GENOMIC AND PHYSIOLOGICAL CHARACTERISATION OF SELECTED ARCTIC, TEMPERATE AND TROPICAL *Pseudanabaena* STRAINS

by

ZOYA KHAN

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

May 2019

ACKNOWLEDGEMENT

Firstly, I am truly grateful to Allah for blessing me the courage and determination to complete my thesis. I would like to express my utmost gratitude to my beloved supervisor Prof. Dr. Wan Maznah Wan Omar for her guidance, critical evaluation, and patience throughout the stages of this work. Her encouragement and support during the time of my graduate study are highly appreciated. I would like to thank my co-supervisors Prof. Peter Convey, Dr. Faradina Merican and Prof. Nazalan Najimudin for their valuable guidance, scholarly inputs, consistent encouragement and unconditional support throughout my research work. Sincere thanks to Dr. Foong Choon Pin who has been very kind to extend support and advices. I am thankful to Universiti Sains Malaysia for giving me an opportunity to complete my study here. My sincere appreciation to my lab mates and friends Basri, Azmir, Wong, CC, Lia, Mohd. Adnan and Kak Nani for their help and guidance. Special thank also goes to my other lab mates Fiqah, Mimie, Diyanah, Ayu, Didi, Mun, Ran, Nana, Dinie and Suraj. All these people have provided support and friendship that I will treasure forever. Thanks are extended to Mr. Tan, friendly staffs in the School of Biological Sciences and also to all my friends.

To my parents Wasiullah Khan and Shamina Khan and my beautiful family, the biggest thanks of all for all the years of love that provided me with inspiration and was my driving force. Special thanks to my brother Dr. Mohd. Jamal Khan and sister-in-law Sahrish for their unconditional support throughout my research work. To Shadab Hussain, for his consistent encouragement throughout my work. Finally, I would like to acknowledge Flagship grant (304/PBIOLOGI/650723/P131) under Ministry of Science, Technology and Innovation, Malaysia for providing the funds for my study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

length. (b) The *cpc2* gene cluster containing six *cpc* genes and approximately 3.5 kb in length.

LIST OF PLATES

LIST OF SYMBOLS AND ABBREVIATIONS

%	Percentage
xg	Times gravity
bp	Base pair
°C	Degree Celsius
Approx.	Approximately
APC	Allophycocyanin
BLAST	Basic local alignment search tool
Chl-*a*	Chlorophyll *a*
cells mL^{-1}	Cells per millilitre
dNTP	Deoxyribonucleotide triphosphate
DNA	Deoxyribonucleotide acid
EDTA	Ethylenediaminetetraacetic acid
d	Day
g	Gram
h	Hour
kDa	KiloDalton
L	Litre
min	Minute
mg	Milligram
mL	Milliliter
nm	Nanometer
NIVA	Norwegian Institute for Water Research
PCR	Polymerase chain reaction
PBP	Phycobiliprotein

TPBP	Total Phycobiliprotein
PC	Phycocyanin
PE	Phycoerythrin
pH	Potential hydrogen
rpm	Rotation per minutes
µg	Microgram
µg/mL	Microgram per milliliter
µL	Microliter
µm	Micrometer

LIST OF APPENDICES

CIRI GENOMIK DAN FISIOLOGI STRAIN *Pseudanabaena* YANG TERPILIH DARI ARTIK, TEMPERAT DAN TROPIKA

ABSTRAK

Tiga strain sianobakteria berfilamen yang diperoleh dari kawasan Kepulauan Artik Svalbard Tinggi, kawasan temperat (England) dan kawasan tropika (Malaysia) telah dicirika dengan menggunakan gabungan pendekatan fenotip dan genetik. Pemerhatian morfologi dan ultrastruktur dilakukan bersama dengan pengukuran pertumbuhan. Dimensi sel, susunan thalakoid dan bentuk sel apikal bagi strain Artik dan temperat adalah selaras dengan keterangan *Pseudanabaena catenata* manakala strain tropika dikenalpasti sebagai *Pseudanabaena amphigranulata*. Analisis gen 16S rRNA menunjukkan strain Artik USMAC16 dan strain temperat NIVA-CYA146 berkongsi identiti jujukan berpasangan yang tinggi (100% dan 98%) dengan *P. catenata* SAG 1464-1. Strain tropika USMAC18 berkongsi hanya 94% jujukan berpasangan dengan *Pseudanabaena* sp. PCC 6802, menunjukkan strain itu berbeza dengan wakil genus yang mewakili daripada strain yang ada dalam pangkalan data sediada. Ketiga-tiga strain menunjukkan konfigurasi 16S-23S ITS yang sama dengan strain *Pseudanabaena* yang lain. Kultur setiap strain didedahkan kepada pelbagai suhu dan fotokala untuk mengenalpasti keplastikan fenotip. Panjang maksimum sel strain Artik dan strain temperat (5.92 ± 0.7 μm dan 5.79 ± 0.26 μm, masing-masing) diperhatikan pada suhu kultur 15 ° C, dan strain tropika (5.7 ± 0.07 μm) pada 25 ° C, semuanya di bawah fotokala 12L: 12D jam (L: cahaya, D: gelap). Strain menunjukkan kepekaan yang tinggi terhadap dimensi sel dan bentuk pada suhu dan fotokala yang berbeza, dengan 15 ° C dan 25 ° C di bawah 12L: 12D h memberikan keadaan yang optimum untuk pertumbuhan mereka. Kuantiti fikobiliprotein (terutamanya fikosianin

