This electronic thesis or dissertation has been downloaded from Explore Bristol Research, http://research-information.bristol.ac.uk

Author:
Druce, Elliot J

Title:
Freshwater Picocyanobacteria and their Organic Diet
An Investigation into Organic Nitrogen Assimilation in Freshwater Picocyanobacteria

General rights
Access to the thesis is subject to the Creative Commons Attribution - NonCommercial-No Derivatives 4.0 International Public License. A copy of this may be found at https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode This license sets out your rights and the restrictions that apply to your access to the thesis so it is important you read this before proceeding.

Take down policy
Some pages of this thesis may have been removed for copyright restrictions prior to having it been deposited in Explore Bristol Research. However, if you have discovered material within the thesis that you consider to be unlawful e.g. breaches of copyright (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please contact collections-metadata@bristol.ac.uk and include the following information in your message:

- Your contact details
- Bibliographic details for the item, including a URL
- An outline nature of the complaint

Your claim will be investigated and, where appropriate, the item in question will be removed from public view as soon as possible.

Freshwater Picocyanobacteria and their Organic Diet

An Investigation into Organic Nitrogen Assimilation in Freshwater Picocyanobacteria

By

ELLIOT DRUCE

School of Geographical Sciences
UNIVERSITY OF BRISTOL

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of DOCTOR OF PHILOSOPHY in the Faculty of Science.

JUNE 2023

Word count: 38,279

Abstract

Freshwater picocyanobacteria are an essential component of our waterbodies, fulfilling a role of global primary producers in waters which are often limited or co-limited by nitrogen. However, they remain relatively poorly understood, especially in comparison to their marine relatives. Also poorly understood is the picocyanobacterial metabolism of organic sources of nitrogen. Research has traditionally focused on inorganic sources such as ammonium or nitrate, however it is becoming increasingly evident that organic nitrogen is bioavailable and an integral component of the freshwater nitrogen pool. In this thesis, we sequence five strains of freshwater picocyanobacteria and compare their encoding capabilities to a common *Synechococcus* model cyanobacterium (*Synechococcus elongatus*). We increased the representation of freshwater picocyanobacteria, and found reduced encoding of nitrite-associated assimilation genes in picocyanobacteria. Further differences in encoded antennae proteins highlight the differences between freshwater picocyanobacteria (*Synechococcus* spp. of the *Syn/Pro* clade) and *Synechococcus elongatus*, highlighting the need for a 'true' freshwater picocyanobacterium model organism. Utilising comparative genomic analyses we investigated the nitrogen assimilation capabilities of freshwater picocyanobacteria, comparing these to picocyanobacteria of different habitats in addition to larger freshwater cyanobacteria. The diversity displayed among nitrogen assimilation capabilities reveals the evolutionary history of the picocyanobacteria, showcasing the environments in which these organisms evolved. These analyses revealed significant variation in the encoded amino acid transporters between freshwater picocyanobacteria and larger freshwater cyanobacteria. With different amino acid transporters having different amino acid preferences and uptake rates, this may have implications for amino acid bioavailability. Other sources of potentially differentially bioavailable nitrogen include novel organic forms, such as chitin and glyphosate. Finally, we conduct a growth assay to determine the assimilation capabilities of amino acids as an organic nitrogen source, and explore the proteomic response to growth on selected amino acids using quantitative proteomic analysis. The growth assay revealed widespread amino acid assimilation as a sole nitrogen source for freshwater picocyanobacteria, enabling a wide N pool which may contribute to their limnetic dominance. However, proteomic analysis revealed a subtle stress response in freshwater picocyanobacteria when grown on selected amino acids, potentially due to the accumulation of metabolites. Together, the work increases our understanding of the nitrogen assimilation capabilities of freshwater picocyanobacteria, with a focus on organic sources of nitrogen.

COVID Statement

The COVID pandemic, amid lockdowns and travel uncertainty, impacted this thesis. This was most keenly felt in the absence of field-work and mesocosm-based experiments. Prior to the COVID lockdowns, preparation and planning was underway to carry out fieldwork in the Lake District to investigate associations between DON composition and the microbial community. In addition, discussions were underway with the CEH Aquatic Mesocosm Facility to carry out amino acid-spiking investigations. However, the restrictions implemented due to COVID, caused by continual and unpredictable lockdowns, necessitated a change in thesis direction towards carrying out lab experiments in the Life Sciences Building at the University of Bristol, which reduced the risk of work interruption. Furthermore, the delays caused by COVID and associated loss of time reduced my capability to cope with equipment delays and failure. I received a grant from the National Environmental Isotope Facility to carry out work investigating the uptake and incorporation of ^{13}C,^{15}N-labelled amino acids in freshwater picocyanobacteria, however extended equipment failure and my approaching thesis deadline meant that the processed samples were not able to be analysed.

Dedication and acknowledgements

This thesis is dedicated to my nephew, Ezra Phillips. Although it will be many years before you know the meaning of half the words in here, I hope it encourages you to follow your interests, have a love for the natural world, and to hang on in there when things get tough.

My supervisors have supported me throughout this experience. Thank you Patricia for all the guidance and encouragement, for never letting me forget to look up from the details and see the bigger picture. You always gave me confidence in my abilities and this would not have happened without your mentorship. Likewise, thank you Stephen for the advice and ecological expertise you were always so happy to share. I'd also like to thank Penny for your enthusiasm in the project and inspiration for its direction. The Sánchez-Baracaldo lab has been star-studded during my time here. Thank you Giorgio and Jo for showing me that this was possible, and for all the help along the way.

A huge thank you goes to Annie. You came into my life and changed it forever. You always kept me sane when times got hard and reminded me to enjoy the small things, all the while being a consistent companion in the hills throughout the years. The same applies to Rebecca. Thank you for putting up with me throughout all this and for always putting a smile on my face, whether wanted or not. I can't express sufficiently how grateful I am for you. Another huge thank you goes to Vicki. We've been on this journey together from the start and I wouldn't have made it past the first year without our hours and hours of chats (and our shared email writing sessions). Of course none of this would have been possible without the support of my family. Thank you Mum, Dad, and Sara for supporting me so much. Your pride is obvious every time I visit, most evident by the fact that phytoplankton has become part of your lexicon. I'm sure Grandmas Lily and Audrey, and Grandpas Len and Stanley, would be delighted to know that there's another Doctor in the family.

Finally, I would like to thank all my friends who have lent an ear, a distraction, or a supporting shoulder throughout this period. I extend much thanks to my friends in Browns, especially Emily, Saskia, Ben, and Isolde. Also my climbing friends - Toby, Sam, Katie, and Zoonii; you helped make this activity a great outlet for a bad day. Lastly, I would like to thank the friends who have been with me through thick and thin, and no matter the distance remain true confidants. Michael Carr and Jonathan (Bonno) Abrahams, for listening to me moan for five years straight; Ziyin Wang, for being my de-facto therapist; and Jonathan Simmons and Sam Specterman, for all our gaming evenings.

A PhD can never be a solo endeavour; I'm glad you were around to help me through.

Author's declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: .... DATE: ....

Table of Contents

List of Tables

List of Figures

Abbreviations

2-OG	2-oxoglutarate
AP	Allophycocyanin
AA	Amino Acid
AAT	Amino Acid Transporter
ABC	ATP-Binding Cassette
AGC	Automatic Gain Control
C	Carbon
DCAA	Dissolved Combined Amino Acid
DEP	Differentially Expressed Protein
DFAA	Dissolved Free Amino Acid
DIN	Dissolved Inorganic Nitrogen
DON	Dissolved Organic Nitrogen
FACS	Fluorescent-Activated Cell Sorting
F-Pcy	Freshwater Picocyanobacteria
GOGAT	Glutamate Synthase
GS	Glutamine Synthetase
HMW	High Molecular Weight
LF-Cy	Larger Freshwater Cyanobacteria
LMW	Low Molecular Weight
M-Pcy	Marine Picocyanobacteria
N	Nitrogen
NH_3	Ammonia
NH_4^+	Ammonium
NO_3^-	Nitrate
NO_2^-	Nitrite
P	Phosphorus
PBS	Phycobilisome
PC	Phycocyanin
PE	Phycoerythrin

OD	Optical Density
SB	Solvent B
SCO	Single-Copy Ortholog
SE-Cy	*Synechococcus elongatus* Strains
T-Pcy	Thermal Picocyanobacteria
TDAA	Total Dissolved Amino Acids
TMT	Tandem Mass Tag

Chapter 1

Freshwater Picocyanobacteria and Their Nitrogen Diet

1.1 Contributions and Acknowledgements

This chapter was written with feedback from Patricia Sánchez-Baracaldo and Stephen Maberly.

1.2 Introduction

Picocyanobacteria comprise the smallest morphological genera of Cyanobacteria and are the most globally abundant photosynthetic organisms [1]. They are the dominating primary producer in marine and freshwater oligotrophic environments [2–4], while also abundant in more productive waters [5]. Nitrogen (N) can be a limiting or co-limiting nutrient in these environments [6]; as such the concentration and speciation of N is of great interest to understand the proliferation of picocyanobacteria [7–9]. The impact of traditional forms of N (such as nitrate (NO_3^-) and ammonium (NH_4^+)) on picocyanobacteria has been extensively studied, though ecological and genomic research has often focused on marine systems [10, 11]. In contrast, the extent of bioavailable N speciation and assimilatory mechanisms in their freshwater relatives is less understood, with a growing awareness of the bioavailability of organic N sources further complicating our understanding of how freshwater picocyanobacteria respond to their nutrient pool [12–14].

This thesis aims to increase our knowledge of the N assimilation capabilities of freshwater picocyanobacteria with a focus on organic sources of N. A foundation for these investigations is provided by this first chapter, introducing picocyanobacteria and their ecological importance in marine and fresh waters, the current understanding of N assimilation capabilities, and a broad overview of amino acid (AA) biosynthesis and degradation pathways in cyanobacteria.

1.3 The Ecological Significance of Picocyanobacteria

Picocyanobacteria are coccoid or bacillus-shaped organisms of the phylum Cyanobacteria found commonly as single cells [15]. Phenotypically characterised by a very small cell size ranging from 0.2 to 2 µm, picocyanobacteria constitute the smallest group of the Cyanobacteria yet have a large impact on global aquatic ecosystems [16]. They are able to thrive in a wide range of habitats, from the ice-covered lakes of the High Arctic [17] to the marine tropics [18], abundant in highly eutrophic [19] and oligotrophic lakes [20], and found in riverine systems [21]. As autotrophic primary producers, picocyanobacteria provide fixed carbon (C) to their ecosystem via photosynthesis, with significant picocyanobacterial contributions to total C common in non-eutrophic environments. In oligotrophic Lake Baikal, 81% of C fixation is carried out by picoplankton [22], while picocyanobacteria account for up to 54% of total photosynthesis in Canadian pre-Alpine lakes [23]. Changes in the picocyanobacteria biomass at the base of the food chain can have a cascading impact on organisms which are magnitudes larger. With fixed C transferred from the picocyanobacteria primary producers up multiple trophic levels, alterations in picocyanobacterial abundance and community structure can result in ecosystem shifts [24, 25]. The focus of this thesis is freshwater picocyanobacteria, however selected information about marine picocyanobacteria is included in this introduction to provide a comparison between picocyanobacteria from the two environments.

1.3.1 Picocyanobacteria and Trophic State

The trophic state of an aquatic habitat will greatly influence its picocyanobacterial population [4, 26]. Though picocyanobacteria are widely prevalent in both freshwater and marine environments, the biomass of picocyanobacteria relative to the total phytoplankton biomass can vary, as can the picocyanobacteria community structure [27, 28]. In marine systems, a productivity gradient from the polar regions to the oligotrophic tropics reveals variations in dominant picocyanobacteria. *Prochlorococcus* numerically dominates offshore tropical environments, with their distribution largely limited to between latitudes 45°N and 40°S [29]. Heavily-streamlined genomes and a very small size (*Prochlorococcus* spp. are the smallest known non-symbiotic photoautotroph) contribute to this niche success, minimising resource requirements in these oligotrophic environments and maximising nutrient exchange and diffusion rate [30, 31]. As marine waters become gradually more productive, in temperate regions and around coastlines [32], *Prochlorococcus* abundance decreases and is replaced by *Synechococcus*, able to out-compete *Prochlorococcus* away from ultra-oligotrophic environments [31]. In contrast, among freshwater environments (where research has typically focused on temperate lakes), *Synechococcus* is the primary genus of picocyanobacteria, with *Cyanobium* and *Vulcanococcus* spp. also present to a lesser extent (*Prochlorococcus* is absent) [15, 33, 34]. In these environments, *Synechococcus* dominates the picocyanobacteria fraction, yet the relative proportion of picocyanobacteria compared to larger phytoplankton can alter considerably and is largely influenced by the trophic state. Nutrient limitation in oligotrophic lakes favour picocyanobacteria over larger phytoplankton species, with similar patterns also displayed in mesotrophic lakes, resulting in *Synechococcus* comprising the bulk of biomass in these environments [35–38]. This success in oligotrophic environments is commonly associated with the small size of picocyanobacteria, their high surface area to volume ratio supporting faster uptake of limited nutrients [39], hence enabling them to out-compete larger phytoplankton. This trait also results in picocyanobacterial sensitivity to small nutrient changes in their environment, facilitating early indications of nutrient enrichment [35].

As freshwater systems become more eutrophic, competition for nutrients is reduced and larger phytoplankton begin to grow in abundance relative to the picocyanobacteria [40]. This ecological shift in phytoplankton community is affected by the depth of the water body. In most eutrophic lakes, the contribution of picocyanobacteria to overall lake biomass does not exceed 10% due to increased abundance of larger phytoplankton groups such as diatoms, dinoflagellates, and larger cyanobacteria [4, 41]. In these habitats, the proportion of picocyanobacterial biomass to overall primary production is diminished, however their actual contribution as primary producers to higher trophic levels may remain considerable. This is dependent on the community structure of the eutrophic-dominating phytoplankton. Some of these groups, such as filamentous cyanobacteria and colonial diatoms, are inedible, leading to a zooplanktonic shift to smaller species (i.e. rotifers) which feed on smaller picoplankton (including picocyanobacteria) [42]. Lake morphology also impacts picocyanobacterial dominance in eutrophic waters. Though picocyanobacterial contribution to total biomass is generally reduced in eutrophic environments, shallow eutrophic lakes can maintain a large picocyanobacterial influence, especially from Spring to Autumn [43, 44].

Picocyanobacterial abundance in varying freshwater environments, from shallow eutrophic to deep oligotrophic, provides evidence of proliferation supported by differential adaptation of photosynthesis-associated molecular mechanisms. As photoautotrophs, picocyanobacterial growth is dependent on the absorption of light energy by photosynthetic pigments and their subsequent excitation; electrons are transferred to photosynthetic reaction centres and ultimately, in the form of ATP and NADPH, towards the Calvin cycle producing glucose and other metabolic intermediates [45]. However, pigments do not universally absorb light, instead absorbing a range of specific wavelengths across the visible (and occasionally far-red [46]) light spectrum [47]. As a result, the composition of photosynthetic pigments in an organism dictates the distribution of photosynthetically-active radiation, which in turn partly determines organism abundance. This is seen in picocyanobacteria among their light capturing antennae of phycobilisomes (PBSs) composed of up to three pigments, a central core of allophycocyanin with stacked rods of phycocyanin (PC) and (if present) phycoerythrin (PE) [48]. Picocyanobacteria populations in deep oligotrophic lakes are abundant in PE (red phycobiliprotein which absorbs visible light at 540 - 570 nm), while populations of picocyanobacteria in eutrophic shallow lakes lack PE, instead containing rods of PC only (able to absorb light at 610 - 630 nm) [49]. These pigment patterns reflect the wavelengths of light most available to the picocyanobacteria in these differing environments. The abundance of PE-rich picocyanobacteria in clear oligotrophic lakes reflects their ability to absorb green light in deeper water layers, with light able to penetrate due to reduced biomass at the surface. In contrast, turbid eutrophic lakes favour the dominance of PC-rich picocyanobacteria, absorbing red light near the surface [49–51]. This feature of environmental selection pressures dictating the necessary molecular equipment for survival is a key theme in this thesis, and is demonstrated in the molecular machinery found in picocyanobacteria to assimilate N.

