2009
Dolichospermum spiroides (Klebahn) Wacklin et al. 2009
Dolichospermum crassum (Lemmermann) Wacklin et al. 2009
Dolichospermum ellipsoides (Bolochoincev) Wacklin et al. 2009
Dolichospermum curvum (Hill) Wacklin et al. 2009
Dolichospermum berezowskii (Usačev) Wacklin et al. 2009
Dolichospermum mucosum (Komárková-Legnerová et Eloranta) Wacklin et al. 2009
Dolichospermum sigmoideum (Nygaard) Wacklin et al. 2009
Dolichospermum mendotae (Trelease) Wacklin et al. 2009
Dolichospermum farciminiforme (Cronberg et Komárková-Legnerová) Wacklin et al. 2009
Dolichospermum longicellulare (Pankow) Wacklin et al. 2009
Dolichospermum lemmermannii (Richter) Wacklin et al. 2009
Dolichospermum shujae-laxum (Komárek et Zapomělová) Wacklin et al. 2009
Dolichospermum delicatulum (Lemmermann) Wacklin et al. 2009
Dolichospermum halbfassii (Bachmann) Wacklin et al. 2009
Dolichospermum heterosporum (Nygaard) Wacklin et al. 2009
Dolichospermum viguieri (Denis et Frémy) Wacklin et al. 2009
Dolichospermum danicum (Nygaard) Wacklin et al. 2009
Dolichospermum solitarium (Klebahn) Wacklin et al. 2009
Dolichospermum macrosporum (Klebahn) Wacklin et al. 2009
Dolichospermum zinserlingii (Kosinskaja) Wacklin et al. 2009
Dolichospermum planctonicum (Brunnthal) Wacklin et al. 2009
Dolichospermum smithii (Komárek) Wacklin et al. 2009
Dolichospermum affine (Lemmermann) Wacklin et al. 2009
Dolichospermum akankoense (M. Watanabe) Wacklin et al. 2009
Dolichospermum bituri (Cronberg et Komárek) Wacklin et al. 2009
Dolichospermum caspicum (Ostenfeld) Wacklin et al. 2009
Dolichospermum fallax (Komárek et Komárková) Wacklin et al. 2009
Dolichospermum fuscum (Hill) Wacklin et al. 2009
Dolichospermum helicoideum (Bernard) Wacklin et al. 2009
Dolichospermum nygaardii (Cronberg et Komárek) Wacklin et al. 2009
Dolichospermum pseudocompactum (M. Watanabe) Wacklin et al. 2009
Dolichospermum arcticum (Kiselev) Wacklin et al. 2009
Dolichospermum bothai (Welsh) Wacklin et al. 2009
Dolichospermum citrisporum (M. Watanabe) Wacklin et al. 2009
Dolichospermum jacuticum (Kiselev) Wacklin et al. 2009
Dolichospermum maximum (Cronberg et Komárek) Wacklin et al. 2009
Dolichospermum nathii (Vasishta) Wacklin et al. 2009
Dolichospermum weneri (Brunnthal) Wacklin et al. 2009

Sphaerospermopsis reniformis (Lemmermann) Zapomělová et al. 2010
Sphaerospermopsis aphanizomenoides (Forti) Zapomělová et al. 2010
Sphaerospermopsis kisseleviana (Elenkin) Zapomělová et al. 2010

Sphaerospermopsis oumiana (M. Watanabe) Tuji et Niiyama 2012
Sphaerospermopsis torques-reginae (Komárek) Werner et al. 2010
Sphaerospermopsis eucompacta (R. Li et M. Watanabe) X. Li et R. Li 2015

Chrysosporum ovalisporum (Forti) Zapomělová et al. 2012
Chrysosporum bergii (Ostenfeld) Zapomělová et al. 2012
Chrysosporum minor (Kiselev) Zapomělová et al. 2012

	Isolate	Origin	Chromosome Mpb / G+C%	Plasmids (kbp / G+C%	Strategy No. contigs / scaffolds	No. protein- coding genes	No. ribosomal operons/ tRNA genes	Comments	Citation / GenBank ID
Finished genomes	<i>Anabaena</i> sp. 90	Lake Vesijärvi, Finland, 1986	4.3 + 0.8 = 5.1 / 38.1, 38.2	80, 56, 20 / 37.2- 37.4	Cloning, Sanger, PCR to finish	4517	5 / 44	Microcystin producer	Wang et al., 2012 GCA_000312705.1
	<i>Anabaena</i> sp. WA102	Anderson Lake, Washington, USA, 2013	5.7 / 38.4	77 / 37.2	PacBio	5271	5 / 44	Anatoxin-a producer	N. Brown & T. Dreher, in prep. CP011456
	<i>Nostoc</i> (<i>Anabaena</i>) sp. PCC 7120	Uncertain, USA, pre-1979	6.4 / 41.3	408, 187, 102, 55, 40, 6 / 40.2 - 44.2	Cloning Sanger, PCR to finish	5866	4 / 48	Nostoc but still widely referred to as <i>Anabaena</i>	Kaneko et al., 2001 GCA_000009705.1
	<i>Anabaena</i> <i>variabilis</i> ATCC 29413	Mississippi, USA, 1964	6.4 + 0.04 / 41.4, 46.5	366, 300, 36 / 38.5 - 42.0		5671	4 / 47	Freshwater Also known as <i>Anabaena flos-</i> <i>aquae</i> UTEX 1444	Unpublished GCF_000204075.1
	<i>Anabaena</i> <i>cylindrica</i> PCC 7122	Cambridge, UK, 1939	6.4 / 38.8	220, 178, 129, 98, 23, 21 / 37.1- 40.1	Roche 454, Illumina, PCR to finish	5856	4 / 61		Shih et al., 2013 GCF_000317695.1
Draft genomes	<i>Dolichospermum</i> <i>circinale</i> (<i>Anabaena</i> <i>circinalis</i>) AWQC310F	Milawa, Victoria, Australia, 1995	4.4 / 37.5		372 / 82 Illumina	4443*		Saxitoxin negative	Unpublished APIZ01
	<i>Dolichospermum</i> <i>circinale</i> (<i>Anabaena</i> <i>circinalis</i>) AWQC131C	Lake Cargelligo, NSW, Australia, 1991	4.4 / 37.3		472 / 121 Illumina	4447*		Saxitoxin producer	Unpublished APIY01
	<i>Aphanizomenon</i> <i>flos-aquae</i> , <i>f.</i> <i>gracile</i> NIES-81	Lake Kasumi- gaura, Japan, 1978	5.9 / 38.5		338 / -103 Roche 454	4972	5 / 42	Saxitoxin positive	Cao et al., 2014 AZY01
	<i>Aphanizomenon</i> <i>flos-aquae</i> 2012/KM1/D3	Curonian Lagoon, Baltic Sea, Lithuania, 2012	5.7 / 38.2		325 / 325 / - Ion Torrent	4415	5 / 38	Brackish water isolate	Sulcius et al., 2015; JSPD01
	<i>Anabaena</i> sp. PCC 7108	Moss Beach, California, USA, 1970	5.9 / 38.8		13 / 3 Roche 454 + Illumina	4950		Intertidal zone	Shih et al., 2013 AJWF01

Table 3 Finished or draft genomes of members of the *Dolichospermum*/*Anabaena*/*Aphanizomenon* group and of isolates currently named *Anabaena* that fall outside this group (shaded).

* Data from D'Agostino et al., 2014