dan fikoeritrin) dan faktor yang mempengaruhi pengeluaran mereka juga dikenalpasti. Pengaruh panjang gelombang cahaya yang berbeza (putih, hijau dan merah) dan tempoh pendedahan (fotokala 12-24 jam (h)) terhadap pengeluaran fikosianin (PC) dan fikoeritrin (PE) dalam tiga strain ditentukan. Seterusnya, pengeluaran fikobiliprotein dibandingkan antara ketiga-tiga strain. Pengeluaran PC dan PE paling tinggi dicapai di bawah cahaya merah dan hijau, masing-masing dengan fotokala 24:00 h L: D bagi strain polar (25.8 ± 2.8 dan 25.5 ± 5.1 mg / L), manakala 12:12 h L: D bagi strain temperat (97.5 ± 12.3 dan 64.31 ± 19.6 mg / L) dan strain tropika (86 ± 14.7 dan 10.1 ± 3.9 mg / L). *P. catenata* (strain temperat) adalah pengeluar fikoeritrin yang baik apabila dikultur di bawah cahaya hijau. Genom ketiga-tiga strain *Pseudanabaena* berjaya dijujukkan dan dianalisis. Saiz genom ketiga-tiga strain adalah kira-kira 5.5 Mb dengan kandungan G + C sebanyak 42-44%. Sejumlah 2293 gen pengekodan protein dikenalpasti. Analisis genom menunjukkan dua kelompok fikosianin (cpc) di mana tiga homolog *cpcA* dan *cpcB* hadir. Sejumlah gen pengekodan untuk ciri adaptasi sejuk telah dijumpai dalam genom strain Artik yang menerangkan tingkah laku strain psikrotoleran ini. Kajian ini telah menggabungkan pendekatan taksonomi morfologi tradisional dengan molekul moden yang memberikan pandangan baru ke atas taksonomi *Pseudanabaena*. Kajian ini juga menunjukkan bahawa ketiga-tiga strain *Pseudanabaena* boleh digunakan sebagai sumber fikosianin dan fikoeritrin yang baik apabila dikultur di bawah cahaya merah dan hijau. Hasil penjujukan seluruh genom telah memberikan maklumat mengenai fisiologi dan penyesuaian strain *Pseudanabaena* terhadap perubahan persekitaran.

GENOMIC AND PHYSIOLOGICAL CHARACTERISATION OF SELECTED ARCTIC, TEMPERATE AND TROPICAL *Pseudanabaena* STRAINS

ABSTRACT

Three filamentous cyanobacteria strains obtained from the High Arctic Svalbard archipelago, temperate (England) and tropical (Malaysia) regions were characterized using combined phenotypic and genetic approaches. Morphological and ultrastructural observations were performed together with growth measurements. Cell dimensions, thylakoid arrangement and apical cell shape of the Arctic and temperate strains were consistent with the description of *Pseudanabaena catenata* while the tropical strain was identified as *Pseudanabaena amphigranulata*. 16S rRNA gene analysis showed that the Arctic strain USMAC16 and temperate strain NIVA-CYA146 shared high sequence similarity (100% and 98%, respectively) with *P. catenata* SAG 1464-1. The tropical strain USMAC18 shared only 94% sequence similarity with *Pseudanabaena* sp. PCC 6802, suggesting that the strain is distinctly different from the strains currently available in the databases. All three strains showed identical internal transcribed spacer (ITS) configuration with other strains of *Pseudanabaena*. Cultures of each strain were exposed to various temperature and photoperiod treatments in order to examine phenotypic plasticity. The maximum cell length of Arctic and temperate strains (5.92±0.7 μm and 5.79±0.26 μm, respectively) was observed at 15°C culture temperature, and that of the tropical strain (5.7±0.07 μm) was at 25°C, all under 12L:12D hours (L: light, D: dark) photoperiod. The strains showed high plasticity in cell dimensions and shape under different temperature and

photoperiod treatments, with 15°C and 25°C under 12L:12D h providing the optimal conditions for their growth. The quantity of phycobiliproteins (mainly phycocyanin and phycoerythrin) and factors that affect their production was also investigated. The influence of different light wavelengths (white, green and red) and exposure duration (photoperiod of 12-24 hours (h)) on phycocyanin (PC) and phycoerythrin (PE) production in the three strains was determined. Subsequently, the production of phycobiliprotein was compared between the three strains. Highest PC and PE production were achieved under red and green light, respectively with photoperiod of 24:00 h L:D in the polar strain (25.8±2.8 and 25.5±5.1 mg/L, respectively), while 12:12 h L: D in the temperate (97.5± 12.3 and 64.31±19.6 mg/L, respectively) and tropical strains (86±14.7 and 10.1±3.9 mg/L, respectively). *P. catenata* (temperate strain) was a good producer of phycoerythrin when grown under green light. The genomes of all three *Pseudanabaena* strains were successfully sequenced and analysed. The genome sizes of all three strains were approximately 5.5 Mb with G+C content of 42-44%. A total of 2293 protein coding genes were identified. Genome analysis identified two phycocyanin (*cpc*) gene clusters in which three homologues of *cpcA* and *cpcB* were present. A number of genes coding for cold adapted features were present in the genome of the Arctic strain which explains its psychrotolerant behaviour. This study has integrated traditional morphological and modern molecular taxonomic approaches providing new insights into the taxonomy of the genera *Pseudanabaena*. The three *Pseudanabaena* strains can be used as a good source of phycocyanin and phycoerythrin when grown under red and green light. The results of whole genome sequencing has revealed information on the physiology and adaptation of *Pseudanabaena* strains towards changing environments.

Chapter 1

GENERAL INTRODUCTION

Cyanobacteria or cyanophytes belong to the eubacterial domain, and are oxygenic, photosynthetic and gram-negative prokaryotes, lacking defined nucleus and other intercellular membrane-bound organelles (Seckbach and Oren, 2007). Their pronounced similarity with green algae earned them the name "blue-green algae". In terms of Earth's history, cyanobacteria occupy a privileged position. They are primary producers and play a significant role in the planetary carbon cycle and, as nitrogen-fixers, prominently in the nitrogen cycle (Knoll, 2008). Cyanobacteria are one of the major components of biodiversity in Arctic and Antarctic regions, where they form benthic mats and films at the bottom of lakes, ponds and streams, and on moist soil surfaces (Zakhia *et al.*, 2008). These communities often dominate ecosystem biomass and productivity, and contend with persistent low temperatures, exposure to repeated freeze–thaw cycles and highly variable light, nutrient and osmotic regimes (Vincent, 2000a). They are also an important phototrophic element of biodiversity in temperate and tropical regions (Whitton and Potts, 2007; Hoffman, 1999).