1.3.2 Picocyanobacteria and Water Quality

An algal bloom is the rapid proliferation of algae caused by a relaxation of factors limiting growth (e.g. N or phosphorus (P) influx) or a reduction in processes causing biomass loss (e.g. reduced water flow). It requires nutrient levels to be adequate to support the biomass produced and thus is typically found in productive sites, though blooms have also been recorded in oligotrophic systems [69]. These events are regularly associated with cyanobacteria, posing risks to ecosystem health through fast population growth. Picocyanobacteria were originally described as a non-blooming group [70], however it is now clear that picocyanobacterial blooms are a world-wide phenomenon (Figure 1.1). Algal blooms can degrade water quality with excessive cyanobacterial growth restricting water transparency. Reduced light availability subsequently inhibits growth of macrophytes which in turn impacts higher trophic levels [71]. Cyanobacteria blooms also lead to abnormally high contributions of cyanobacteria to primary production. This can negatively impact food webs due to the low concentrations of long chain-essential fatty acids in cyanobacteria compared to other primary producers (e.g. eukaryotic algae) [72]. This disrupts the transfer of long chain-essential fatty acids across trophic levels, negatively impacting

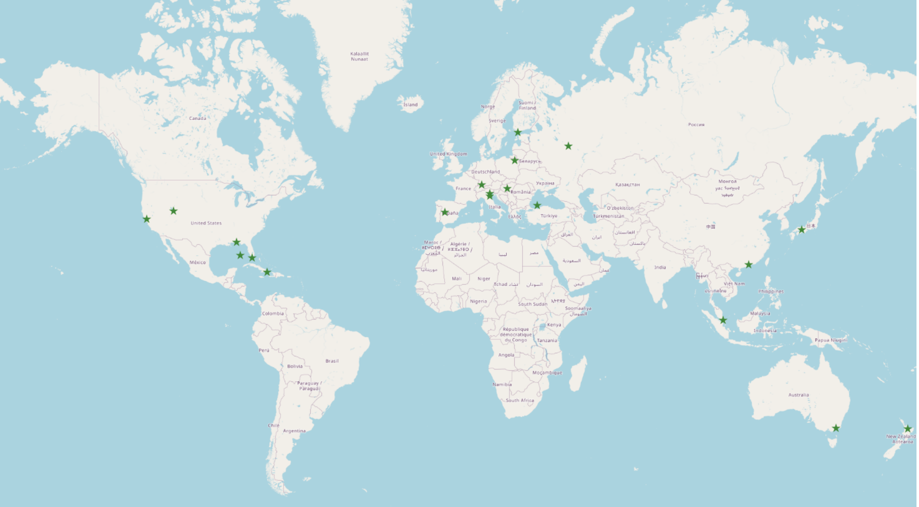

Figure 1.1: **Locations of recorded picocyanobacteria blooms.** Locations include Lake Constance (Germany) [52], Lake Zab-szék (Hungary) [53], Gippsland Lakes (Australia) [54], Gorky Reservoir (Russia) [55], Rosarito Reservoir (Spain) [56], Utah Lake (USA) [5], New Zealand lakes [26], Lake Lazduny (Poland) [57], Venice Lagoon (Italy) [58], Comacchio Lagoon (Italy) [59], Pearl River Estuary (China) [60], San Francisco Bay (USA) [28], Pensacola Bay (USA) [61], Florida Bay (USA) [62], Singapore [63], Guantanamo Bay (Cuba) [64], Seto Inland Sea (Japan) [65], Baltic Sea [66], Black Sea [67], and Gulf of Mexico [68]. Note that this list is not comprehensive.

organisms of higher trophic levels reliant on essential fatty acid uptake [73]. Bloom decay is primarily caused by physical conditions becoming unsuitable (e.g. seasonal shift, reduction of stratification, nutrient depletion) [74]; cyanobacterial die off releases organic matter, promoting bacterial growth which is often coupled with anoxia [75]. The decay of substantial blooms can cause intense localised anoxia, resulting in fish death during extreme conditions [75].

Though rapid cyanobacteria population increase can directly impact the ecosystem through obstructing light and a decrease in nutritional quality, cyanotoxin accumulation is the major public health concern. Cyanotoxins are toxic secondary metabolites produced by some cyanobacterial genera, and are generally split into hepatotoxins, cytotoxins, and neurotoxins [76]. The most ecologically important cyanotoxin is microcystin, a liver-targeting heptatoxin which is ubiquitously found in aquatic environments worldwide [77, 78]. Microcystins are most commonly produced by *Microcystis* spp., but production has also been detected from freshwater and marine picocyanobacteria [79–81]. Population growth leads to an increased concentration of microcystin during algal blooms; associated toxin ingestion mostly impacts pets and livestock due to increased exposure to active algal blooms, with mortality rates increasing [82]. In addition to animal poisoning, human microcystin exposure through insufficient reservoir treatment or contaminated water activities can lead to extensive organ damage, with death recorded in isolated incidents [83]. Freshwater *Synechococcus* strains have also been found to produce cylindrospermopsin (cytotoxin) and anatoxin-a (neurotoxin), highlighting risks to public health when exposed to picocyanobacteria blooms [84]. In addition to harmful toxins, *Synechococcus* can produce 2-methylisoborneol and geosmin [80, 85]. These compounds have a potent odour (akin to mould) which negatively impacts potable water quality and necessitates their costly removal from water sources before becoming suitable for consumer use.

1.3.3 Picocyanobacteria and Global Processes

Global physical processes can influence, and be influenced, by picocyanobacteria. Microbially-induced calcite precipitation involves the formation of calcium carbonate crystals from solution, a vital process in the natural calcification of aquatic environments which can result in large-scale C storage [86]. Freshwater *Synechococcus* spp. in oligotrophic environments can trigger this precipitation, and though the exact mechanism is yet to be elucidated, Kosamu *et al* suggests that the picocyanobacterial cell wall provides a template for the precipitation reaction to occur [87, 88]. While picocyanobacteria impact their environment through calcite precipitation, the environment is directly impacting picocyanobacteria through a changing climate. Increasing global temperatures will promote picocyanobacteria growth, with *Synechococcus* growth rates enhanced with increased CO_2 concentration and higher temperature [89–91]. This can lead to greater incidence of picocyanobacteria blooms, threatening potable water quality and ecosystem health, and requiring investment to mitigate cyanotoxin risks.

1.4 Freshwater Picocyanobacteria vs. Marine Picocyanobacteria

1.4.1 Picocyanobacteria Phylogenetics

Cyanobacteria are an ancient phylum, diversifying through billions of years to occupy a plenitude of ecological niches [93]. Estimates for the divergence of Cyanobacteria from their closest relatives (Vampirovibrionia) range from 2.54 to 3.37 billion years ago [94–96], with the earliest diverging taxa (*Gloeobacter*) preserving a number of primitive characteristics (unicellular morphology and thylakoid absence [92]). The emergence of multicellularity in filamentous forms of cyanobacteria (i.e. *Pseudanabaena* and *Leptolyngbya*, among others) is thought to have contributed to the increase in cyanobacterial diversification around the time of the Great Oxidation Event (2.4 billion years ago), resulting in the widespread colonisation of aquatic and terrestrial habitats [97, 98] (1.2). The majority of extant cyanobacteria are derived from two clades - the Macrocyanobacteria and Microcyanobacteria (Figure 1.2) [92]. These are monophyletic clades which differentiated around the time of the Great Oxidation Event, primarily characterised by cell diameter [99]. Macrocyanobacteria include taxa with cell diameters > 3 µm and is a highly diverse clade containing the greatest phylogenetic diversity within Cyanobacteria [98]. Meanwhile, Microcyanobacteria maintained smaller cell diameters (between 2 and 3 µm), with a sub-group eventually evolving very small cells and radiating into picocyanobacteria [98].

Physically, picocyanobacteria are defined as unicellular cyanobacteria with a cell diameter smaller than 2 µm, comprising taxa within the Microcyanobacteria, Macrocyanobacteria, and basal clades associated with hot springs (i.e. convergent evolution of pico-morphology at various stages of cyanobacterial history). However, the primary lineage of picocyanobacteria form a monophyletic clade in the Microcyanobacteria, termed the *Syn/Pro* clade [100]. This clade was first identified in 2005 and is in essence,

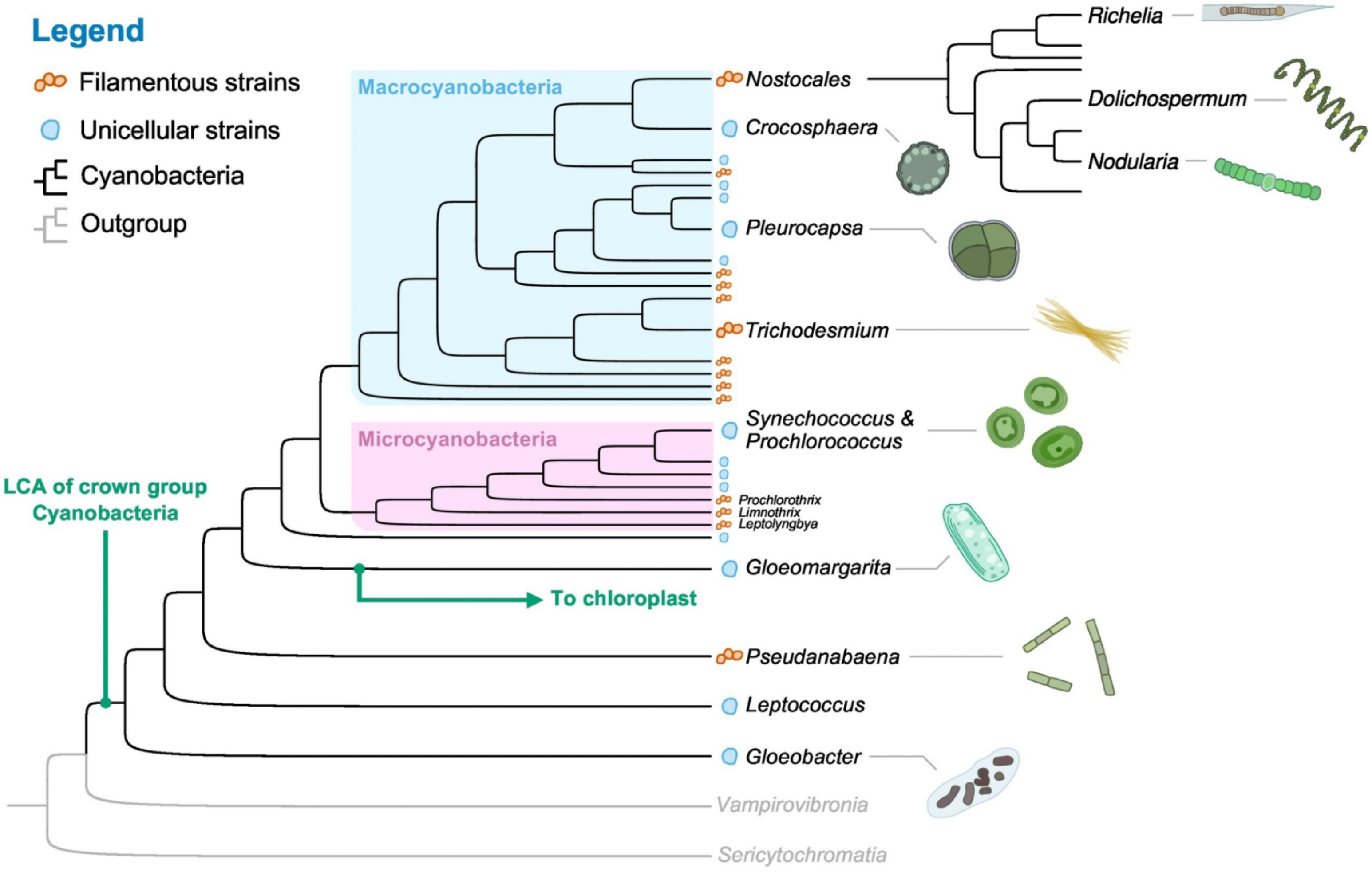

Figure 1.2: **The evolutionary relationships of Cyanobacteria** (taken from Sánchez-Baracaldo *et al.*, 2021 [92]). This phylogeny displays the relationship between the Macrocyanobacteria (highlighted in blue) and the Microcyanobacteria (highlighted in pink). Picocyanobacteria are members of the Microcyanobacteria, labelled as *Synechococcus* and *Prochlorococcus* in this figure.

the 'true' picocyanobacteria, exhibiting the smallest sizes. Though initially reliant on marine taxa, subsequent sampling has improved the resolution and four sub-clusters are now evident in the *Syn/Pro*, generally split between habitat (Figure 1.3). Sub-cluster 5.1 contains most marine *Synechococcus* spp., prevalent throughout the global oceans, and is further sub-divided into 20 sub-clades [101]. A sister group to sub-cluster 5.1 contains the *Prochlorococcus*, the most abundantly sampled sub-cluster. This is further divided into high-light and low-light clades (and further sub-ecotypes), specifically adapted to light niches in the tropical ocean [102]. Freshwater picocyanobacteria are found in sub-clusters 5.2 and 5.3, though both sub-clusters contain marine picocyanobacteria as well. Sub-cluster 5.2 contains the majority of known freshwater picocyanobacteria and has recently been significantly expanded by widespread sampling from Cabello-Yeves *et al* [103]. No ecotypes or sub-clades are identified with sub-cluster 5.2 as yet, though the increased representation now enables greater elucidation of this sub-cluster. This also applies to sub-cluster 5.3 - previously a very small sub-cluster with minimal representation, now expanded [103]. However, taxonomic names do not represent well-defined groups in phylogenetic terms [104]; while the *Syn/Pro* clade contain the majority of picocyanobacteria, a small number of *Synechococcus* are also represented among basal clades and the Macrocyanobacteria, phylogenetically distant from the *Syn/Pro* clade [105–107]. These strains were isolated from hot spring (basal) and estuarine (Macrocyanobacteria) locations in contrast to the mix of habitats of *Syn/Pro Synechococcus* and may indicate a case of convergent evolution of pico-sized morphology.

1.4.2 Representation of Freshwater Picocyanobacteria

Freshwater picocyanobacteria are widespread and abundant, yet their importance to the global ecosystem is not fully understood as a result of their historic under-representation in ecological research and genomic sequencing. The early adoption of *Synechococcus elongatus* PCC 7942 and *Synechococcus elongatus* PCC 6301 as freshwater model organisms resulted in the common utilisation of these strains to understand cyanobacterial photosynthesis and circadian rhythms [109–111]. Their general use as a *Synechococcus* model organism has led to the common misconception that *Synechococcus elongatus* are picocyanobacteria [112–118]. Not a member of the *Syn/Pro*, *Synechococcus elongatus* are a sister group to this clade of 'true' picocyanobacteria, and therefore are not representative of the wider freshwater picocyanobacteria community. The recent focus on freshwater picocyanobacteria has begun to shift

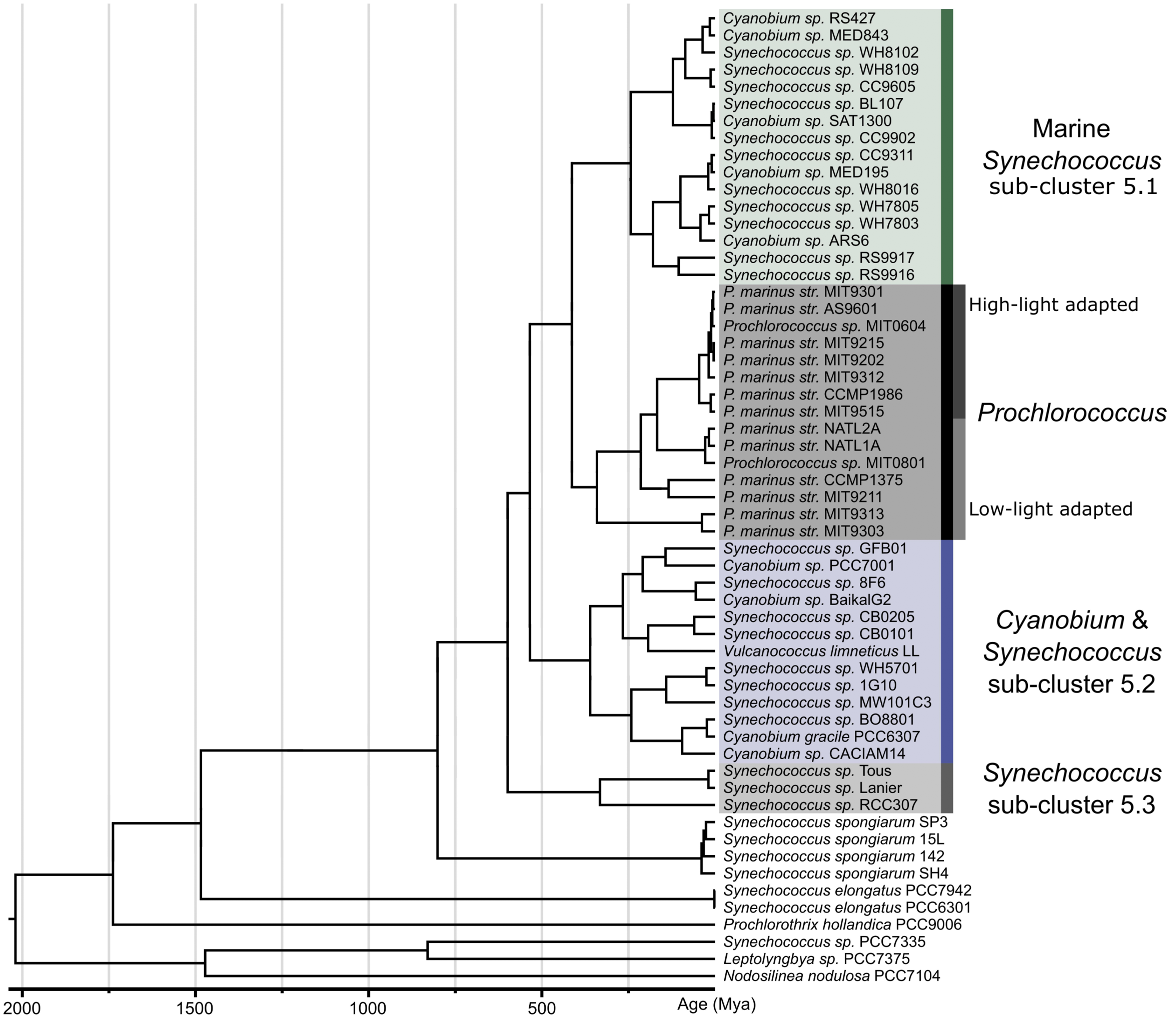

Figure 1.3: **The evolutionary relationships of picocyanobacteria** (adapted from Sánchez-Baracaldo *et al.*, 2019 [108]). This phylogenetic tree displays relationships between the sub-clusters of the *Syn/Pro* clade. Marine picocyanobacteria are located in sub-clusters 5.1, 5.2, and 5.3, and the *Prochlorococcus* sub-cluster. Freshwater picocyanobacteria are found in sub-cluster 5.2 and sub-cluster 5.3. A group of marine picocyanobacteria β-*Synechococcus* are not found in the *Syn/Pro* clade, instead located among the Macrocyanobacteria (not shown). All strains shown are in the Microcyanobacteria clade

.

attention away from *Synechococcus elongatus* and on to the *Syn/Pro* clade, achieved through sequencing of freshwater *Syn/Pro* genomes and phylogenetic investigations to understand the genetic variety within the *Syn/Pro* and the resulting ecological influence [103, 108, 119].

1.4.3 Sequenced Picocyanobacteria Genomes

Picocyanobacterial genomic research has historically focused on marine environments, until recently neglecting freshwater habitats. With the first marine picocyanobacteria genome sequenced in 2003, it took 11 years to reach 100 sequenced genomes in 2014, exponentially increasing to 1,688 genomes in 2023 (Figure 1.4). Meanwhile, the first freshwater picocyanobacteria was sequenced in 2012, eight years after *Synechococcus elongatus* PCC 6301, reaching 100 freshwater picocyanobacteria genomes by 2020, and 282 in 2023 (Figure 1.4). There are currently almost 6x the amount of available marine genomes than freshwater, a discrepancy which has until recently hindered the ability to study picocyanobacteria in freshwater habitats from a genomics perspective. Further difficulties have arisen due to the metagenomic approach to picocyanobacteria sequencing, resulting in smaller than expected genomes which are likely incomplete [120]. Therefore, only a subset of currently available genomes are suitable for comparative genomic analyses, reducing the available pool of genomes substantially. Therefore, increased genomic representation of freshwater picocyanobacteria is important to enhance our understanding of the molecular mechanisms present in freshwater environments. These environments are typically comprised of a greater diversity of habitats (compared to marine environments) due to increased nutrient heterogeneity and varied spatial and physical factors [121, 122]. Consequently, this may have enhanced the diversity of freshwater picocyanobacteria genomes, enabling greater adaptability to grow in diverse niches.

1.5 The Diversity of Nitrogen

1.5.1 The Importance of Nitrogen

N is an essential nutrient that is necessary for life on Earth as a component of proteins and nucleic acids. It is the most abundant element in our atmosphere in the form of N_2, however this form of N is only bioavailable to N-fixing bacteria (diazotrophs) [123, 124]. Due to this, N is often sparse in natural

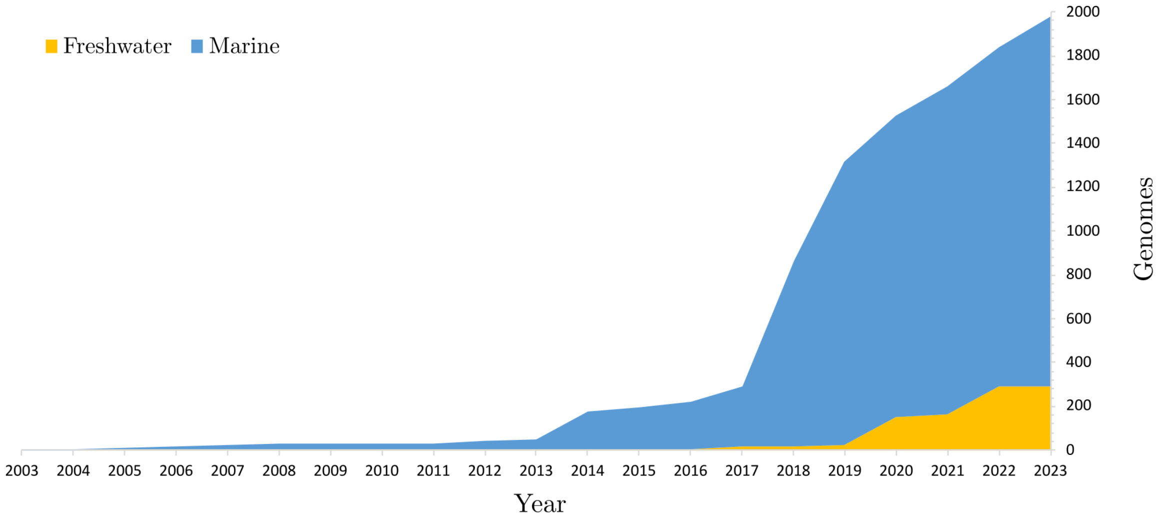

Figure 1.4: **Availability of freshwater picocyanobacteria vs. marine picocyanobacteria genomes.** The cumulative number of genomes, including metagenome-assembled genomes, belonging to the picocyanobacterial families of *Synechococcaceae* and *Prochloraceae* are plotted based on their isolation habitat and year submitted to NCBI. *Synechococcus elongatus* genomes are not included.

aquatic and terrestrial systems and acts as one of two long-established principal limiting factors of phytoplankton production alongside P [125]. The classic paradigm of nutrient limitation states that P is limiting in freshwaters while N is limiting is marine environments [126]. However, it is now generally accepted that the role of N in freshwater has been understated [6, 8], thus the co-limitation of N and P is considered the dominant nutrient state in freshwater (and marine [7]) environments. This includes nutritionally-starved oligotrophic habitats in addition to eutrophic lakes with high productivity [9].

The largest inputs of N in aquatic environments are terrestrial leaching and point source pollution, atmospheric deposition, and microbial N fixation [127]. In the UK alone, 2 Tg N yr^{-1} flows from the land to fluvial and limnetic systems, while 0.8 Tg N yr^{-1} travels further to estuarine and thalassic environments [128]. This land-based N is heavily influenced by anthropogenic activities which will be explored below. Another substantial source of aquatic N is atmospheric deposition, emitted through agricultural activity and the combustion of fossil fuels [129]. This atmospheric N is typically found in the forms of NH_4^+, NO_3^-, and nitric oxides, with resulting deposition leading to acidification and nutrient enrichment of water bodies potentially hundreds of kilometers away [129, 130]. On the other hand, N fixation is a naturally large source of N which feeds primary producer communities by fixing atmospheric N (unreactive N_2) to NH_4^+, providing bioavailable N upon subsequent NH_4^+ release. N fixation is a relatively common trait in cyanobacteria, found in both filamentous and unicellular morphologies, with the marine cyanobacterium *Trichodesmium* alone supplying over 100 Tg N y^{-1} to marine N budgets [131] - an important source of N to picocyanobacteria in the oligotrophic open ocean [132]. Additionally, evidence of picocyanobacterial genetic capability of N_2 fixation has been found though attempts to express these genes have as yet been unsuccessful [33, 133].

Anthropogenic N loads have significantly altered the global N cycle, depositing over 140 Tg N yr-1 onto the terrestrial system and directly into water bodies from point sources such as factories or sewage treatment plants [127, 134]. Terrestrial N loading primarily consists of agricultural non-point pollution sources including N-rich fertiliser applied to overcome nutrient limitation and organic waste generated from livestock [135]. However, this is not fully contained on land – large quantities of N are leached every year (up to 35 Tg N yr-1 [136]), entering into riverine systems and significantly enriching fluvial N transport. On the other hand, direct fluvial input from industry and sewage treatment are the primary contributors of N point source pollution, typically releasing less N than combined non-point sources (though often major sources of N pollution near urban areas) [135, 137]. In addition to an increased frequency of high-magnitude harmful algal blooms, elevated concentrations of N itself, especially in the form of NO_3^-, can compromise drinking water quality and endanger public health via increased carcinogenic risk [138–140]. The vastness of the open ocean has not rendered it immune to the impact of anthropogenic pollution, with N load increasing relative to P by 0.24 µmol kg^{-1} yr^{-1} in the North Pacific Ocean [141].

1.5.2 Diversity of Nitrogen Forms in the Environment

The prevailing sources of bioavailable N in freshwater have typically been dissolved inorganic N (DIN) forms, such as NH_4^+ and NO_3^- [142]. These accumulate due to the release of sewage and leaching of fertiliser, with ammonium nitrate (NH_4NO_3) the most commonly applied fertiliser in the UK [143]. However, the use of urea as an organic fertiliser is increasing, and now accounts for over 50% of worldwide fertiliser application [144]. This has raised awareness of urea, naturally only found in low concentrations, as an organic driver of phytoplankton growth [144, 145].

On the other hand, the contribution of other forms of dissolved organic N (DON) to phytoplankton productivity has received less attention [146]. Historically, N budgets and water quality measurements have neglected DON bioavailability and diminished its role as a N source, instead focusing on mitigating NH_4^+ and NO_3^- concentrations to prevent eutrophication. However, research has shown that cyanobacteria can grow utilising DON and AAs specifically as their sole N source [147, 148] (Table 1.1). The DON pool is a heterogenous mixture of nitrogenous compounds with significant concentrations of urea, AAs (both free and combined), nucleic acids, and humic substances amid many thousands of other compounds [146]. DON originates from a variety of allochthonous sources, including human and livestock excretion, cellular decay, soil leachate, and atmospheric deposition, and the contribution of DON to the dissolved N pool can be considerable (22 - 74%) [131, 149, 150]. Among freshwater catchments in both temperate and tropical environments, DON regularly represents the bulk of total dissolved N in oligo- and meso-trophic waterbodies [151–153]. Meanwhile, the abundance of DON remains high in eutrophic freshwater environments, though the DON:DIN ratio decreases as DIN concentration increases due to increased anthropogenic effects [153, 154]. DON is generally split into low molecular weight (LMW; $<$ 1,000 Da) and high molecular weight (HMW; $>$ 1,000 Da) classes [155]. The bioavailability of these two groups vary, with greater molecular complexity (i.e. HMW) resulting in recalcitrant compounds and reduced bioavailability [156, 157], while the LMW pool is labile and highly bioavailable, often found at decreased concentrations in oligotrophic environments compared to HMW molecules due to its low retention time [158]. From the total DON pool, up to 60% is thought to be readily metabolised by primary producers, significantly increasing the bioavailable N load in downstream waters and contributing to eutrophication [159, 160]. Many European legislative frameworks have been implemented which focus on inorganic N sources to quantify and manage eutrophication ([161–163]), yet the effective regulation of waters requires the consideration of additional N sources. Incorporation of organic N into the nutrient paradigm will allow us to enhance our understanding of the dangers of natural and anthropogenic DON on these systems and begin to effectively alleviate and mitigate them [146].

AAs are an essential bioavailable component of DON, found as both readily consumed dissolved free AAs (DFAA) or dissolved combined AAs (DCAA) forming polypeptides of varying size. The concentration of DFAAs in surface waters are typically in the nM range, yet their rapid turnover

Table 1.1: **Amino acids as a sole nitrogen source in cyanobacteria**

Strain	**Effective Sole N Source AAs**	**Ineffective Sole N Source AAs**	**Reference**
Synechococcus sp. PCC 7002	Ala; Arg; Asn; Asp; Glu; Gln; Gly; His; Phe; Pro; Ser; Thr	Trp	[164]
Synechococcus sp. PCC 6301	Gln	Gly; Glu; Asp; Asn; His; Met; Leu; Ala; Ser; Pro; Arg	[165]
Synechocystis sp. PCC 6714	Gln; Asn; Arg	Gly; Glu; Asp; His; Met; Leu; Ala; Ser; Pro	[165]
Anabaena sp. PCC 7118	Gln; Arg	Gly; Glu; Asp; Asn; His; Met; Leu; Ala; Ser; Pro	[165]
Pseudanabaena sp. PCC 6903	Asn; Arg	Gly; Glu; Gln; Asp; His; Met; Leu; Ala; Ser; Pro	[165]
Pseudanabaena sp. PCC B2	Gln; Asn	Gly; Glu; Asp; His; Met; Leu; Ala; Ser; Pro; Arg	[165]
Anabaena sp. PCC 7122	Ala; Arg; Asn; Asp; Gly; Glu; Gln; Pro; Phe; Ser; Val	Cys; His; Ile; Leu; Lys; Met; Thr; Trp; Tyr	[166]
Nostoc muscorum	His, Gln, Asn; Trp; Ser; Leu; Ile; Met; Val; Lys	Glu; Ala; Tyr; Cys; Asp; Thr; Gly; Arg; Pro; Phe	[167]
Nostoc ANTH	Gln; Asn; Arg	-	[168]
Microcystis aeruginosa	Ala; Arg; Leu; Glu; Asp	Lys	[169]
Spirulina platensis	Lys; Arg; Pro; Thr; Trp; Glu; Tyr; Asp; Gln; Ala; Gly; Asn; Leu; His; Ser; Val; Cys; Met	-	[170]

and efficient microbial uptake suggests a disproportionately large contribution of DFAAs to satisfy N requirements [171, 172]. DFAAs can be found at much greater concentrations in lake sediments (30 - 60 µM [173]), yielding a large supply of N upon sedimentary disturbance. DCAAs are generally found at higher concentrations than DFAAs (low µM range), however are not as readily available for N assimilation, first requiring sufficient extracellular catabolism to enable intracellular uptake [174]. As a proportion of total DON, the total dissolved AA (TDAA) pool (TDAA = DFAA + DCAA) in freshwater environments (5 - 28 % [175, 176]) makes up a greater proportion of DON than in marine environments (1 - 12 % [177]). Furthermore, oligotrophic waterbodies have a greater proportion of TDAA than eutrophic waters, increasing the role of TDAA as a N source in low nutrient environments [178]. The diversity of AAs may also play a role in the dominance of freshwater picocyanobacteria, with the composition of other bacterial communities shown to be influenced by preferential uptake of specific DFAAs [179]. Differing microbial community, catchment land use, and the variety of nutrient inputs can impact the AA composition, resulting in waterbodies with varying dominant TDAA profiles [171, 173, 174].

1.5.3 Energetics of Nitrogen Assimilation

The energy required to assimilate N may provide insights into the preferential status of different nitrogenous forms. The most reduced form of inorganic N is NH_4^+, commonly thought to be the preferred cyanobacterial N source [180]. This is due to the immediate incorporation of NH_4^+ into the C skeleton via the glutamine synthetase/glutamate synthase (GS-GOGAT) pathway, yielding two molecules of glutamate (Figure 1.5). Meanwhile, NO_3^- and nitrite (NO_2^-) are oxidised forms of N requiring active transport and reduction to NH_4^- for subsequent incorporation [181]. This is achieved through a two-step process (Figure 1.5); first the reduction of NO_3^- to NO_2^- requiring two electrons and nitrate reductase (encoded by *narB*), followed by the reduction of NO_2^- to NH_4^+ requiring six electrons and nitrite reductase (encoded by *nirA*). Though less energetically favourable than NH_4^+, NO_3^- and NO_2^- assimilation is much less costly than N_2 fixation. In the absence of suitable N sources, diazotrophic cyanobacteria can fix N_2 into NH_4^+, catalysed by nitrogenase [182]. However, this reaction is energetically demanding, requiring 16 molecules of ATP and eight electrons (provided by reduced ferredoxin) per molecule of N_2 fixed. As such, other sources of N are preferred when available, though diazotrophs are often pioneering species in nutrient deficient environments [183].