The classification of cyanobacteria has long proved to be problematic (Waterbury, 2006). Earlier studies assumed that phenotypic and genotypic characters, made accessible by the availability of pure cultures, would simplify the classification of cyanobacteria. However, this in large part has not proved to be the case (Waterbury, 2006). Therefore, comparative molecular studies of cyanobacteria can contribute significantly to the revision of their taxonomy. Classification should reflect phylogenetic relationships, for example as encoded in DNA gene sequence data or ribosomal RNA sequence data such 16S or 23S (Mur *et al.*, 1999). The combination of phenotypic, genotypic and phylogenetic information generates a consensus form of

taxonomy known as polyphasic or integrated taxonomy (Vandamme *et al.*, 1996). Further combination of phenotypic, chemotypic, and genotypic features is commonly used to characterise cyanobacterial taxa at the species or lower level (Caroppo *et al.*, 2012; Yu *et al.*, 2015).

Cyanobacteria have a number of distinct properties that determine their relative importance in natural microbial communities. Many of these have been widely studied for their possible applications in industries such as mariculture, food, feed, fuel, fertilizer, biopolymers, natural colorants, vitamins, toxins, enzymes, pharmaceuticals, pharmacological fluorescent probes, and pollution abatement (Shrivastava *et al.*, 2010). Phycobiliproteins like phycocyanin and phycoerythrin are the accessory photosynthetic pigments of cyanobacteria, and such natural pigments and fluorescent proteins have versatile applications (Ojit *et al.*, 2015). However, among various cyanobacterial taxa, only a few such as *Spirulina*, *Nostoc*, *Porphyra*, and *Porphyridium* (Soni *et al.*, 2006; Mishra, 2007) are well-characterized and used commercially. Many cyanobacteria are highly adaptable towards different light wavelengths, intensities and photoperiods in different environments, which can also be economically exploitable. various studies have integrated the use of different light quality and intensity to enhance pigment production (Prasanna *et al.*, 2004; Chaneva *et al.*, 2007; Stomp *et al.*, 2008; Mishra *et al.*, 2012; Ojit *et al.*, 2015).

1.1 Problem statement

Many cyanobacteria, having relatively large colony or filament sizes and the ability to float near the water surface through their gas vesicles have been studied comprehensively. However, potential applications of microscopic cyanobacteria species, including those naturally occurring in extreme environments such as the polar regions, have largely been ignored to date because of difficulties in isolation and observation due to their relatively small cell/filament sizes. Occasionally, these cyanobacteria dominate phytoplankton biomass (Sthapit *et al.*, 2008), although generally they have been characterised based on few characters or only recorded as species names probably because of poor understanding of their features (Mischke and Nixdorf, 2003; Villena and Romo, 2003).

The non-heterocystous genus *Pseudanabaena* is also a lesser known microscopic cyanobacteria with potentially useful pharmaceutical and commercial properties (Oufdou *et al.*, 2001). Members of the genus play key roles in ecosystem dynamics and primary productivity (Bertos-Fortis *et al.*, 2016). Despite the cosmopolitan distribution of the genus, taxonomic information, such as ultrastructure features, pigment composition, DNA sequence data and confirmation of interspecific phylogenetic relationships for most *Pseudanabaena* species remains incomplete (Yu *et al.*, 2015). Therefore, increasing our knowledge of the ecology, morphological variability and phylogenetic status of members of this key genus from different places is of particular interest and importance.

Taking advantage of the rapid current developments in molecular and systematic approaches, the purpose of this study was threefold (Figure 1.1). First, strains of *Pseudanabaena* from the High Arctic Svalbard archipelago (polar), England (temperate) and Malaysia (tropical) were subjected to detailed polyphasic

characterisation using a combination of electron microscopy together with 16S rRNA and 16S-23S internal transcribed spacer (ITS) sequencing. The morphological plasticity of these strains was documented under various temperature and light (photoperiod) regimes in order to confirm the morphological features that are stable and are of highest taxonomic importance. Second, this study characterises the pigmentation of the three strains to investigate their adaptation under different conditions of light quality and photoperiods. Third, this study described distinctive features of the genes and the genome of these *Pseudanabaena* strains through whole genome characterisation.

1.2 Objectives

The primary aim of this study was to characterize three filamentous cyanobacteria of the genus *Pseudanabaena* using combined morphological and molecular approaches. The objectives are as follow:

i. To identify selected *Pseudanabaena* strains using a polyphasic approach (morphological and molecular characterization, ultrastructure, phylogeny and phenotypic plasticity).

ii. To investigate the composition and chromatic adaptation of pigments in *Pseudanabaena* strains under exposure to different light quality and photoperiods.

iii. To determine the physiological characteristics of *Pseudanabaena* strains through whole genome sequencing.

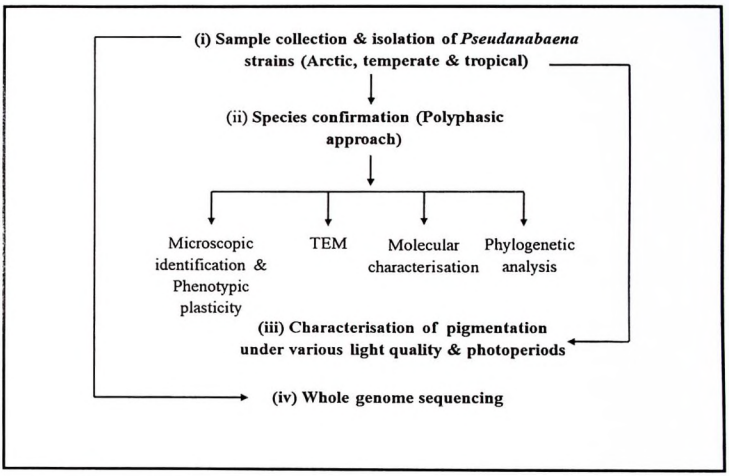

Figure 1.1. The flow of ideas and aims integrated in this study.

1.3 Thesis structure

'What is *Pseudanabaena* and why is it important?' is the central question of this thesis. The morphological, molecular and physiological characterisation of three strains of *Pseudanabaena* originating from Arctic, temperate and tropical regions were thoroughly examined in this study.

Chapter 2 provides a background literature reviewed for the studies described.

Chapter 3 introduces and describes the selected polar (Arctic), temperate (England) and tropical (Malaysia) sample collection sites as well as the sample collection, isolation and maintenance procedures.