The assimilation of organic forms of N may be just as energetically favourable as inorganic forms, if not more so, though research in this area is less well explored and focused on urea [12, 184]. Urea can be imported into the cell via passive diffusion or active transport, though low natural concentrations of urea commonly results in uptake by the high-affinity Urt transport system [185] (Figure 1.5). Urea is hydrolysed by the constitutively expressed Ni^{2+}-dependent urease, producing two NH_4^+ molecules for efficient N incorporation plus a by-product of carbon dioxide which can be further incorporated into the Calvin Cycle [186, 187]. Assimilation of AAs, especially glutamate, may be the most advantageous for cyanobacteria. Gluatamate is the N donor for all AAs, providing an amine-group directly to eight of the 20 proteinogenic AAs and indirectly, via AA intermediates, to the remaining 12 [188]. Additionally, glutamate and glutamine together provide all the N necessary for nucleotide biosynthesis [188]. The assimilatory benefits of other AAs are based on the products of their catabolism and length of their degradative pathways, with proline and aspartate an additional effective source of glutamate (described in more detail in section 1.7.2) [189]. The greater N content of AAs compared to inorganic sources may provide another advantage; six AAs contain two or more amine groups, arginine bearing the largest number of N atoms for assimilation. This increased N content is utilised as a N store, with long chains of arginine and aspartate forming cyanophycin in larger cyanobacteria (though not present in picocyanobacteria) [190, 191].

1.5.4 Nitrogen Forms and the Phytoplankton Community

The dominant forms of N can impact the composition of the phytoplanktonic community [192]. It has previously been determined that the trophic status of a water body can alter which taxa are able to dominate their environment (see section 1.3.1), with macrocyanobacterial growth favourable in eutrophic concentrations [193–195], while picocyanobacteria dominate oligotrophic environments [38, 196]. However, the composition of bioavailable N can also modify the species community. Mesocosm experiments have showed that the proportion of NH_4^+ to NO_3^- has direct impact on a mixed community, favouring cyanobacteria at elevated NH_4^+ concentrations and eukaryotic phytoplankton (especially diatoms) at increased NO_3^- concentration [197]. This pattern of cyanobacterial preference of reduced forms of N extends to DON, where marine and freshwater urea and DFAA uptake are associated with cyanobacteria and cryptophyte growth [198–200]. However, few studies have investigated the different components of DON and their role in influencing microbial composition, outside of urea. The energetic advantages of AA assimilation, and their increased importance in oligo- and meso-trophic environments, may suggest strong potential influences on picocyanobacteria and cyanobacteria as a whole.

1.6 Molecular Basis of Cyanobacterial Nitrogen Assimilation

Model cyanobacteria such as *Synechocystis* sp. PCC 6803, *Nostoc* sp. PCC 7120, and *Synechococcus elongatus* PCC 7942 have been the conduits for the majority of molecular cyanobacterial research.

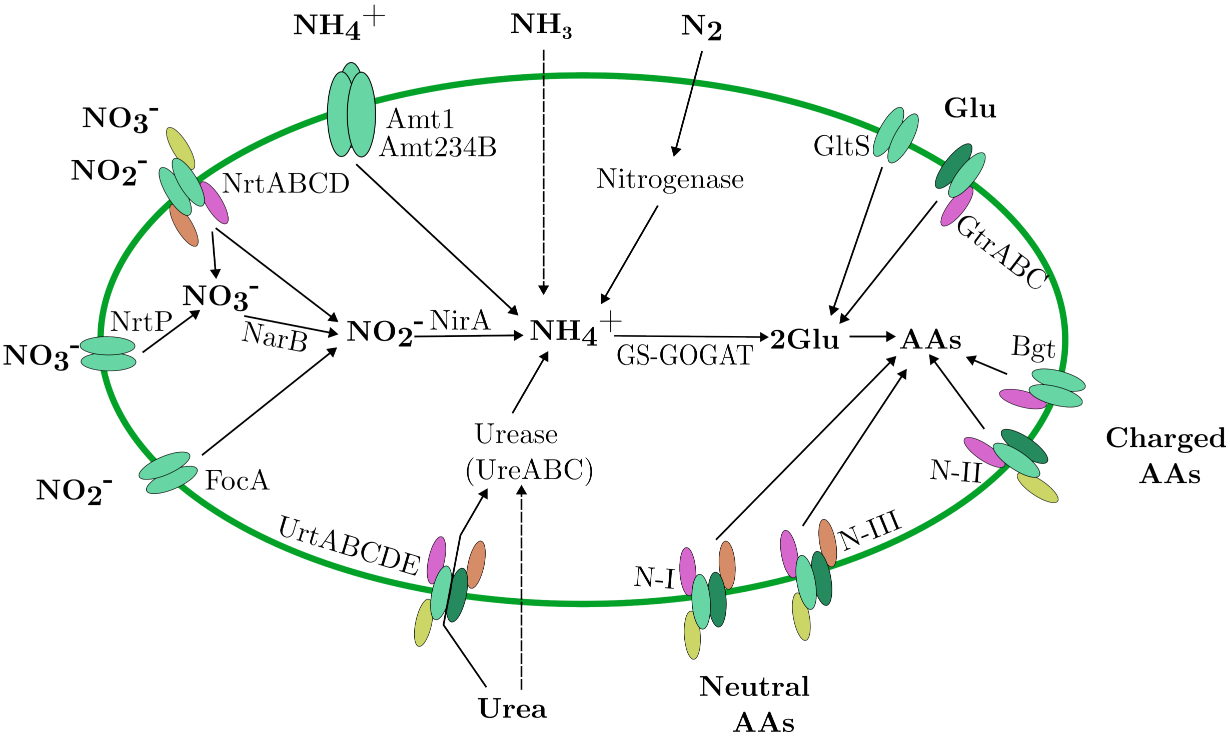

Figure 1.5: **The assimilation of nitrogen in cyanobacteria.** N is taken up from the environment in many forms. Most N sources are metabolised into NH_4^+ which is subsequently incorporated into C skeletons via the glutamine synthetase/glutamate synthase pathway (GS-GOGAT). Atmospheric N_2 is fixed to NH_4^+ by nitrogenase in diazotrophs. NH_4^+ is imported through the Amt1 transporter, with Amt2/3/4/B being accessory transporters to aid uptake. At high concentrations and alkaline pH, NH_4^+ in the form of ammonia (NH_3) can be taken up passively. NO_3^- and NO_2^- are imported via different systems in freshwater and marine cyanobacteria. The Nrt complex in freshwater cyanobacteria displays high affinity for both NO_3^- and NO_2^- while marine cyanobacteria utilise NrtP to import NO_3^- and FocA to import NO_2^-. Once imported, NO_3^- is reduced to NO_2^- by nitrate reductase (NarB), which is subsequently reduced to NH_4^+ by nitrite reductase (NirA), all encoded on a singular operon (*nirA-nrtABCD-narB* / *nirA-nrtP-narB*). Urea can be transported passively or imported via the Urt ATP-binding cassette (ABC) transporter system, then hydrolysed by urease (UreABC) to produce two molecules of NH_4^+ and one molecule of CO_2 (not shown). Uptake of AAs depends on transporter preferences, with N-I and N-III transporters able to uptake neutral AAs, while N-II and Bgt transporters can uptake charged AAs. GltS and GtrABC are glutamate-specific transporters. Transporter component functions are indicated by colours. Green: transmembrane proteins (different shades indicate hetero-oligomers); purple and orange: ATPases; yellow: periplasmic substrate-binding proteins. Dashed arrows indicate passive uptake. Other sources of DON are not included due to unknown assimilation mechanisms. This figure displays N assimilation pathways found in cyanobacteria from freshwater and marine environments, thus not all cyanobacteria have all these assimilation capabilities.

While it is expected that, broadly, freshwater picocyanobacteria behave similarly to these strains (especially non-filamentous genera such as *Synechocystis* or *Synechococcus elongatus*), the necessary research on assimilation pathways has not been carried out. Therefore, this section primarily references cyanobacteria as a whole, with known picocyanobacterial research and mechanistic differences mentioned.

Most bioavailable nitrogenous forms require active transport to be assimilated into the cell. This primarily takes place via ATP-binding cassette (ABC) transporters or active permeases. An ABC transporter utilises the energy supplied from the hydrolysis of ATP to an ATP-binding domain to transport substrates across a membrane. These transporters typically contain three distinct subunits - an extracellular/periplasmic substrate-binding domain which selectively binds to extracellular nutrients. This brings the bound substrate to the trans-membrane domain which then imports the substrate via a conformational change triggered by ATP hydrolysis in the ATP-binding domain [201]. On the other hand, the mechanism behind active permease transport utilises antiporter action, importing high-affinity substrates while co-transporting ions (such as Ca^{2+}, H^{+}, Na^{+}) out of the cell [202].

1.6.1 Regulation of Nitrogen Assimilation

The regulation of N assimilation in cyanobacteria is controlled by the concentration of intracellular N in the form of NH_4^+ and its incorporation into glutamate via the GS-GOGAT pathway. This is generally achieved through three regulatory proteins, and the monitoring of one metabolite. The global N regulator NtcA (encoded by *ntcA*), signal transduction protein PII (encoded by *glnB*) and co-activating factor (encoded by *pipX*) form a network of interactions to respond to the internal indicator of N availability, 2-oxoglutarate (2-OG) [203, 204]. An intermediate of the TCA cycle and synthesised by isocitrate dehydrogenase (encoded by *icd*), 2-OG is a key regulating metabolite which links C metabolism to N assimilation [205]. Acting as the C backbone of NH_4^+ incorporation into glutamate, fluctuations in N availability (NH_4^+ concentration) lead to resultant change in 2-OG abundance (Figure 1.6). In times of sufficient N, the concentration of 2-OG remains steady due to constant production via the TCA cycle and incorporation into glutamate. However, when the NH_4^+ substrate is not sufficiently available, intracellular 2-OG builds up, resulting in the rapid activation of signal pathways to promote N uptake [206].

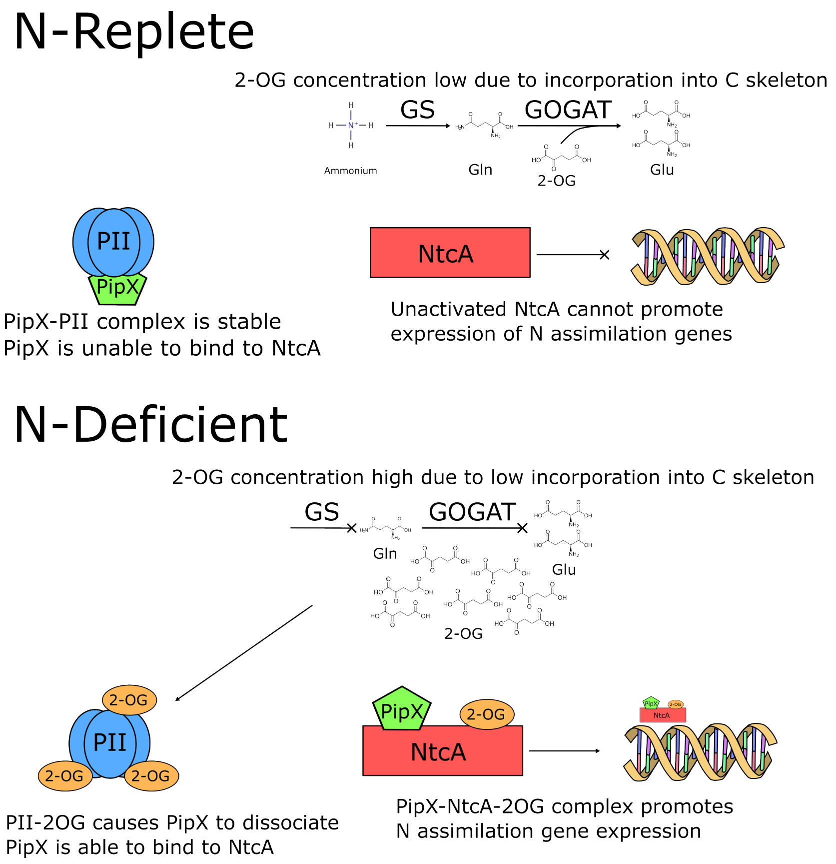

Figure 1.6: **Intracellular regulation of the global nitrogen regulator NtcA**. In sufficient N conditions, 2-OG concentration remains stable due to incorporation into the C skeleton via the GS-GOGAT pathway. This enables PII to bind PipX, resulting in an inactivated NtcA. When N is limited, 2-OG accumulates and interacts with PII, causing dissociation of PipX. Free PipX interacts with NtcA leading to enhanced activity of NtcA, enabling transcription of the NtcA regulon present in many N assimilation genes. GS: glutamine synthetase; GOGAT: glutamate synthase.

The N signalling network relies on PII as the direct sensor of 2-OG concentration [207]. During NH_4^+ abundance, PII and PipX form a stable complex [208]. However, an increase in 2-OG concentration due to N depletion results in the binding of accumulated 2-OG to PII, altering PII-PipX structure and promoting the dissociation of PipX [209] (Figure 1.6). Free PipX is now available to form a complex with NtcA and 2-OG [210]. The NtcA-PipX-2-OG complex has the capability to activate transcription of the NtcA regulon, promoting expression of many genes involved with N assimilation (including *glnB* (encodes PII) and *ntcA*) [211–213]. Once N limitation is relieved, the corresponding decrease in ATP:ADP ratio leads to ADP interacting with PII, in turn reducing the interaction between PII and 2-OG, resulting in the subsequent formation of PII-PipX complexes to deactivate NtcA [214].

Though this is the primary method of N assimilation regulation in cyanobacteria, there are additional factors which may be involved. An additional transcriptional regulator, NtcB (encoded by *ntcB*), has been identified in *Synechococcus elongatus* PCC 7942 and Macrocyanobacteria but not found in marine picocyanobacteria [215–217]. NtcB enhances the ability to transcribe the NO_3^-/NO_2^- assimilation operon (*nirA-nrtABCD-narB*) alongside NtcA, activated in response to NO_2^- [218]. Furthermore, though most research on the regulation of N assimilation has focused on DIN, genes involved with DON assimilation and metabolism have been identified within the NtcA regulon [213, 219]. This includes a urea transporter and charged AA transporter, in addition to genes involved with AA biosynthesis and degradation [213]. Though comprehensive studies of NtcA regulation of DON assimilation genes are lacking, the role of NtcA as a promoter of DFAA and DCAA uptake is likely and supports the utilisation of these N sources as essential nutrients when available.

1.6.2 Ammonium/Ammonia Assimilation

Extracellular NH_4^+ uptake is primarily carried out via active permeases (Figure 1.5). The major NH_4^+ transporter in cyanobacteria is Amt1, encoded by *amt1* (part of the NtcA regulon) and first characterised in *Synechocystis* sp. PCC 6803 [220]. The Amt1 transporter displays high affinity to NH_4^+ and enables uptake at low nutrient concentrations (µM range) [220]. A number of accessory proteins are also associated with the Amt1 complex. These include AmtB (*amtB*) from *Synechococcus elongatus* PCC 7942 [221], Amt2 (*amt2*) and Amt3 (*amt3*) from *Synechocystis* sp. PCC 6803 [220], and Amt4 (*amt4*) from *Nostoc* sp. PCC 7120 [222]. The expression of these proteins are within the NtcA regulon, with the exception of *amt3* where no such NtcA-binding site has been found [213]. The function of these proteins is suggested to assist with Amt1 NH_4^+ uptake via incorporation into the Amt trimer, maximising the activity of Amt1 to enable uptake at reduced NH_4^+ concentrations [220].

In addition to active transport, ammonia (NH_3) can diffuse across the lipid bilayer or through porins providing a passive source of NH_4^+ following protonation, demonstrated by the growth of Amt1 deletion (Δ*amt1*) strains when grown under high NH_4^+ concentrations [223]. However, passive NH_3 uptake is insufficient to sustain N requirements in all but the most nutrient-enriched environments, suggesting porins are involved with NH_3 leakage rather than uptake, with active NH_4^+ transport the primary method of uptake [224].

As described previously, intracellular NH_4^+ is incorporated into the C skeleton of 2-OG through the GS-GOGAT pathway. The first phase of this pathway utilises glutamine synthetase (GS) to catalyse the amination of glutamate to glutamine via the addition of NH_4^+. GS activity is present in two forms in cyanobacteria, GSI (encoded by *glnA*) and GSIII (encoded by *glnN*) (GSII is primarily found in eukaryotes [225]). The expression of these two proteins are determined by the intracellular N status [226]. GSI is the primary form of GS in cyanobacteria, responsible for more than 95% of GS activity in *Synechocystis* sp. PCC 6803 [206]. On the other hand, GSIII accounts for only 3% of *Synechocystis* sp. PCC 6803 GS activity when grown under N-replete conditions, rising to 20% when N limitation is induced [227]. Thus, though GSIII activity is only auxiliary to the constitutively expressed GSI in sufficient N conditions, the onset of N limitation requires the ability to scavenge the maximum amount of NH_4^+ [228]. Furthermore, unlike GSI, *glnN* (GSIII) is not present in all cyanobacteria, suggesting this additional boost for N assimilation is not ubiquitous among this group [228]. The second step of the GS-GOGAT pathway is catalysed by glutamate synthase (encoded by *glsF*), which catalyses the formation of two glutamate molecules via the incorporation of 2-OG into the previously produced glutamine.