Chapter 4 briefly describes the identification of selected *Pseudanabaena* strains using a polyphasic approach that combines morphological and molecular characterization, and the investigation of ultrastructural studies, phylogeny and phenotypic plasticity. The results presented answer the '*what*' question by confirming the identities of the *Pseudanabaena* strains at species level. The *why* question will be answered by advances in knowledge of this genus generated by this chapter.

Chapter 5 addresses the effects of different light wavelengths and photoperiods on the pigments of *Pseudanabaena* strains studied. The data obtained in this part also answer the '*why*' question, as pigments are of great commercial importance.

Chapter 6 describes physiological characterization of the *Pseudanabaena* strains through whole genome sequencing.

Chapter 7 and 8 provide general discussion and conclusions integrating the different parts of this study and place them in wider context.

Chapter 2

LITERATURE REVIEW

2.1 Cyanobacteria

Cyanobacteria, also known as cyanophytes belonging to the eubacterial domain, are oxygenic, photosynthetic and gram-negative prokaryotes, lacking defined nucleus and other intercellular membrane-bound organelles (Seckbach and Oren, 2007). Their functional similarity with green algae earned them the name "blue-green algae". Cyanobacteria have been present on the Earth for 3.5 billion years (Tomescu *et al.* 2006), and for most of this time were the sole photosynthesizers, contributing oxygen to the prehistoric anaerobic landscape (Seckbach and Oren, 2007), dramatically different from the present day.

Microfossils of cyanoprokaryotes found at the Apex Chert in Western Australia are believed to be 800 million years old (Schopf 1993). Fossil evidence from Early Silurian sediments in Virginia, USA, confirms that cyanobacteria were amongst the initial colonisers of extreme continental habitats (Tomescu *et al.* 2006). Studies from 2 billion year old fossils contain both unicellular and multicellular morphotypes of cyanobacteria (Amard and Bertrand-Sarfati, 1997). Moreover, an assumption can be drawn from the studies of Blankenship (2010), Bekker *et al.* (2004) and Allen and Martin (2007) that they were responsible for the rapid increase in atmospheric oxygen levels, known as the "Great Oxygenation Event".

Multicellular fossils of cyanobacteria have also been documented from the late Precambrian (Butterfield, 2009). Other multicellular filamentous fossils older than 3.0 billion years have been reported (Walsh, 1992; Tice and Lowe, 2004), some of which are morphologically analogous to species from the cyanobacterial order Oscillatoriales (Schopf *et al.*, 2007). These fragmentary fossil records of cyanobacteria, while

intriguing, are not sufficient to unravel the origin of cyanobacteria and their morphological phenotypes. Modern polyphasic approaches provide the possibility to gather further evidence about the evolution of this complex phylum (Schirrmeister *et al.*, 2011).

2.2 Occurrence in nature

Cyanobacteria inhabit limnic and marine environments. They flourish in extreme habitats such as saline, brackish or fresh waters, cold and hot springs, and in environments where no other microalgae can survive (Mur *et al.*, 1999). Cyanobacteria include a large component of marine plankton with cosmopolitan distribution (Gallon *et al.*, 1996). Numerous freshwater species are also able to endure relatively high salinity.

Many cyanobacteria isolated from coastal environments are capable of withstanding saline environments (halotolerant) and are capable of growing at salt concentrations as high as 3-4 molar rather than specifically requiring salinity (halophilic) (Mur *et al.*, 1999). Numerous species characteristically dwell in, and occasionally dominate the epilimnic, euphotic and hypolimnic water zones of lakes (Whitton, 1992). Some of them colonise the substrates by attaching to rocks or sediments, forming mucilaginous mats that float to the surface. These mats exhibit a remarkably high biodiversity compressed into a few millimetres (Stal and Caumette, 2013).

A wide range of phototrophs such as bacterial and cyanobacterial, heterotrophs and chemoautotrophs are found within these mucilaginous mats (Zehr *et al.*, 1995; Steppe *et al.*, 1996). Studies have shown that the role of microbial mats has been crucial throughout the history of the Earth for the composition and modification of its

atmosphere and also for the production of O_2, H_2, and CH_4 (Hoehler *et al.*, 2001). Therefore, microbial mats are, certainly, a natural laboratory where microbial diversity (community structure), evolutionary processes, and their adaptation to stress environments can be studied (Inskeep *et al.*, 2013; Villanueva *et al.*, 2007).

Cyanobacteria have an extraordinary capability to inhabit bare and infertile substrates such as desert sand, volcanic ash and rocks (Dor and Danin, 1996). The ability to produce scytonemin, a pigment that provides protection against short wavelength solar ultraviolet (UV) radiation, increases their fitness in the relatively exposed terrestrial environment (Rastogi *et al.*, 2015). Many species inhabit soil and other terrestrial habitats, where they perform important functional processes in ecosystems such as nitrogen fixation in N cycle (Latysheva *et al.*, 2012). Cyanobacteria are also dominant phototropic component of biota and major primary producers in Arctic and Antarctic regions (Comte, 2007) for example in Svalbard archipelago (Norway). The most common groups in these regions are Oscillatoriales and Nostocales, with some Chroococcales (Garcia-Pichel, 1999).

The most striking characteristics of polar region is extremely low temperature resulting in snow cover, ice bergs and restricted vegetations (Convey *et al.*, 2018). The contemporary climate of Svalbard is modulated by the North Altantic Drift, derived from the Gulf Stream in the Atlantic. Annual air temperature is about 0°C and positive diurnal values are only reached for few summer months or weeks, and even not at all at the most extreme locations (Convey *et al.*, 2018). February is the coldest month with a mean sea level air temperature of −15.2°C, and the warmest month is July at 4-6°C (Ingólfsson, 2004). Precipitation is low, only about 200 mm in central Spitsbergen and somewhat higher along both the western and eastern coasts of the island, about 400-600 mm (Ingólfsson, 2004). February to March and August to September are

comparatively humid (Ingólfsson, 2004). Precipitation is dominantly in the form of snow, with December and January receiving heavy snowfall in some years. Winter seasons are very windy, while fog is common during the summer in coastal areas. Long periods of calm and sunny weather can be experienced in April and May (Ingólfsson, 2004). Lying above the Arctic Circle, the sun remains above the horizon continuously for several months in summer, and conversely below it in winter (Marziali *et al.*, 2009). Despite of these extreme conditions, cyanobacteria show remarkable adaptation to prolonged freezing (Sabacká, 2006) and some are capable of active metabolism at temperatures as low as -20°C (Vincent, 2000a). The most common cyanobacterial groups are filamentous types that occur in various polar freshwater and terrestrial habitats (Table 2.1).