1.6.3 Nitrate/Nitrite Uptake

The uptake of NO_3^-/NO_2^- is carried out via alternative mechanisms in freshwater and marine cyanobacteria (Figure 1.5). The first major cyanobacterial NO_3^- transport system was identified in *Synechococcus elongatus* PCC 7942 and is a member of the ABC superfamily [229]. This transporter, termed NRT, is encoded by four genes - *nrtABCD* - located in the NtcA-regulated *nir* operon, and displays high affinity for both NO_3^- and NO_2^- [230]. Comparative genomic analyses have identified NRT in most freshwater cyanobacteria strains, yet it is absent from marine cyanobacteria, regardless of phylogenetic affiliation [217]. Instead, marine cyanobacteria encode different transporters to import NO_3^- and NO_2^-. NrtP (encoded by *nrtP*) is a permease of the major facilitator superfamily which has high affinity for NO_3^- only, displaying reduced affinity to NO_2^- [231]. Meanwhile, FocA (encoded by *focA*) is the high affinity NO_2^- transporter present in marine cyanobacteria [232], thus the co-expression of FocA and NrtP contributes the required machinery for high-affinity uptake of both NO_3^- and NO_2^-.

Additional NO_3^-/NO_2^- transporters have also been identified. Some *Synechococcus* spp. encode a novel transporter (NrtS) with a lower affinity for NO_3^-/NO_2^- than the primary NRT and NrtP/FocA transporters, however NrtS was only actively expressed in the marine *Synechococcus* sp. PCC 7002 encoding two copies of *nrtS* [233]. Expression of *nrtS* was not achieved in *Synechococcus elongatus* PCC 7942 which encodes only a single copy of the gene [233]. However, as neither strain studied are *Syn/Pro* clade *Synechococcus*, the wider distribution of NrtS in picocyanobacteria is unclear. Furthermore, transporters for other substrates may also import NO_2^-. The cyanate transporter of *Synechococcus elongatus* PCC 7942, encoded by *cynABD*, was found to actively transport NO_2^- [234]. Though displaying lower affinity for NO_2^- than the NRT transport system, the cyanate transporter was able to contribute 30% of required N for optimum growth, though only in the absence of cyanate [234]. However, the presence of cyanate also constitutes a source of N, with cyanase (encoded by *cynS*) catalysing the bicarbonate-dependent conversion of cyanate to NH_3 and CO_2 [235].

1.6.4 Nitrogen Fixation

The nitrogenase complex is the central component of N fixation, responsible for reducing atmospheric N_2 into NH_3 and H_2. Nitrogenase consists of two distinct enzymes - dinitrogenase reductase and dinitrogenase - encoded by the *nifHDK* operon [236]. N fixation requires a substantial amount of reducing energy, supplied by the action of dinitrogenase reductase via reduced ferredoxin and ATP. Electrons are transferred to the N_2-bound FeMo-cofactor (molybdenum component can be substituted with vanadium or iron [123]) of dinitrogenase and six sequential reduction events fully reduce one molecule of N_2 to two molecules of NH_3. Many accessory proteins are involved with this reaction, encoded by three other *nif* operons (*nifENXW*, *nifB-fdxN-nifSU*, and *nifVZT*). The function of these proteins primarily include structural assistance for FeMo-cofactor assembly (NifE, NifN, NifX) [237] and precursor synthesis (NifS, NifU, NifB) [238]. Nitrogenase is inactivated by oxygen (O_2), thus cyanobacteria have evolved mechanisms to reduce O_2 exposure. These include the temporal separation of photosynthesis and nitrogenase function [239, 240], or spatial separation through the development of heterocysts (filamentous cyanobacteria) and diazocytes (*Trichodesmium* spp.) [241, 242]. These specialised cells produce an environment for nitrogenase which lacks photosynthetic machinery, enabling effective N_2 reduction in localised anaerobic conditions [243, 244]. Once N is fixed into NH_3 in these specialised cells, it must be transported to vegetative cells; in heterocysts, high levels of GS alongside low concentrations of GOGAT suggest fixed N is transferred in the form of glutamine [245]. Additionally, the accumulation of cyanophycin in heterocyst cells provides a reservoir of arginine and aspartate to utilise when N is scarce [246].

1.6.5 Urea Assimilation

Urea can be imported through passive diffusion or active transport, the latter being the dominant method of uptake due to generally low environmental concentrations [131]. The active transport system in cyanobacteria is Urt, a five-component ABC-type transporter (UrtABCDE) that has high affinity for urea enabling uptake at low concentrations (< 1 µM) [247]. The Urt components are encoded by the *urtABCDE* operon and under NtcA regulatory control. However, recent research has suggested there may be additional regulatory mechanisms for *urt* expression, due to seemingly independent up-regulation compared to other NtcA-controlled genes [248]. This may be due to the use of urea as a C source in addition to N, while it has also been suggested that Urt may transport additional molecules [184].

The hydrolysis of urea into NH_3 and CO_2 is carried out via urease and its accessory proteins. These are encoded by the NtcA-regulated *ureABCDEFG* operon, with *ureABC* encoding the urease enzyme itself while *ureDEFG* encode chaperone proteins which stabilise urease conformation and aid in metallocentre active site formation [249–251]. Cyanobacteria can encode the *ure* operon without the *urt* operon, with urease necessary for breaking down internally generated urea to prevent toxic accumulation [184].

1.6.6 Amino Acid Uptake

Environmental DFAAs can be imported into the cell through eight different transporters in cyanobacteria (Figure 1.5) [252]. These transporters vary depending on which AAs are selectively taken up, differing based on characteristics such as hydrophobicity and charge of the AA. This results in different transporter preferences, potentially impacting the diversity of AAs imported. Much of this research has been carried out with the model cyanobacteria of *Synechocystis* and *Dolichospermum*, and thus knowledge is lacking in picocyanobacterial characterisation.

There are four major ABC-type broad-specificity AA transporters (AAT) in cyanobacteria [253–255]. Of these four, two are primarily associated with neutral AA uptake - N-I and N-III [253, 255]. The most common method for AAT characterisation is the production of a deletion mutant, knocking out one or more components of the transporter system and determining changes in AA uptake. Knockout and full deletion of the five-component N-I transporter (Table 1.2) resulted in severely impaired uptake of neutral AAs, in addition to histidine (likely due to its weak basic side chain (pKa of 6.0) resulting in partial protonation at pH 7) [253]. Uptake assays have also revealed that proline is the preferred substrate of N-I, alongside leucine and phenylalanine [253, 256, 257]. The N-III AAT fulfils a similar function to N-I transporting neutral hydrophobic AAs, however different preferred substrates are found [255]. Inactivation of this transporter resulted in a 50% decrease of glycine uptake in *Nostoc* sp. PCC 7120 (a species which also harbours the N-I transporter) compared to wildtype, in addition to reductions in proline, glutamate, and alanine among others [255].

Charged AAs are imported into the cell via the action of other transporters. Acidic AAs are imported by N-II, this four-protein complex encoded by the *natF-natG-natH-bgtA* gene cluster (Table 1.2) [254]. Removal of the N-II transporter resulted in severely reduced uptake of aspartate and glutamate, indicating these are the preferred substrates for this acidic AAT [254]. Further competition assays and deletion studies suggest N-II can recognise acidic and neutral polar AAs, displaying a range of AA uptake potential [254, 258]. The only basic AAT in cyanobacteria is the Bgt transporter which shares its ATPase with the N-II transporter (Table 1.2). The uptake of basic AAs (arginine, lysine, and histidine) significantly decreased upon *bgtB* deletion, though the magnitude of histidine uptake reduction was lessened due to its import from neutral transporters at pH 7.0 [254, 259]. Additionally, aspartate and glutamine uptake dropped (to a lesser extent) upon *bgtB* deletion, suggesting limited capability of importing other AAs [254].

In addition to broad-specificity ABC-type AATs, two glutamate-specific transporters have been identified in cyanobacteria [259]. The first is a tripartite ATP-independent periplasmic (TRAP) transporter, consisting of two integral membrane proteins (encoded by *gtrA* and *gtrB*) and a periplasmic substrate-binding protein, identified in *Synechocystis* sp. PCC 6803 (*gtrC*) [259]. Δ*gtrAB* and Δ*gtrC* strains resulted in the impaired uptake of glutamate with a 30% reduction in import compared to wildtype

Table 1.2: **Components of cyanobacteria ABC-type amino acid transporters**

Transporter	Protein	Function
N-I (Neutral)	NatA	ATP-binding
	NatB	Periplasmic substrate binding
	NatC	Transmembrane protein
	NatD	Transmembrane protein
	NatE	ATP-binding
N-II (Acidic)	NatF	Periplasmic substrate binding
	NatG	Transmembrane protein
	NatH	Transmembrane protein
	BgtA	ATP-binding
Bgt (Basic)	BgtA	ATP-binding
	BgtB	Transmembrane protein and periplasmic substrate binding
N-III (Neutral)	NatI	Periplasmic substrate binding
	NatJ	Transmembrane protein
	NatK	Transmembrane protein
	NatL	ATP-binding
	NatM	ATP-binding

[259]. The second glutamate transporter is GltS, a mono-component permease also characterised in *Synechocystis* sp. PCC 6803 [259]. Similarly to the TRAP transporter, a Δ*gltS* strain resulted in the disruption of Na-dependent glutamate uptake (56 - 69% of wildtype) while other AAs were unaffected. The presence of glutamate-specific transporters is noteworthy due to the central role of glutamate in N incorporation, responsible for providing amine groups for subsequent AA biosynthesis.

Further AATs have been putatively identified though not yet fully characterised. An AA permease (DmeA) of the DMT superfamily was identified in *Synechococcus elongatus* PCC 7942 [258]. The deletion of *dmeA* resulted in significantly reduced uptake (30%) of aspartate, glutamine, and glutamate compared to wildtype [258]. Interestingly, deletion of the N-II transport system and presence of DmeA resulted in an almost compete impairment of DmeA substrates ($<$ 10% of wildtype). As Δ*dmeA* did not affect the expression of the N-II transporter, it is thought that DmeA is dependent on the activity of N-II for function and may enhance uptake of AAs through this ABC-type AAT though the mechanism of this function is not known. Another novel cyanobacterial AAT is the Ala/Gly:cation symporter (AgcS), discovered in the halotolerant cyanobacterium *Aphanothece halophytica* [260]. This is encoded by *apagcs1* with high homology to the AGCS superfamily. Though no deletion mutants for this gene have been created, transformation of *apagcs1* into *E. coli* strain JW4166 (which is deficient in glycine import) resulted in the Na-dependent uptake of glycine [260]. Co-addition of other AAs (aspartate and glutamine) reduced the uptake of glycine, indicating an interaction between these additional AAs and the transporter.

Transporters involved with the uptake of small chains of DCAA have been putatively identified in cyanobacterial genomes, though also not experimentally characterised. An open reading frame in *Synechococcus elongatus* PCC 6301 displays 43.6% identity to *oppC* from *Salmonella typhimurium* [261]. This encodes an oligopeptide permease membrane protein which imports peptides into the cell. Expression of this gene in *Synechococcus elongatus* PCC 6301 has not yet been detected and typical promoter motifs missing upstream may suggest that this gene may be a pseudogene or regulated by an uncommon regulator [261]. Additionally, a component of a dipeptide transporter has been found in marine *Synechococcus* and *Prochlorococcus* genomes [10]. This component is homologous to *dppB* which encodes an integral membrane subunit of the ABC-type transporter Dpp. The Dpp transporter is linked to the utilisation of di/tripeptides and may present a further novel method of uptake for combined amino acids [262].

1.7 Molecular Basis of Cyanobacterial Amino Acid Metabolism

An understanding of the cyanobacterial ability to biosynthesise and metabolise AAs is vital to understand the mechanisms behind AA assimilation. The recent mass sampling of freshwater picocyanobacteria and subsequent comparative genomics has demonstrated the presence of enzymatic biosynthesis for all 20 proteinogenic AAs [103]. However, despite the importance of these essential processes for biological function, our understanding of entire pathways are relatively poor, especially for AA catabolism [189].

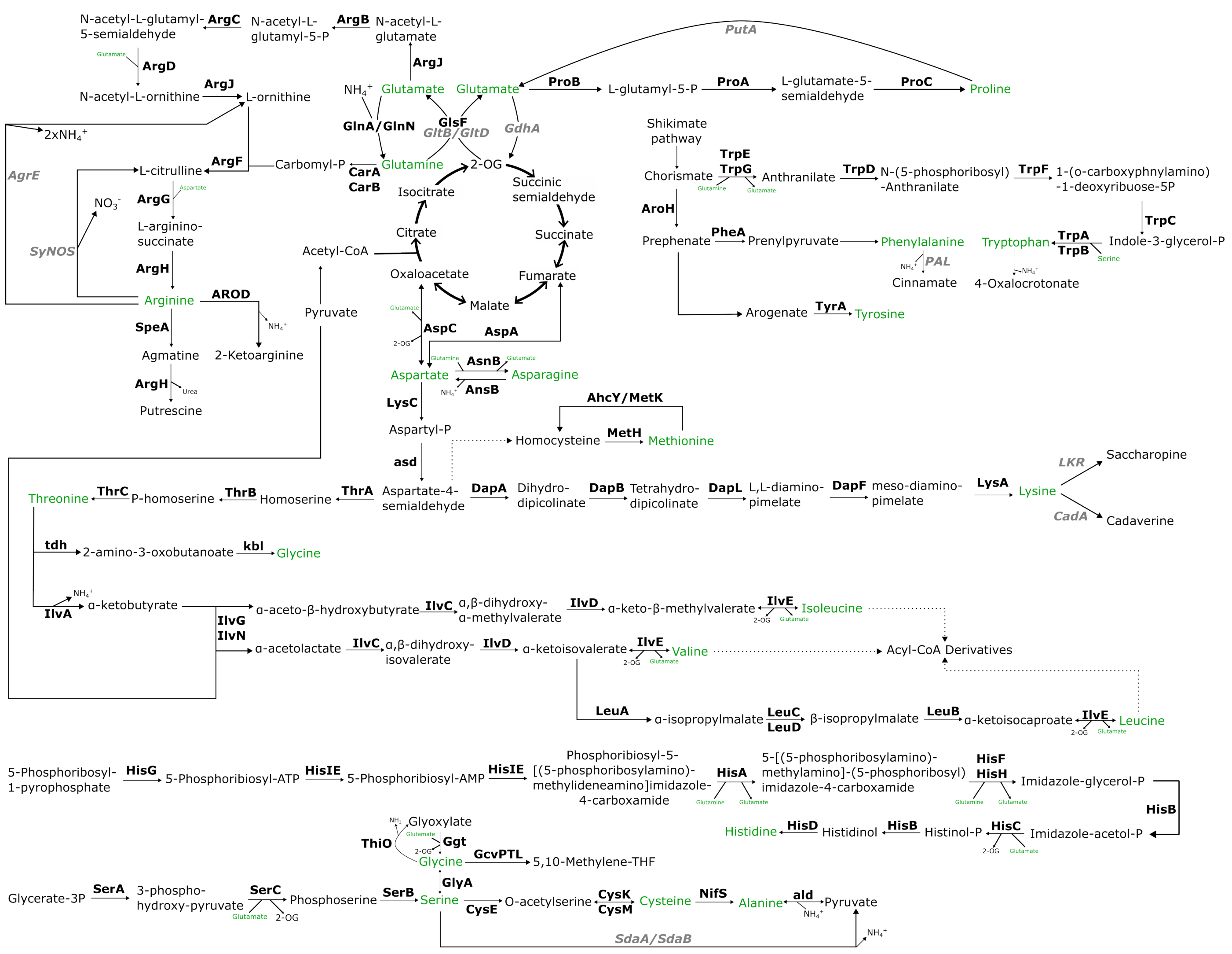

Figure 1.7: **Biosynthesis pathways of the 20 proteinogenic amino acids in cyanobacteria**. The 20 proteinogenic L-AAs are highlighted in green. Co-factors are not included except for amino acids and 2-OG. Dashed lines indicate condensed pathways. Enzymes are in bold while greyed out enzymes are those not known to be found in freshwater picocyanobacteria. Adapted from Mills et al. [189].