In general, high latitude and altitude cyanobacteria tend to be cold-tolerant (psychrotrophs), with suboptimal growth under low temperatures, rather than psychrophiles that grow optimally at low temperature (Tang and Vincen,t 1999). They have a variety of mechanisms that allows them to tolerate and continue to grow, in the cold and to tolerate freeze–thaw conditions (Vincent, 2007). To maintain membrane fluidity at low temperatures, polyunsaturated fatty acids with decreased chain-lengths are incorporated into the membrane. In addition, the production of compatible solutes (e.g., trehalose) helps to reduce the freezing point of the intracellular fluid. This strategy also reduces cell desiccation as less water is needed to retain the osmotic equilibrium (Welsh, 2000). Furthermore, extracellular compounds such as polymeric substances can reduce ice nucleation around the cells (Vincent, 2007). Generally, Antarctic region shows high cyanobacteria diversity than Arctic region (Zakhia *et al.*, 2008). Most available studies have so far focused on the Canadian Arctic, whereas no information is yet available from the Russian Arctic (Zakhia *et al.*, 2008).

Table 2.1. Few representatives of cyanobacterial taxa reported from polar regions.

Species	Location	Reference
Aphanocapsa cf. *hyalina*	Canadian High Arctic	Jungblut *et al.*, 2010
Calothrix sp.	Mac.Robertson Land, Antarctica	Broady, 1981
Chroococcus cf. *prescotii*	Canadian High Arctic	Jungblut *et al.*, 2010
Chroococcidiopsis sp.	Eastern Antarctica	Wood *et al.*, 2008
Acaryochloris marina	Ross Sea coast, Antarctica	De Rios *et al.*, 2007
Cylindrospermum	Svalbard island, Arctic	Kleinteich *et al.*, 2018
Dichothrix sp.	Canadian High Arctic	Jungblut *et al.*, 2010
Gloecapsa sp.	Ross Sea coast, Antarctica	De Rios *et al.*, 2007
Leptolyngbya sp.	Antarctica	Komàrek and Komàrek, 1999
Cyanothece	McMurdo Dry Valleys, Antarctica	De Rios *et al.*, 2004
Microcystis sp.	Svalbard island, Arctic	Kleinteich *et al.*, 2018
Myxosarcina	Mac.Robertson Land, Antarctica	Broady, 1981
Nostoc sp.	McMurdo Dry Valleys, Antarctica	Taton *et al.*, 2003
Nodularia cf. harveyana	McMurdo Dry Valleys, Antarctica	Taton *et al.*, 2003
Oscillatoria sp.	High Arctic	Mueller *et al.*, 2004
Pseudanabaena sp.	Canadian High Arctic	Jungblut *et al.*, 2010
Phormidium sp.	King George Island Antarctica	Elster, 2002
Schizothrix sp.	McMurdo Dry Valleys, Antarctica	Taton *et al.*, 2003
Synechocystis sp.	McMurdo Dry Valleys, Antarctica	De Rios *et al.*, 2004
Synechococcus sp.	Ross Sea coast, Antarctica	De Rios *et al.*, 2007
Tolypothrix sp.	Svalbard island, Arctic	Kleinteich *et al.*, 2018

Cyanobacteria are also commonly found in temperate regions, where their communities show important seasonal variations (Whitton and Potts, 2007). They are able to survive in the average monthly temperatures of temperate regions during winter range from 0.5 to -4°C. The longest daylength occurs in June with 16 hours (approx.), during summer. Cyanobacteria are widely reported in many temperate ecosystems (Arango *et al.*, 2009). Some of the genera are presented in Table 2.2.

Table 2.2. Few representatives of cyanobacterial taxa reported from temperate regions.

Species	Location	Reference
Anabaena sp.	River Swale, England	Whitton and Potts, 2007
Anabaena bergii	Northeast Germany	Mehnert *et al.*, 2010
Anabaena macrospora	Northeast Germany	Mehnert *et al.*, 2010
Aphanizomenon gracile	Northeast Germany	Mehnert *et al.*, 2010
Aphanizomenon flos-aquae	Northeast Germany	Mehnert *et al.*, 2010
Calothrix sp.	River Swale, England	Whitton and Potts, 2007
Chamaesiphon sp.	River Swale, England	Whitton and Potts, 2007
Cylindrospermopsis raciborskii	Germany	Wiedner *et al.*, 2007
Cylindrospermopsis sp.	France	Sukenik *et al.*, 2012
Cylindrospermopsis sp.	Beijing, china	Xie *et al.*, 2018
Lyngbya sp.	Lake Okeechobee, Florida, USA	Paerl *et al.*, 2011
Microcystis sp.	Lake Okeechobee, Florida, USA	Paerl *et al.*, 2011
Pseudanabaena sp.	North America	Kling and Watson, 2003
Cylindrospermopsis raciborskii	Ukraine	Rzymski *et al.*, 2018

Tropical ecosystems are also rich in specific cyanobacterial flora where temperature most probably is the major factor limiting their geographic distribution (Hoffman, 1999). The growth of cyanobacteria is also influenced by light, climate and nutrient availability (Carrick and Steinman, 2001). Tropical ecosystems are characterized by tropical rainforest weather that includes uniform temperature, high humidity and abundant rainfall. The average annual temperature is approximately 27°C (Hock, 2007). Daylength varies little over the year and is close to 12 h. Tropical climates also provide all the necessary conditions to support cyanobacterial blooms, and few representative bloom-forming cyanobacteria are indicated in Table 2.3.

Table 2.3. Few cyanobacterial species reported from tropical regions.

Species	Location	Reference
Cylindrospermopsis raciborskii *Aphanizomenon* sp. *Anabaena* sp. *Anabaena tenericaulis*, *Microcystis panniformis*	Australia	McGregor and Fabbro, 2000 Saker and Griffiths, 2001 White *et al.*, 2003
Anabaena flos-aquae *Planktothrix* sp.	Bangladesh	Ahmed *et al.*, 2008 Jahan *et al.*, 2010
Leptolyngbya sp. *Oscillatoria tenuis* *Scytonema evanescens* *Spirulina maior* *Phormidium* sp. *Jaaginema* sp.	Mexico	Rejmankova and Komarkov'a', 2000
Microcystis sp. *Cylindrospermopsis raciborskii* *Planktothrix agardhii*	Malaysia	Harith and Hassan, 2011 Mansoor *et al.*, 2011 Wan Maznah and Makhlough, 2015

Anabaena, Oscillatoria, Nostoc, Chroococcus, Pseudanbaena amphigranulata		Khan *et al.*, 2018
Oscillatoria perornata	Vietnam	Nguyen *et al.*, 2007
Cylindrospemopsis sp. *Anabaena* sp. *Microcystis* sp. *Aphanizomenon flos-aquae*, *Oscillatoria limnetica* *Anabaena spiroides*	Nigeria	Anadu *et al.*, 1990 Kemdirim, 2000 Ezra and Nwankwo, 2001 Odokuma and Isirima, 2007 Chia *et al.*, 2009 Ajuzie, 2012
Arthrospira fusiformis *Anabaenopsis abijatae*	Kenya	Ballot *et al.*, 2005 Haande *et al.*, 2007
Anabaena sp. *Chrococcus sp.* *Cylindrospermum sp.* *Gloeocapsa gelatinosa* *Hapalosiphon fontinalis* *Nostoc muscorum,* *Oscillatoria* sp. *Plectonema* sp. *Pseudanabaena* sp.	India	Hemlata and Fatma, 2009 Mishra *et al.*, 2012

2.3 History of cyanobacterial taxonomy and classification approaches

Traditionally, cyanobacterial classification was based on morphological features such as cell dimensions, shape, trichome/filament length, shape of apical cell, presence/absence of polar aerotopes described under light microscope (Komárek, 2003). Morphologically similar specimens were classified as the same species, and this approach resulted in the identification of more than 2000 morphospecies in 150 genera (Waterbury, 2006).

Geitler (1925) based on morphological feature, first proposed the erection of the orders Chroococcales, Entophysalidales, Pleurocapsales, Dermocarpales, Siphononematales, Nostocales and Stigonematales within the Cyanobacteria. But later he adopted the simpler classification suggested by Frémy (1929), that included only three orders, Chroococcales, Chamaesiphonales and Hormogonales (Geitler 1932). Ten years later, Geitler (1942) removed Chamaesiphonales and included Dermocarpales and Pleurocapsales (Geitler, 1942), giving a four order classification system still in use 24 years later with only minor amendments (Starmach, 1966). Desikachary (1959) reviewed the order Stigonematales and grouped non-branching filamentous taxa into one order, Nostocales. An alternative approach was introduced by Drouet (1968), who believed that morphospecies were the environmentally-induced varieties of a limited number of genotypes. This approach significantly reduced the number of species and genera (Castenholz, 2001) and led to the loss of valuable ecological information (Whitton, 2008).

Stanier *et al.* (1978) became convinced that cyanobacteria should be treated like bacteria and introduced the bacteriological approach to the classification of cyanobacteria. This approach was based on morphological, physiological and genetic characteristics of cultured and axenic strains. This proposal caused a conflict with the

botanical nomenclature, as cyanobacteria (Cyanophyta, blue-green algae) are also named under the International Code of Botanical Nomenclature (ICBN) (Oren, 2004). This latter approach provides additional information that is useful for identification purposes (Whitton, 2011), but has limited utility in ecological studies. Subsequently, the five sections system recommended by Rippka *et al.*, (1979) (Table 2.4) became the primary basis for the cyanoabacterial classification in Bergey's Manual of Systematic Bacteriology, which recognized five subsections instead of orders, I (= Chroococcales), II (= Pleurocapsales), III (= Oscillatoriales), IV (= Nostocales) and V (= Stigonematales) (Komárek et al., 2014).

Table 2.4. Classification of cyanobacteria proposed by Rippka *et al.*, (1979)

S.No.	Section	Cell morphology	Examples
I	Chroococcales	Unicellular, isopolar	*Synechoccous, Synechocystis*
II	Pleurocapsales	Pseudoparenchymatous	*Dermocarpa, Xenococcus*
III	Oscillatoriales	Multicellular, trichal, heterocysts not present	*Oscillatoria Spirulina, Phormidium*
IV	Nostocales	Multicellular, trichal, heterocysts present	*Nostoc, Anabaena, Scytonema*
V	Stigonematales	Multicellular, trichal, with branches, heterocysts present	*Chlogloeopsis, Fischerella, Mastigocladus*

The traditional taxonomic classification of cyanobacteria was based only on morphological and, more rarely, ecological features, but morphology alone appears to be an insufficient tool in modern taxonomy (Komárek, 2006). Therefore, a combination of traditional and bacteriological approaches as described by Anagnostidis and Komárek (1988, 1990) has been increasingly adopted. This is essentially an elaboration of the classical system with reassessment of various important characters, and also incorporating new information, to confidently identify different taxa. This 'polyphasic approach' combines morphological, molecular, ultrastructural and ecological information, and has been widely used in recent cyanobacterial studies (e.g. Ballot *et al.*, 2008; Zapomělová *et al.*, 2009; 2010; Heath *et al.*, 2010; Dadheech *et al.*, 2012; Sciuto *et al.*, 2012; Yu *et al.*, 2015).

Hoffmann *et al.* (2005a, b) divided the class Cyanophyceae into four subclasses: Gloeobacteriophycidae, Synechococcophycidae, Oscillatoriophycidae and Nostochopycidae. This development reflects the group's phylogeny and is drastically different to the previous system of classification. Komárek *et al.* (2014) advocated a further development in the taxonomic classification system, incorporating a polyphasic approach that reflects evolutionary history and includes monophyletic taxa. They divided cyanobacteria into eight orders as shown in Table 2.5.

Table 2.5. Most recent cyanobacterial classification system for Orders and Families (Komárek *et al.*, 2014).