1.7.1 Glutamate, Glutamine, Proline, and Arginine Metabolism

Glutamate is the central molecule of AA biosynthesis, itself synthesised from glutamine and 2-OG. This reaction is catalysed by the ferredoxin-dependent GOGAT in freshwater picocyanobacteria (GlsF) though an NAD(P)H-dependent GOGAT (GltB/GltD) is also present in *Synechocystis* [263] (Figure 1.7). The deamination of glutamate to yield 2-OG and NH_4^+ is catalysed by glutamate dehydrogenase (GdhA), an enzyme not present in freshwater picocyanobacteria, preventing the direct recycling of glutamate to NH_4^+ [103, 264]. Glutamate is converted to glutamine by the incorporation of NH_4^+ catalysed by GS, with two isoforms as discussed in section 1.6.2. Glutamate is also the substrate for proline biosynthesis with the successive action of three enzymes - ProB, ProA, and ProC - phosphorylating glutamate followed by two reduction steps to yield proline. Larger cyanobacteria encode proline oxidase, PutA, which can catabolise proline to glutamate via reduction of $NADP^+$, however the presence of this enzyme in picocyanobacteria is undetermined [265]. The universal importance of glutamate to metabolic processing highlights its favourability as a N source, supported by growing evidence for glutamate assimilatory preference over DIN forms [154].

Arginine, synthesised through a glutamate precursor, is composed of 33% N making it the most N-rich AA. Eight successive reactions, including the incorporation of N from glutamate, carbomyl phosphate (synthesised from glutamine), and aspartate, yield arginine (Figure 1.7) with all enzymes present in freshwater picocyanobacteria [103]. However, the distribution of arginine catabolism pathways are less well understood. Due to the large N pool available in arginine, multiple pathways are known for potential N liberation. Research has not yet been conducted on the freshwater picocyanobacteria encoding capabilities of these pathways, though *Synechococcus elongatus* encodes none in full [266, 267]. Of two partially encoded pathways of arginine degradation in *Synechococcus elongatus*, the first is via the arginine decarboxylase pathway, decarboxylating arginine to agmatine (SpeA) and subsequent conversion to putrescine and the release of urea (ArgH) [103, 266]. Alternatively, arginine oxidation to 2-ketoarginine releases NH_4^+ through the action of arginine oxidase (AROD) [266]. An additional pathway of arginine-N catabolism has been identified in marine *Synechococcus*, via the oxidation of arginine by nitric oxide synthase (SyNOS) yielding citrulline and NO_3^- [267]. Further metabolism of these metabolic products is yet to be determined although multiple methods of arginine-N degradation and its high N-content suggest a key role of arginine in N mobilisation.

1.7.2 Aspartate, Asparagine, Lysine, Threonine, and Methionine Metabolism

The transamination (catalysed by AspC) of glutamate with the TCA intermediate oxaloacetate results in the generation of aspartate and 2-OG (Figure 1.7) [268]. This reaction can also be performed in the reverse direction, transaminating aspartate to yield glutamate. Aspartate can be additionally catabolised by deamination, catalysed by aspartase (AspA), to produce the TCA intermediate fumarate and NH_4^+, with these enzymes encoded by freshwater picocyanobacteria [103]. With AAT preferences often dictated by charge, the uptake and assimilation capabilities of negatively-charged aspartate may be concomitant with glutamate, potentially resulting in adaptations to efficiently utilise aspartate as a nutrient source. Aspartate also functions as the precursor to asparagine biosynthesis, with aspartate amination by glutamine yielding asparagine and glutamate (catalysed by AsnB). The hydrolysis of asparagine, catalysed by asparaginase (AnsB), yields aspartate and NH_4^+ [269]. Mechanisms for both asparagine biosynthesis and catabolism are encoded by freshwater picocyanobacteria [103].

Aspartate can be further metabolised to aspartate-4-semialdehyde, the substrate for biosynthesis of threonine and lysine. The three-step reaction pathway for threonine generation (ThrA, ThrB, and ThrC) and the five-step pathway for lysine (DapA, DapB, DapL, DapF, and LysA) are all encoded by freshwater picocyanobacteria [103, 270, 271]. Threonine can be degraded through two pathways, both encoded by freshwater picocyanobacteria [103]. Threonine dehydratase (IlvA) catalyses the degradation of threonine, releasing NH_4^+ and α-ketobutyrate, an important intermediate in isoleucine biosynthesis. The secondary catabolic pathway is via the oxidation of threonine by threonine dehydrogenase (tdh), yielding 2-amino-3-oxobutanoate which is subsequently cleaved to form glycine and acetyl Co-A [272]. On the other hand, lysine catabolism is less well understood. Two pathways have been identified in bacteria, one degrading lysine via a saccharopine intermediate, the other via a cadaverine intermediate [273, 274]. However, genomic analysis of *Microcystis aeruginosa*, a Macrocyanobacterium often associated with cyanotoxin production, revealed neither of these two pathways for cyanobacterial lysine degradation [275].

Cyanobacterial methionine biosynthesis is unclear. In other bacteria, methionine is produced from an aspartate-4-semialdehyde precursor, yielding methionine through the action of MetA, MetB, and MetC [276]. However, cyanobacteria do not encode these enzymes, nor homologs, though the final reaction of methionine biosynthesis catalysed by methionine synthase (MetH) is present [189, 276]. This knowledge gap remains to be elucidated but may suggest an alternative precursor for methionine in cyanobacteria. Methionine degradation is primarily through conversion to S-adenosylmethionine, an important cofactor for biological methylation, and subsequent recycling to homocysteine catalysed by MetK and AhcY, both identified in cyanobacteria [189]. This lack of understanding over methionine biosynthesis resembles research into methionine bioavailability. Evidence is mixed, with methionine found to be a preferred source of organic matter in marine environments though also determined to be one of the least abundant AA among DON [277, 278].

1.7.3 Isoleucine, Valine, and Leucine Metabolism

The precursor for valine, leucine, and isoleucine biosynthesis is pyruvate while the latter additionally utilises α-ketobutyrate (generated through threonine catabolism) (Figure 1.7). All enzymes associated with these biosynthesis pathways are encoded by freshwater picocyanobacteria [103]. However, degradation pathways are not fully explored. Though common degradation pathways to acyl-CoA derivatives generally begin with a reversed final biosynthesis reaction [279], subsequent degradation mechanisms are understudied in cyanobacteria. While degradation pathways are not known, the abundance of valine and leucine in freshwater environments (isoleucine is routinely found at very low concentrations) indicates these AAs are potential nutrient sources [280, 281]. A greater understanding of their degradative pathways will be beneficial to understand picocyanobacterial ecology and underlying mechanisms behind DON nutrient uptake.

1.7.4 Aromatic AA (Tyrosine, Phenylalanine, and Tryptophan) Metabolism

The aromatic AAs are produced from a chorismate precursor, generated via the shikimate pathway [282]. The tryptophan biosynthetic pathway begins with the conversion of chorismate to anthranilate (TrpE/TrpG), with all enzymes conserved in picocyanobacteria and larger cyanobacteria [103, 189]. Though the first step of phenylalanine and tyrosine biosynthesis is the conversion of chorismate to prephenate (AroH), subsequent steps are not yet fully determined in cyanobacteria. For phenylalanine, the initial step from prephenate to phenylpyruvate is encoded (PheA) in picocyanobacteria, though the following conversion to phenylalanine is unknown. In contrast, the final step of tyrosine biosynthesis is encoded in picocyanobacteria (TyrA), though the enzyme responsible for transamination of prephenate to arogenate is as yet unclear [283]. Information of the cyanobacterial degradation of aromatic AAs is also sparse. Tryptophan may be degraded to 4-oxalocrotonate, and onwards to pyruvate and acetate while releasing NH_4^+, though this is yet to be demonstrated in cyanobacteria [284]. Phenylalanine may be degraded by its deamination to cinnamate catalysed by phenylalanine ammonia lyase (PAL), identified in two filamentous cyanobacteria [285]. Though many unknowns remain regarding aromatic AA biosynthesis and degradation, the presence of enriched phenylalanine in freshwater is linked to limnetic systems influenced by agriculture, suggesting increased concentrations of phenylalanine in areas of nutrient abundance [286].

1.7.5 Glycine, Serine, Cysteine, and Alanine Metabolism

Glyoxylate is a precursor molecule to glycine and serine [189]. The light-dependent photorespiratory pathway yields glycine via glyoxylate amination, subsequently converting glycine to serine by the reversible action of serine hydroxymethyltransferase (GlyA). Catabolism of glycine can be carried out directly through glycine oxidase, producing glyoxylate and NH_3, or decarboxylated by the glycine cleavage system yielding 5,10-methylenetetrahydrofolate, an important cofactor and an intermediate in one-C metabolism [287]. Serine can also be synthesised via a light-independent pathway from a glycerate-3-P precursor, which undergoes oxidation (SerA), amination (SerC), and dephosphorylation (SerB) to form serine [288]. The catabolism of serine is via serine deaminase (SdaA/SdaB), producing pyruvate and NH_4^+, though the presence of this pathway in cyanobacteria is not known. The importance of glycine as a putative N source is emphasized by its abundance in both freshwater and marine environments (glycine is commonly the most abundant AA in these DON pools) [278, 280, 289, 290], and its ability to be utilised as a single nutrient source by a variety of bacterial species [291].

Serine acts as a precursor for cysteine and alanine. Cysteine is synthesised through the acetylation (CysE) and sulfonation (CysK/CysM) of serine, while cysteine is desulfonated (NifS) to yield alanine [292, 293]. Alanine is also produced by the reductive amination of pyruvate, catalysed by alanine dehydrogenase (adh) [294]. This reaction is reversible, thus also yields pyruvate and NH_4^+ as alanine (and cysteine) degradation products. These enzymes are conserved in cyanobacteria [103, 189]. Though not generally found at the elevated concentrations of glycine, alanine and serine are significant components of the DON pool, especially pronounced in marine environments [277, 278]. Their abundance combined with short degradation pathways may indicate a favourable status as a nutrient source.

1.7.6 Histidine Metabolism

The biosynthesis pathway of histidine is well-conserved in cyanobacteria, forming histidine from the nucleotide precursor 5-phosphoribosyl-1-pyrophosphate [103, 189]. However, cyanobacterial histidine catabolism is unknown. A common pathway of histidine degradation is deamination to urocanate and NH_4^+, then subsequent conversion to formiminoglutamate. These reactions are catalysed by the histidine utilisation system (Hut) proteins, not encoded by cyanobacteria [295]. Histidine is present in the freshwater DON pool in low to moderate amounts, yet is N-rich (consisting of 3 N atoms) and thus may be nutritionally beneficial [280, 290]. However, the lack of Hut proteins and subsequent unknown method for histidine catabolism requires clarification to elucidate the bioavailable capability of histidine.

1.8 Summary

Freshwater picocyanobacteria are an ecologically significant group, globally abundant and key primary producers in both oligotrophic and eutrophic freshwater environments. Unlike their marine relatives, relatively little is known about the genomic capabilities of freshwater picocyanobacteria, with the non-picocyanobacterial *Synechococcus elongatus* strains historically dominating freshwater *Synechococcus* biological and ecological research. With picocyanobacteria dominating freshwater environments, a greater understanding of the factors (physical, morphological, biological) contributing to this is essential. The constraint of N limitation is often present in environments where picocyanobacteria thrive, suggesting the freshwater picocyanobacteria N uptake capabilities may contribute to their dominance. While the concentration of N strongly influences the cyanobacterial community, so do the molecular forms with which N is found. The bioavailability of DON has attracted greater interest in recent years, with AAs in particular providing diverse sources of N especially in oligotrophic environments. The increased sampling and sequencing of freshwater picocyanobacteria has improved our understanding of the genomic capabilities of this essential group. An investigation into the N assimilation machinery of freshwater picocyanobacteria will provide insights into their dominance and increase understanding of their ecologic role and capabilities.

1.9 Objectives

Objective 1: Increase the representation of freshwater picocyanobacteria by sequencing five strains and comparing these genomes to *Synechococcus elongatus*

Picocyanobacteria represent an essential source of freshwater biomass yet remain understudied, especially in comparison to marine environments. The scarcity of freshwater picocyanobacteria genomes has led to a reduced understanding of the ecology and molecular capabilities of this important group, and greater representation is necessary to elucidate their dominant role in freshwaters. In addition to increased genomic representation, the utility of *Synechococcus elongatus* as a model freshwater *Synechococcus* is questioned by the greater prevalence of *Syn/Pro* clade *Synechoccocus* picocyanobacteria. In Chapter 2, draft genomes are produced of five freshwater picocyanobacteria strains from the UK and the Netherlands. These genomes are compared to *Synechococcus elongatus*, both with a broad focus on genomic content and with in-depth investigation into photosynthetic and NO_3^-/NO_2^- assimilatory machinery. Differences between the *Syn/Pro Synechococcus* and *Synechococcus elongatus* are discussed, including the representativeness of *Synechococcus elongatus* as a model organism for freshwater *Synechococcus*.

Objective 2: Investigate the encoded N assimilation capabilities of freshwater picocyanobacteria, comparing these to marine picocyanobacteria, thermophylic picocyanobacteria, *Synechococcus elongatus*, and larger freshwater cyanobacteria

The uptake and assimilation of the important nutrient N is of great importance to picocyanobacteria productivity, however it is little understood compared to marine picocyanobacteria and larger freshwater cyanobacteria. Recent increased sampling of freshwater picocyanobacteria, aided by strains sequenced in Chapter 2, have increased the representation of this keystone group, and enabled greater analysis of their genomic features. The encoded molecular machinery associated with N assimilation is investigated in Chapter 3 and compared to that of picocyanobacteria in other habitats, and larger freshwater cyanobacteria, especially the sister group of freshwater picocyanobacteria, *Synechococcus elongatus*. A particular focus is on AA uptake. These molecules are a ubiquitous component of DON, regularly constituting a large percentage of bioavailable DON yet they remain under-researched, particularly in comparison to other forms of organic N like urea. This work aims to increase our understanding of how freshwater picocyanobacteria import sufficient N to thrive in oligotrophic environments, and discusses the evolutionary adaptations associated with their environment.

Objective 3: Investigate the capability of freshwater picocyanobacteria to utilise AAs as sole N sources and explore the proteomic response to growth on glutamate, proline, arginine, and asparagine

The results from Chapter 3 prompted our investigation into the AA uptake capabilities of freshwater picocyanobacteria. A diverse set of AATs are found in freshwater picocyanobacteria, with their specific set of encoded AATs not found in larger freshwater cyanobacteria. We hypothesised that the AATs present in a genome may determine the AAs which can act as sole N sources. To investigate this, a growth assay featuring 20 proteinogenic AAs and novel organic substrates was carried out for two strains of freshwater picocyanobacteria. This assessed the ability of freshwater picocyanobacteria to import these AAs, and to relate this back to the AATs which they encode. We further wanted to investigate how picocyanobacteria respond to growth on AA substrates. Utilising quantitative proteomic analysis, the proteomic response to picocyanobacterial growth on arginine, asparagine, glutamate, and proline compared to NO_3^- was explored and differentially expressed pathways identified.

Chapter 2

Draft Genome Sequences of Five Freshwater *Syn/Pro* Clade Picocyanobacteria

2.1 Contributions and Acknowledgements

This chapter has been adapted from an article published in the Journal of Genomics [34]. The figures, tables, methods, and results/discussion remain the same. The introduction (section 2.2) has been expanded to provide more background to *Synechococcus* strains and introduce genome assembly from non-axenic cultures. Furthermore, the conclusion (section 2.5), limitations (section 2.6), and future directions (section 2.7) paragraphs are new to this chapter. The submitted article was edited with contributions from Patricia Sánchez-Baracaldo, Henk Bolhuis, Michele Grego, and Penny Johnes. The full reference of the published article is below:

Druce, E., Grego, M., Bolhuis, H., Johnes, P. J. & Sánchez-Baracaldo, P. 2023. Draft Genome Sequences of *Synechococcus* sp. strains CCAP1479/9, CCAP1479/10, CCAP1479/13, CCY0621, and CCY9618: Five Freshwater *Syn/Pro* Clade Picocyanobacteria. Journal of Genomics, 11, 26-36.

All sequenced strains were received from culture collections. Henk Bolhuis and Michele Grego sent *Synechococcus* spp. CCY0621 and CCY9618 from Culture Collection Yerseke (CCY) while *Synechococcus* spp. CCAP1479/09, CCAP1479/10, and CCAP1479/13 were sent from the Culture Collection of Algae and Protozoa (CCAP). Jane Coghill and Christy Waterfall at the Bristol Genomics Facility carried out genome sequencing. Giorgio Bianchini assisted with phylogenomic tree construction. All other work, including DNA extraction, genome assembly, and comparative genomic analysis, is my own.