S.No.	Order	Family
I	Gloeobacterales	*Gloeobacteraceae*
II	Synechococcales	*Synechococcaceae, Merismopediaceae, Prochloraceae* *Coelosphaeriaceae, Acaryochloridaceae, Chamaesiphonaceae, Romeriaceae,* *Pseudanabaenaceae, Leptolyngbyaceae, Heteroleibleiniaceae, Schizotrichaceae*
III	Spirulinales	*Spirulinaceae*
IV	Chroococcales	*Microcystaceae, Aphanothecaceae, Cyanobacteriaceae, Cyanothrichaceae,* *Stichosiphonaceae, Chroococcaceae, Gomphosphaeriaceae, Entophysalidaceae*
V	Pleurocapsales	*Hydrococcaceae, Dermocarpellaceae, Xenococcaceae, Pleurocapsaceae,*
VI	Oscillatoriales	*Cyanothecaceae, Borziaceae, Coleofasciculaceae, Microcoleaceae, Homoeotrichaceae,* *Oscillatoriaceae, Gomontiellaceae*
VII	Chroococcidiopsidales	*Chroococcidiopsidaceae*
VIII	Nostocales	*Scytonemataceae, Symphyonemataceae, Rivulariaceae,* *Tolypothrichaceae, Godleyaceae, Chlorogloeopsidaceae, Hapalosiphonaceae,* *Capsosiraceae, Stigonemataceae, Gloeotrichiaceae,* *Aphanizomenonaceae, Nostocaceae,*

Progressively more complex analytical methods and approaches such as electron microscopy and molecular analyses have become crucial tools in the characterization of cyanoprokaryotes and assessing their evolutionary relationships. These approaches commonly involve the sequencing of various genes. In bacteria, comparison of 16S rDNA sequences potentially allows recognition of different genera (Clarridge, 2004), and this approach can also be used for identification and classification of cyanobacteria (Castenholz, 2001). The development of small ribosomal subunit gene sequencing has also become a keystone in unravelling phylogenetic relationships in microbes (Webb *et al.*, 2002). SSU rRNA's are universal molecules that comprise both highly conserved and rapidly evolving regions. 16S rDNA is one of the three genes that form ribosomal RNA (rRNA) operons in bacteria (Iteman *et al.*, 2000). It consists of nearly 1550 nucleotides and provides a large number of characters that allow statistically valid evaluations between taxa (Clarridge, 2004). This is assumed to be related to its critical function in protein synthesis (Clarridge, 2004). Nubel *et al.* (1997) developed and tested a set of oligonucleotide primers for the specific amplification of the 16S rRNA gene from cyanobacteria, using it to investigate cyanobacterial diversity in cultures, lichens, and complex microbial communities. Association between morphology and 16S rDNA phylogenies has been recorded in heterocytous cyanobacteria (Tomitani *et al.*, 2006) as well as in some genera of Oscillatoriales, including *Planktothrix* (Komárek, 2006), *Oscillatoria* (Luke *et al.*, 2008), *Trichodesmium* (Wilmotte and Herdman, 2001) and *Arthrospira* (Komárek, 2006).

However, the use of 16S rRNA gene also has drawbacks. Due to its high level of sequence conservation, its use in the determination of closely-related species is limited even when species are morphologically distinct (Rosselló-Mora and Amann,

2001). In contrast, in some cases extremely high intraspecific 16S rDNA diversity has also been documented (Lokmer, 2007). Although threshold values of 16S rDNA similarity required to support species delimitation in bacteria are not clear (Coenye *et al.*, 2005), taxa with more than 98% 16S rDNA sequence similarity are commonly considered to be the same species (cf. Stackebrandt, 2011).

To obtain better resolution for closely related taxa, many studies have turned to analyses of the 16S-23S rRNA internal transcribed spacer (ITS) region (Boyer *et al.*, 2001; Brown *et al.*, 2005; Iteman *et al.*, 2000; Novis and Visnovsky, 2011). The ITS region is another part of the rRNA operon and has variable sequence, length and secondary structure, sometimes also having multiple copies within a single genome (Gugger *et al.*, 2002). Three types of 16S-23S ITS sequence configurations have been reported in cyanobacteria (Boyer *et al.*, 2001). The most common type contains both $tRNA^{Ile}$ and $tRNA^{Ala}$ (Boyer *et al.*, 2001), but there are also several examples of 16S-23S ITS with either only the $tRNA^{Ile}$ gene (Novis and Visnovsky, 2011) or completely lacking tRNA genes (Iteman *et al.* 2000, Boyer *et al.* 2001, Taton *et al.*, 2003).

16S-23S ITS sequence has been successfully utilized to differentiate the closely related taxa within the genera *Arthrospira* (Ballot *et al.*, 2004), *Phormidium* (Comte *et al.* 2007), *Synechococcus* (Becker *et al.* 2004), *Phormidium* (Novis and Visnovsky, 2011), *Anabaena* (Gugger *et al.*, 2002), *Nostoc* (Iteman *et al.*, 2000), *Calothrix* (Boyer *et al.*, 2001) contains both $tRNA^{Ile}$ and $tRNA^{Ala}$. Several examples containing only $tRNA^{Ile}$ include *Microcystis* (Novis and Visnovsky, 2011), *Synechococcus* (Novis and Visnovsky, 2011) and *Spirulina* (Novis and Visnovsky, 2011). Whereas another composition shows a lack of tRNA genes and has been reported in members of *Nostoc* (Iteman *et al.*, 2000), *Calothrix* and *Scytonema* (Boyer *et al.*, 2001).

Moreover, the existence of unicellular cyanobacterial ecotypes in different microenvironments of hot springs has also been revealed by ITS sequencing (Ward *et al.*, 2006). The high degree of divergence typically within ITS sequences makes phylogenetic analyses possible for closely related strains (Taton *et al.* 2006), and the sequence configuration of ITS can be a valuable tool in the understanding of population structure of cyanobacteria and in high-level phylogeny studies (Boyer *et al.*, 2001).

Microbial whole genome sequencing, developed in the 1990s, is a comprehensive method of mapping genomes of novel organisms, finishing genomes of known organisms, or comparing genomes across multiple samples (Bruce *et al.*, 2008). *Haemophilus influenzae* was the first organism to have its entire genome sequenced (Fleischmann *et al.*, 1995). The genomes of other bacteria and some archaea soon followed, largely due to their small genome size. Many tools have since been developed enabling the sequencing of whole genomes. The first generation of such tools were basically sanger sequencing (Sanger *et al.*, 1977), but the completion of the draft human genome in 2001 (Lander *et al.*, 2001) prompted the development of second generation or next-generation sequencing (NGS) technologies, to provide cheaper, higher throughput and more reliable sequencing. Of the major players in the NGS market, Illumina is widely used

Illumina sequencing technology is a proprietary reversible terminator technology for rapid and accurate large-scale sequencing. This innovative and flexible sequencing system enables a broad array of applications in genomics, transcriptomics, and epigenomics (Loman *et al.*, 2012).

Genome sequencing using different technologies varies in the length of reads produced. Most current protocols (e.g. Illumina) produce read lengths in the range of 100-500 bp (Jünemann *et al.*, 2013). Pacific Biosciences platforms produce approximately 1500 bp long reads (Jünemann *et al.*, 2013). Longer read length improves the resolution of *de novo* genome assembly and also detection of structural variants (Chhangawala *et al.*, 2015). However, the currently high cost per base prohibits large-scale sequencing using Pacific Biosciences platforms (Quail *et al.*, 2012). Illumina DNA sequencing technologies have emerged as a cost-effective and convenient approach which allows researchers to characterize a bacterial genome to address many microbiological questions (Didelot *et al.*, 2012).

The occurrence of horizontal gene transfer (HGT) can be identified by whole genome studies. Horizontal (or lateral) gene transfer, potentially followed by recombination, is now documented as a key force shaping evolutionary histories of both prokaryotes (Zhaxybayeva *et al.*, 2006) and eukaryotes (Mitreva *et al.*, 2005). A high degree of HGT has been reported in cyanobacteria (Lodders *et al.*, 2005), making it more difficult to construct meaningful phylogenies. However, there is a core of genes that remain closely associated and resistant to HGT, possibly permitting separation of true phylogenetic signals from 'noise' (Shi and Falkowski, 2008).

The use of modern molecular, cytomorphological and ecological methods in cyanobacterial taxonomic studies is necessary and is recommended as the only method for the elaboration of their modern systematics (Komárek, 2016). Modern methods of taxonomy must include molecular sequencing as the basic approach, and other criteria including morphological and ecological (if available) should be added when they provide distinct and recognizable information (Komárek, 2016). Application of this polyphasic approach is an exclusive, modern, unambiguous, and a fully suitable

methodological procedure, but it is not yet widely used in cyanobacterial studies. A further complication is that working with cultivated strains is challenging for taxonomic studies, because long cultivation of cyanobacteria under unified conditions can lead to changes, often resulting in modified morphology and physiology and subsequent loss of important ecological features (Komárek & Kaštovský, 2003). Polyphasic approach has been utilised to identify many cyanobacterial genus for example *Anabaena circinalis*, *Anabaena crassa* (Zapomelova *et al.*, 2008), *Cylindrospermopsis raciborskii* (Soares *et al.*, 2013), *Pseudanabaena mucicola* (Yu *et al.*, 2015), *Pseudanabaena amphigranulata* (Khan *et al.*, 2018) etc.

2.4 Cell structure of cyanobacteria

The cell structure and organisation of cyanobacteria are studied with the aid of light and electron microscopes.

2.4.1 Basic Morphology

The basic morphology of cyanobacteria includes unicellular, colonial and multicellular filamentous forms (Mur *et al.*, 1999). Unicellular forms, for example members of order *Chroococcales*, have spherical, ovoid or cylindrical cells. They occur singly when the daughter cells separate after reproduction by binary fission. The cells may form irregular colonies, becoming embedded in the slimy mucilaginous matrix secreted during the growth of the colony (Mur *et al.*, 1999). Members of the order *Chamaesiphonales* produce exospores, budded off from the upper ends of cells (Rippka *et al.*, 1979). In the order *Pleurocapsales*, the principal mode of replication is by a series of successive binary fission events that convert a single mother cell into many minute daughter cells (baeocytes or endospores). Filamentous morphology results due to repeated cell divisions occurring in a single plane at right angles to the

main axis of the filament. The multicellular structure consisting of a uniseriate row of cells is called a trichome. Cell size and shape varies greatly among the filamentous forms of cyanobacteria. Species in the order *Oscillatoriales* possess unbranched trichomes (Komárek and Anagnostidis, 2005). The other order with a filamentous organisation (order *Nostocales*) consist of trichomes with heterogenous cells. Trichomes are composed of two types of vegetative cells, heterocysts (with thick wall and hyaline protoplast, capable of nitrogen fixation) and akinetes (large thick-walled cells, containing reserve materials that enable their survival under unfavourable conditions). Members of the family *Stigonemataceae* in the order *Nostocales* (Komárek *et al.*, 2014) are characterised by multiseriated and branched filaments. Heterocysts and akinetes are also present (Komárek and Anagnostidis, 2005).

2.4.2 Cell Structure

Many monographs and reviews discuss the detailed structural and functional organization of cyanobacteria (Hoiczyk and Hansel 2000; Herrero and Flores 2008). The structural composition of a cyanobacterial vegetative cell is illustrated in Figure 2.1. The cells are Gram-negative, surrounded by a cell wall comprising a peptidoglycan layer and the outer membrane. On the cell surface, sheaths, sometimes capsules, paracrystalline protein or glycoprotein S-layers, pili and spines can be found (Baulina, 2012). Thylakoids, responsible for energy supply, fill most of the cytoplasm. These membrane structures form a single system, termed as thylakoid networks (Nevo *et al.* 2009). The thylakoids are lamellae formed by paired membranes containing chlorophyll *a* as a component of the photosystems. Phycobilisomes are found in rows on the outer surface of the thylakoids (Douglas, 1994). The nucleoid with (poly)ribosome is located preferentially in the central part of the cell (Baulina, 2012).

Cyanobacteria have a remarkable capacity to store various structures and inclusions within the cells, including: gas vesicles that promote buoyancy of the cell within the water column; carboxysomes, containing the CO_2 fixation catalysing enzyme, ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO); cyanophycin granules that provide an alternative nitrogen source; lipid granules and granules of poly-b-hydroxybutyrate which act as sources of carbon and energy; and polyphosphate granules as sources of phosphorus (Gorelova *et al.*, 1996; Gusev *et al.*, 2002; Baulina, 2012).