2.2 Introduction

Picocyanobacteria play a key role in aquatic ecosystems, contributing a significant proportion of total primary production in both marine and fresh waters [25, 296, 297]. These unicellular cyanobacteria, sized between 0.5 and 2 μm, are distributed globally, from temperate and tropical open oceans to alpine lakes and eutrophic reservoirs [298–300]. Freshwater picocyanobacteria are predominately *Synechococcus* strains which can dominate the picophytoplankton component (1 - 99% [70]) and total biomass (10 - 70% [27]) depending on trophic status and depth [52, 122]. Other taxonomic names associated with freshwater picocyanobacterial strains are *Cyanobium* spp. [301] and *Vulcanococcus* spp. [33].

For decades, *Synechococcus elongatus* strains have been the primary focus for cyanobacterial molecular and ecological research. *Synechococcus elongatus* PCC 7942 (isolated from a pond in Berkeley) was the first cyanobacterium to be transformed by exogenous DNA [302], while the first complete genome sequence of *Synechococcus* was achieved for *Synechococcus elongatus* PCC 6301 (isolated from Waller Creek, Texas) [303]. In addition to their role in pioneering genomic research, the light-harvesting mechanisms in cyanobacterial photosynthesis were explicated through characterisation of the phycobilisome (PBS) in *Synechococcus elongatus* PCC 6301 [304]. The popularity of *Synechococcus elongatus* strains has further led to their use as model organisms to investigate the interactions between freshwater cyanobacteria and the surrounding environment. This includes the molecular characterisation of carbon (C) and nitrogen (N) assimilation systems and their regulation [209, 234, 305], the response to nutrient deprivation and stresses [306–309], and the ability to adapt to variations in light intensity and temperature [310–313]. Though *Synechococcus elongatus* cells are larger than those of the *Syn/Pro* clade *Synechococcus* [314, 315], and do not fall under the 'pico-' threshold (¡ 2 μm), they are often used as models for freshwater picocyanobacteria [221, 223, 229, 233, 258]. However, the emergence of the

Synechococcus elongatus strains as a deep branching sister group to the monophyletic *Syn/Pro* clade suggests *Synechococcus elongatus* provides an unrepresentative model of freshwater picocyanobacteria and freshwater *Synechococcus* [316].

A sequenced genome presents a myriad of opportunities to understand how an organism has adapted to its environment. Analysing genomes can reveal unknown nutrient preferences, mechanisms for essential life processes, and responses to environmental stressors. Of special interest in cyanobacteria are two essential metabolic pathways - photosynthesis and N metabolism. Due to their distinction as the only prokaryote capable of performing oxygenic photosynthesis and their role in the Great Oxidation Event (2.4 billion years ago [317]), considerable research interest has focused on cyanobacterial photosynthesis, both in understanding the function of the photosynthetic apparatus and elucidation of the early evolution of photosynthesis [318, 319]. On the other hand, N (alongside phosphorus) is often the limiting factor for cyanobacterial growth in aquatic systems [9]. The sequenced genomes of *Synechococcus elongatus* strains have revealed insights into the capabilities of mitigating N-starvation in freshwater *Synechococcus* (i.e. identification of a cyanate transporter for nitrite assimilation [234, 320]), though significant differences between *Synechococcus elongatus* and *Syn/Pro* strains have already been identified, such as differential capabilities of urea utilisation and the presence of N-fixing capability in a freshwater picocyanobacterium [33, 184].

The availability of sequenced freshwater *Synechococcus* genomes from the *Syn/Pro* clade has lagged behind that of marine picocyanobacteria [108]. The first marine *Synechococcus* and *Prochlorococcus* genomes became available two decades ago [48, 232, 321], and since then many studies have utilised genomics to understand the ecology and evolution of marine picocyanobacteria, ranging from the evolutionary mechanisms behind niche adaptation to the diversity of PBSs among marine *Synechococcus* [10, 322–324]. Only recently has attention turned towards freshwater picocyanobacteria, with representatives sequenced from Italy [33], Spain [301], and Central and South America [108], before a substantial global increase through the efforts of Cabello-Yeves *et al.* [103]. This increased representation of *Syn/Pro* clade freshwater picocyanobacteria will aid in reducing the knowledge gap between marine and freshwater picocyanobacterial environmental adaption, itself enlarged due to the heterogenous nature of freshwater environments [325], and determine further genomic distinctions between *Syn/Pro* freshwater picocyanobacteria and *Synechococcus elongatus* strains.

Genome sequencing has traditionally relied on first obtaining an axenic culture, a time-consuming task for many cyanobacteria with strongly associated heterotrophic bacteria [326]. However, the rise of next-generation sequencing and advancements in genome assembly have made genomes cheaper, quicker, and easier to decipher while providing new possibilities for the separation of a desired genome from a pool of multi-genomic reads [327]. The first barrier to overcome in non-axenic cyanobacterial genome sequencing is the generation of a mono-phototrophic culture [328]. This is generally much easier than achieving an axenic culture, disregarding the requirement for complete contamination elimination, instead focusing on growth of the single desirable species for DNA extraction and sequencing. Once sequenced, the primary challenge is separation of cyanobacterial sequences from sequences originating from the associated heterotrophic bacteria. This can be achieved through the utilisation of metagenome binning for heavily contaminated samples or analysis of de Bruijn graphs from genome assembly [327, 329]. Incorporated into genome assembler software (such as SPAdes [330]), de Bruijn graphs generate assemblies by tracing a path through reads (nodes) of k-mer length (i.e. splitting a 150-nucleotide read into 86 overlapping 65-mers) based on overlapping sequences [331]. This would ideally provide a single chromosome displaying one path between contigs, however duplicated genes and repetitive DNA sequences result in more complex contig alignments. Though not assembled into a single chromosome, these multiple alignments can aid in identifying relationships between contig clusters, especially when paired with other contig characteristics such as sequence coverage and presence of expected taxon-specific genes.

To increase freshwater *Syn/Pro* clade genomic representation and identify genetic diversity between these and *Synechococcus elongatus* strains, we sequenced five picocyanobacteria held in culture collections, originally isolated from European lakes. *Synechococcus* sp. CCY0621 (Leiden) and *Synechococcus* sp. CCY9618 (Vinkeveen) were isolated from the Netherlands, and *Synechococcus* sp. CCAP1479/9, *Synechococcus* sp. 1479/10, and *Synechococcus* sp. 1479/13 were isolated from Windermere in the UK. Phylogenomic analyses revealed their positions in the *Syn/Pro* clade, while their genome content was compared with *Synechococcus elongatus* strains, especially with regards to photosynthetic machinery and nitrate (NO_3^-) metabolism where multiple differences were identified.

2.3 Methodology

2.3.1 Culture Collection and Maintenance

Three *Synechococcus* strains were obtained from the Culture Collection of Algae and Protozoa: *Synechococcus* sp. CCAP1479/9, *Synechococcus* sp. CCAP1479/10, and *Synechococcus* sp. CCAP1479/13, all isolated from Windermere, UK. Two *Synechococcus* strains were obtained from the Culture Collection Yerseke: *Synechococcus* sp. CCY0621 and *Synechococcus* sp. CCY9618, isolated from ponds in the Netherlands (Leiden and Vinkeveen respectively) (Appendix Figure A.1). All strains were mono-phototrophic and grown in 25 cm^2 vented flat-sided cell culture flasks with BG-11 medium [332]. An LMS Series 2 Cooled Incubator (Appleton Woods, UK) was used at 20 °C with 10-20 µmol m^{-2} s^{-1} of white light under a 16 h: 8 h light:dark cycle.

2.3.2 Genome Extraction

Aliquots of 1.8 mL of each mono-phototrophic culture were harvested to extract genomic DNA using DNeasy UltraClean Microbial Kits (Qiagen, Germany) according to the manufacturer's instructions. Once purified, genomic DNA was stored at -80 °C in 10 mM Tris buffer at pH 8. DNA concentration and quality was measured using a NanoDrop 2000 spectrophotometer (Thermo Scientific, USA) and a Qubit 2.0 Fluorometer (Thermo Scientific, USA).

2.3.3 DNA Sequencing and Genome Assembly

Whole genome library preparation and sequencing was carried out by the University of Bristol Genomics Facility, UK. DNA libraries were prepared for each strain using Truseq Nano LT Kit (Illumina, USA) and sequenced using Illumina NextSeq 500/550 Mid Output Kit v2 (300 cycles) (Illumina, USA) to generate paired-end reads (2x150 bps). Raw reads were trimmed using Trimmomatic v0.39 [333] with parameters Leading: 20, Trailing: 20, SlidingWindow:4:20, MinLen: 20, and assembled de novo using SPAdes v3.14.1 [330] with k-mers of 67, 77, 87, 97 and a coverage cutoff of 20 in −careful mode. A BLAST [334] database was generated at the amino acid sequence level for each assembly and searched against a collection of 1,054 core cyanobacterial genes (CCGs) [335]. Bandage v0.8.1 [336] was used to visualise strain assemblies and separate out cyanobacterial sequences based on contiguous CCGcontaining nodes as demonstrated in previous assemblies [329] (Appendix figures A.2 - A.6). Contigs which did not contain cyanobacterial genes were discarded, in addition to short (<200 bp) contigs. The assembled genomes had overall coverages ranging from 552x to 939x (Table 2.1) and were structurally annotated with GeneMark.hmm-2 v1.05 [337], Prodigal v2.6.3 [338], INFERNAL v1.1.2 [339], and tRNAscan-SE v2.0.5 [340]. Genome completeness was estimated by identifying cyanobacteria-specific single-copy orthologous genes using BUSCO v3.0.2 [341]. The draft genomes were submitted to JGI IMG/ER [342] (GOLD Analysis Project IDs: Ga0436386, Ga0436387, Ga0436388, Ga0436389, and Ga0436390). The five draft genomes were deposited to the DDBJ/Genbank/ENA repositories with accession numbers JAFKRG000000000 (CCY9618), JAFKRH000000000 (CCY0621), JAFKRI000000000 (CCAP1479/13), JAFKRJ000000000 (CCAP1479/10), and JAFKRK000000000 (CCAP1479/9). The average nucleotide identity (ANI) between these newly sequenced strains is shown in Appendix Table B.1.

2.3.4 Comparative Genomics

Functional annotation was determined through the eggNOG web server [343]. Two-tailed t-tests were applied to carry out statistical analysis on total COG numbers and COGs normalised as a proportion of total genome. JGI IMG/ER was used to carry out KEGG [344] comparative genomic analysis for photosynthesis and NO_3^- metabolism pathways between *Synechococcus elongatus* (*Synechococcus elongatus* PCC 7942, *Synechococcus elongatus* UTEX 2973, *Synechococcus elongatus* PCC 6301, *Synechococcus elongatus* FACHB-242, *Synechococcus elongatus* FACHB-1061) (average ANI pairwise for all strains > 99 %) and the sequenced *Synechococcus* strains.

2.3.5 Phylogenomic Analysis

The evolutionary relationships of the newly sequenced strains with a selection of cyanobacterial taxa sampling a broad range of morphologies, lifestyles, and metabolisms, were estimated through phylogenomic analysis. Our dataset included 373 cyanobacteria genomes (no metagenome-assembled genomes) and ortholog sequences from 143 protein-coding genes (the latter based on previously published studies [94, 97, 98]). We performed BLAST searches with these ortholog sequences against the 373 genomes using BLASTP v2.11.0+ [345] with an E-value threshold of 10^{-5}, retaining the hit with the highest score and extracting the corresponding protein sequences. The resulting sequences were aligned using MAFFT v7.511 [346] with the –localpair –maxiterate 1000 parameters. Maximum-likelihood gene trees were constructed using IQ-TREE 2.2.0 [347], implementing the *LG* protein evolution model and the –fast option. These gene trees were used to identify the clusters of sequences that were most closely associated with the BLAST query sequences – these clusters were assumed to be ‘true’ orthologs. These true orthologs were re-aligned with MAFFT (same parameters as above) and inspected with mis-aligned columns and alignment positions with a gap content higher than 80% removed from each alignment. The best evolutionary model for each gene was determined by using IQ-TREE with the -m MF option [348], selecting the model with the lowest BIC score. A maximum-likelihood partitioned phylogenomic analysis was performed using IQ-TREE [349]. Using the previously determined evolutionary models, partitioned analysis was carried out with IQ-TREE using -p and -B 1000 parameters with each gene assigned to its own partition. The -p option constrains all partitions to the same topology and branch length but allows each partition to have a different overall evolutionary rate, while -B 1000 produces ultrafast bootstrap support values [350]. This analysis was carried out twice with the two resulting trees compared to confirm no significant differences between them.

2.4 Results and Discussion

2.4.1 Genome Statistics

The newly sequenced picocyanobacteria genomes consist of 88 to 133 contigs (average of 112) and range in size from 2.9 Mbps to 3.3 Mbps (average of 3.2 Mbps), significantly larger than *Synechococcus elongatus* strains ($p < .001$). *Synechococcus* sp. CCY9618 has the smallest genome and is composed of

Table 2.1: **Genomic features of the sequenced freshwater picocyanobacteria**

	***Synechococcus* sp. CCAP1479/9**	***Synechococcus* sp. CCAP1479/10**	***Synechococcus* sp. CCAP1479/13**	***Synechococcus* sp. CCY0621**	***Synechococcus* sp. CCY9618**
Genome size (bp)	3,288,920	3,313,705	3,299,582	3,230,971	2,927,161
Contigs	88	108	132	101	133
N50 (bp)	207,208	151,487	78,719	105,719	94,487
Genome coverage	825X	939X	552X	818X	865X
DNA coding (%)	92	92	92	91	91
DNA G+C (%)	69	69	69	67	68
Total genes	3,423	3,502	3,507	3,471	3,165
Protein encoding genes	3,364	3,441	3,446	3,407	3,109
Completeness (%)	98	99	98	99	99
Average Nucleotide Identity to *Synechococcus elongatus* PCC 7942	73.5808	73.4985	73.5014	73.3739	73.3276

the largest number of contigs, with an N50 value of 94,487 (Table 2.1). *Synechococcus* sp. CCAP1479/10 has the largest genome, while *Synechococcus* sp. CCAP1479/9 contains the fewest contigs (88) and the largest N50 (207,208). Genome coverage is high among the assemblies (552x − 939x) with genome completeness estimated at 98.2 – 98.7%. It should be noted that these genomes have not been completely closed yet a high genome completeness suggests that the 'missing' part of the genome is limited.

All five genomes contain high GC contents ranging from 67.45 - 69.36% (Table 2.1). This is consistent with previously sequenced freshwater picocyanobacteria and *Synechococcus elongatus*, regularly featuring a high (>60%) GC content [103, 108]. Compared to marine *Syn/Pro* strains, freshwater *Synechococcus* have significantly larger genome sizes ($p < .001$; primarily due to genomic streamlining of *Prochlorococcus* spp. [351]) and higher GC content ($p < .001$) (Figure 2.1). Meanwhile, the trend of increasing GC content with increasing genome size present in freshwater and marine *Synechococcus* is not found in larger cyanobacteria (cell size greater than 2 μm). Higher genomic GC contents have been linked with increased horizontal gene transfer and protection against DNA damage through higher resilience against UV irradiation, contributing to picocyanobacterial genomic plasticity and environmental adaptability [352, 353]. Conversely, lower GC contents in marine picocyanobacteria may indicate selection in N limited environments due to the reduced N requirement for AT pairs [354].

2.4.2 Phylogenomic Analysis

Phylogenomic analysis was carried out to identify the closest relatives of the newly sequenced freshwater picocyanobacteria. All five strains belong to the *Cyanobium* and *Synechococcus* freshwater sub-cluster 5.2 of the *Syn/Pro* clade (Figure 2.2). *Synechococcus* sp. CCAP1479/10, *Synechococcus* sp. CCAP1479/13, and *Synechococcus* sp. CCAP1479/13 form a monophyletic clade, with *Synechococcus* sp. BO8801 (Lake Constance, Germany) and *Synechococcus* sp. FACHB-909 (Baohu Lake, China) the closest related strains (a sister group to these three newly sequenced picocyanobacteria). *Synechococcus* sp. CCY0621 and *Synechococcus* sp. CCY9618 are more distantly related and appear as outgroups to the CCAP newly sequenced strains. In contrast, *Synechococcus* elongatus strains are a sister group of the *Syn/Pro.*

Interestingly, the phylogenomic tree produced in this study differs in the placement of the Nodosilin-

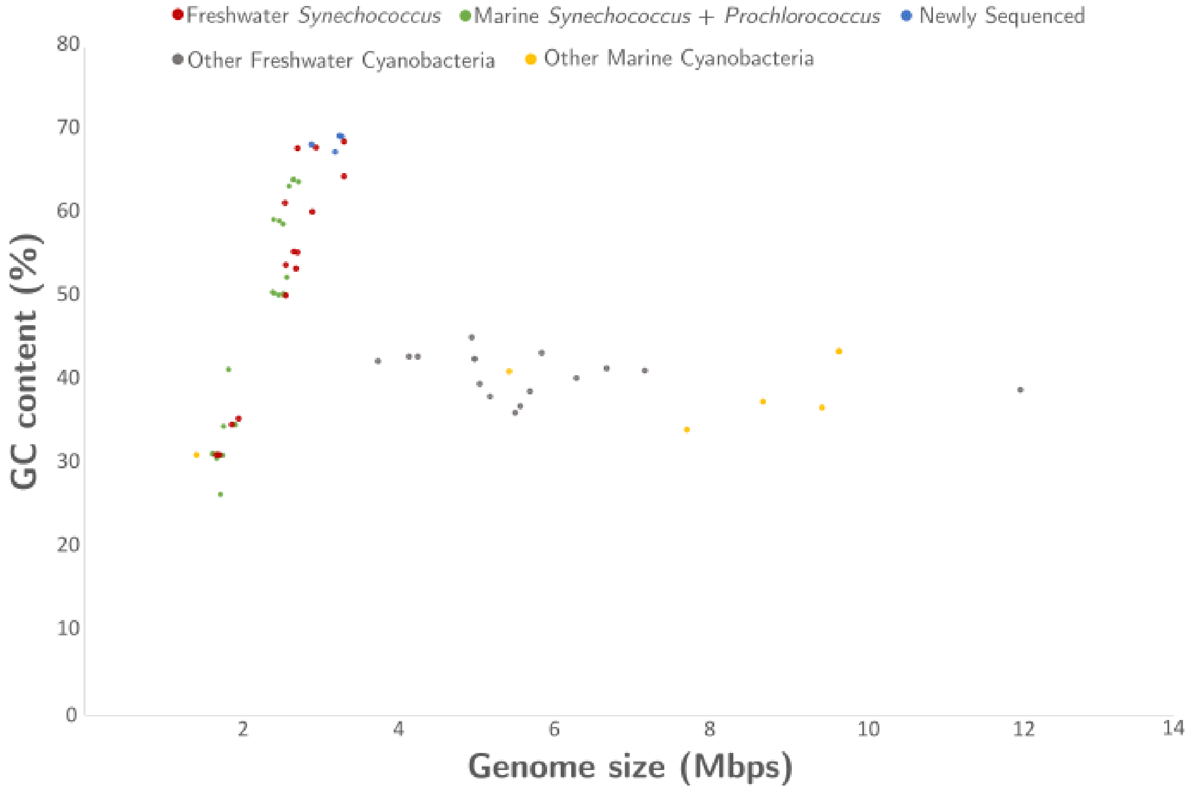

Figure 2.1: **GC content and genome size of cyanobacteria characterised into habitat and phylogeny**. Freshwater *Synechococcus* include picocyanobacteria found in the *Syn/Pro* clade and *Synechococcus elongatus* strains. The newly sequenced strains are clustered with the freshwater *Synechococcus*.

eales compared to other phylogeneomic analyses (Appendix Figure A.7) [94, 104, 329]. The Nodosilineales are often found as a sister group to the *Syn/Pro* + *Prochlorothrix* clade, however in this analysis the Nodosilineales are a sister clade to the Macrocyanobacteria. There are two reasons for this. Firstly, the bootstrap support for the Nodosilineales as a sister group to the *Syn/Pro* is low in some analyses (i.e. 50% support [104]), questioning the 'true' location of the Microcyanobacteria. Secondly, this study focuses on the *Syn/Pro*, over-representing taxa to improve the resolution of this clade. This may have introduced artefacts into the surrounding taxa, namely the positioning of the Nodosilineales.

2.4.3 Comparative Genomics

Freshwater picocyanobacteria from the *Syn/Pro* clade are derived taxa that specialised in a planktonic habitat. The newly sequenced genomes were functionally annotated with eggNOG and KEGG, in addition to five *Synechococcus elongatus* genomes (*Synechococcus elongatus* PCC 7942, *Synechococcus*

Figure 2.2: **Maximum likelihood phylogeny showing the relationship of *Synechococcus* sp. CCAP1479/9, *Synechococcus* sp. CCAP1479/10, *Synechococcus* sp. CCAP1479/13, *Synechococcus* sp. CCY0621, and *Synechococcus* sp. CCY9618 within the *Syn/Pro* clade**. New sequenced picocyanobacteria are highlighted in red. The tree was constructed from 373 cyanobacteria and 143 orthologous proteins. Bootstrap values less than 100 are displayed at branching nodes while blank nodes have a support of 100. The tree is rooted using *Melainabacteria* spp. as an outgroup. An expanded tree is shown in Appendix Figure A.7.

elongatus UTEX 2971, *Synechococcus elongatus* PCC 6301, *Synechococcus elongatus* FACHB-242, and *Synechococcus elongatus* FACHB-1061). This enabled insights into the genomic capabilities of the scarcely researched freshwater sub-cluster 5.2 of the *Syn/Pro* clade compared to *Synechococcus elongatus*.

Table 2.2: **Number of eggNOG classifications of proteins encoded by the five sequenced sub-cluster 5.2 *Synechococcus* genomes and five selected *Synechococcus elongatus* strains.** Percentage of genes as proportion of the genome is provided in brackets. J: Translation, ribosomal structure and biogenesis; K: Transcription; L: Replication, recombination and repair; B: Chromatin structure and dynamics; D: Cell cycle control, cell division, chromosome partitioning; V: Defence mechanisms; T: Signal transduction mechanisms; M: Cell wall/membrane/envelope biogenesis; N: Cell motility; U: Intracellular trafficking, secretion, and vesicular transport; O: Posttranslational modification, protein turnover, chaperones; C: Energy production and conversion; G: Carbohydrate transport and metabolism; E: Amino acid transport and metabolism; F: Nucleotide transport and metabolism; H: Coenzyme transport and metabolism; I: Lipid transport and metabolism; P: Inorganic ion transport and metabolism; Q: Secondary metabolites biosynthesis, transport and catabolism; S: Function unknown.

COG	***Synechococcus* sp. CCAP1479/9**	***Synechococcus* sp. CCAP1479/10**	***Synechococcus* sp. CCAP1479/13**	***Synechococcus* sp. CCY0621**	***Synechococcus* sp. CCY9618**	***Synechococcus elongatus* PCC 7942**	***Synechococcus elongatus* UTEX 2973**	***Synechococcus elongatus* PCC 6301**	***Synechococcus elongatus* FACHB-242**	***Synechococcus elongatus* FACHB-1061**
J	**163** (4.7)	**161** (4.7)	**161** (4.7)	**161** (4.7)	**160** (5.1)	**163** (6.1)	**166** (6.1)	**165** (6.5)	**167** (6)	**167** (6)
K	**161** (4.8)	**155** (4.5)	**154** (4.5)	**138** (4.1)	**116** (3.7)	**114** (4.3)	**115** (4.2)	**110** (4.4)	**114** (4.1)	**114** (4.1)
L	**116** (3.4)	**121** (3.5)	**122** (3.5)	**127** (3.7)	**146** (4.7)	**108** (4.1)	**116** (4.3)	**112** (4.4)	**114** (4.1)	**114** (4.1)
B	**2** (0.1)	**2** (0.1)	**2** (0.1)	**2** (0.1)	**2** (0.1)	**2** (0.1)	**2** (0.1)	**2** (0.1)	**2** (0.1)	**2** (0.1)
D	**35** (1)	**42** (1.2)	**42** (1.2)	**37** (1.1)	**28** (0.9)	**27** (1)	**27** (1)	**25** (1)	**27** (1)	**27** (1)
V	**41** (1.2)	**43** (1.2)	**43** (1.2)	**51** (1.5)	**47** (1.5)	**31** (1.2)	**31** (1.1)	**31** (1.2)	**31** (1.1)	**31** (1.1)
T	**92** (2.7)	**100** (2.9)	**98** (2.8)	**89** (2.6)	**62** (2)	**85** (3.2)	**117** (4.3)	**109** (4.3)	**117** (4.2)	**117** (4.2)
M	**199** (5.9)	**201** (5.8)	**203** (5.9)	**204** (6)	**186** (6)	**134** (5)	**148** (5.4)	**145** (5.7)	**148** (5.3)	**148** (5.3)
N	**19** (0.6)	**21** (0.6)	**20** (0.6)	**17** (0.5)	**18** (0.6)	**26** (1)	**26** (1)	**26** (1)	**26** (0.9)	**25** (0.9)
U	**71** (2.1)	**73** (2.1)	**75** (2.2)	**71** (2.1)	**62** (2)	**34** (1.3)	**55** (2)	**53** (2.1)	**55** (2)	**54** (2)
O	**111** (3.3)	**114** (3.3)	**115** (3.3)	**114** (3.3)	**108** (3.5)	**94** (3.5)	**99** (3.6)	**101** (4)	**100** (3.6)	**100** (3.6)
C	**220** (6.5)	**229** (6.7)	**230** (6.7)	**220** (6.5)	**224** (7.2)	**199** (7.5)	**199** (7.3)	**202** (8)	**198** (7.2)	**199** (7.2)
G	**112** (3.3)	**111** (3.2)	**113** (3.3)	**104** (3.1)	**103** (3.3)	**71** (2.7)	**82** (3)	**80** (3.2)	**82** (3)	**82** (3)
E	**181** (5.4)	**184** (5.3)	**181** (5.3)	**174** (5.1)	**163** (5.2)	**138** (5.2)	**137** (5)	**132** (5.2)	**138** (5)	**138** (5)
F	**86** (2.6)	**89** (2.6)	**89** (2.6)	**88** (2.6)	**83** (2.7)	**98** (3.7)	**97** (3.6)	**96** (3.8)	**98** (3.5)	**98** (3.5)
H	**186** (5.5)	**183** (5.3)	**185** (5.4)	**185** (5.4)	**181** (5.8)	**167** (6.3)	**171** (6.3)	**171** (6.8)	**173** (6.3)	**173** (6.3)
I	**79** (2.3)	**80** (2.3)	**79** (2.3)	**88** (2.6)	**70** (2.3)	**54** (2)	**54** (2)	**51** (2)	**54** (2)	**54** (2)
P	**142** (4.2)	**151** (4.4)	**151** (4.4)	**162** (4.8)	**122** (3.9)	**152** (5.7)	**163** (6)	**157** (6.2)	**163** (5.9)	**164** (5.9)
Q	**46** (1.4)	**45** (1.3)	**45** (1.3)	**46** (1.4)	**39** (1.3)	**27** (1)	**44** (1.6)	**42** (1.7)	**44** (1.6)	**44** (1.6)
S	**669** (19.9)	**711** (20.7)	**710** (20.6)	**704** (20.7)	**629** (20.2)	**572** (21.5)	**737** (27.1)	**715** (28.3)	**745** (26.9)	**742** (36.8)

2.4.3.1 Overall Genome Content

Of the 19 functional COG categories identified, 11 categories differed significantly between our sequenced genomes and *Synechococcus elongatus* strains, in terms of total gene number and genes as a percentage of the total genome (Table 2.2, Appendix Tables B.3 and B.4).

Five of these categories were found to be significantly increased in our sequenced genomes (defence mechanisms (V), cell wall/membrane/envelope biogenesis (M), carbohydrate transport and metabolism (G), amino acid transport and metabolism (E), lipid transport and metabolism (I)), while three were significantly decreased (translation, ribosomal structure and biogenesis (J), cell motility (N), nucleotide transport and metabolism (F)). The total number of genes associated with three categories (post-translational modification, protein turnover (O), energy production and conversion (C), coenzyme transport and metabolism (H)) were significantly greater in our sequenced genomes ($p < .001$), though as a proportion of their genome were significantly greater in *Synechococcus elongatus* strains ($p = .006$, $p = .007$, $p < .001$ respectively). Additionally, KEGG analysis revealed 1,425 KO terms within at least one of the sub-cluster 5.2 freshwater picocyanobacteria of which 183 terms were not identified in *Synechococcus elongatus* strains. Meanwhile, 162 KO terms are found in *Synechococcus elongatus* but absent from our newly sequenced strains (Appendix Table B.5).

Our sequenced sub-cluster 5.2 strains encode significantly more genes involved in carbohydrate (G), amino acid (E), and lipid (I) transport and metabolism than *Synechococcus elongatus* strains ($p < .001$). Conversely, *Synechococcus elongatus* strains encode significantly more nucleotide transport and metabolism genes (F; $p < .001$). As the *Synechococcus elongatus* genome size is smaller than that of our sub-cluster 5.2 freshwater strains, it may be expected to encode a reduced number of nucleotide-associated genes, though this is not found. These genomic differences may be caused by the different environmental niches these two clades inhabit. Fresh waters are spatially diverse and exhibit a greater amount of nutrient heterogeneity than ocean environments [6]. Multiple other factors contribute to freshwater habitat niches, including light availability, temperature, water retention time, and composition of the surrounding microbial community [15]. However, while sub-cluster 5.2 and *Synechococcus elongatus* strains have been isolated from geographically distant locations, they occupy the same position in the water column (limnetic zone [108]) and are more dominant in temperate waters (though the full biogeographic distribution of *Synechococcus elongatus* is not clear). Increased genomic sequencing of taxa from sub-cluster 5.2 will aid in understanding freshwater picocyanobacteria ecology and the evolutionary context of these divergent lineages.

Further differences have been identified in the number of genes responsible for information storage and cellular processes between our sequenced strains and *Synechococcus elongatus* strains. Genes encoding defence mechanisms (V) and cell wall biogenesis-related (M) proteins are significantly increased in our newly sequenced strains ($p < .05$). Meanwhile, *Synechococcus elongatus* strains have significantly higher numbers of genes involved in translation (J) and cell motility (N) ($p < .001$). Research on cyanobacterial chemo- and photo-taxis has focused on *Synechocystis* spp. which exhibit a 'gliding' form of motility utilising a type IV pilus system [355]. Motility among marine *Synechococcus* spp. is achieved through multiple mechanisms, the most common through S-layer rotation [356, 357], while recent findings have identified phototactic behaviour in *Synechococcus elongatus* [358]. However, the motility of sub-cluster 5.2 is yet to be determined. These differences in core cellular control may represent subtle changes in clade behaviour. As *Synechococcus elongatus* PCC 7942 is traditionally used as a model for freshwater *Synechococcus*, the genotype and associated phenotypes for these strains may distort expectations of the *Syn/Pro* clade due to genomic differences between these two groups.

2.4.3.2 Photosynthesis Machinery

A comparison of the photosynthesis pathway between the newly sequenced picocyanobacteria and *Synechococcus elongatus* reveals a number of differences. Among core Photosystem II (PSII) components, the gene for the D2 protein (*psbD*) is surprisingly absent from the newly sequenced strains (in addition to two recently sequenced *Synechococcus elongatus*) (Table 2.3). The D2 protein forms part of the PSII reaction core alongside D1 (encoded by *psbA*) and is essential in binding the necessary redox-active cofactors for electron transfer [359]. The presence of *psbD* in other sub-cluster 5.2 strains is likewise unclear - absent from *Synechococcus* sp. BO8801 yet found in *Synechococcus* sp. 1G10 and *Cyanobium gracile* PCC 6307 (data not shown). However, *psbD* is essential for photosynthetic growth, suggesting the absence of *psbD* from our sequenced picocyanobacteria is a result of the unclosed nature of the genome. *psbC* is found clustered with *psbD* in other cyanobacteria (e.g. *Synechococcus elongatus* PCC 7942 and *Synechocystis* sp. PCC 6803), though the contig encoding *psbC* in our sequenced *Synechococcus* spp. is truncated upstream (where the *psbD* locus is usually found). Other genes encoding photosynthesis electron transport proteins that are absent from our newly sequenced sub-cluster 5.2 strains include *petL*, encoding the cytochrome b6f complex subunit 6, and *petE* encoding plastocyanin, responsible for transferring electrons from cytochrome b6f to Photosystem I (PSI). Cytochrome b6f is an intermediate in the transport of electrons from PSII to PSI, however the role of PetL in the complex is unclear. A function linked to stability of the dimeric state of the cytochrome b6f complex has been suggested while the non-essential nature of PetL in cyanobacteria has been demonstrated [360, 361]. Accepting electrons from cytochrome b6f, copper-containing plastocyanin is another essential component of the photosynthesis electron transport chain. However, most cyanobacteria also contain Fe-containing cytochrome c6 (encoded by *petJ*). Expression of these two electron carriers is regulated by copper availability, a response to Fe-limitation [362]. Plastocyanin absence in sub-cluster 5.2 strains appears to reduce adaptability in low-Fe environments, though heterocyst-forming cyanobacteria have been shown to preferentially utilise cytochrome c6 for electron transport, even in the presence of copper [363]. Further notable differences include the presence of *psb28-2* (a PSII assembly factor with a distinct function from *psb28-1* [364]) only in *Synechococcus elongatus* and an additional copy of *psaI* (encoding a PSI transmembrane protein [111]) in the newly sequenced strains (Table 2.3). While the deletion of *psbD* is an artefact and must be resolved by the generation of closed freshwater picocyanobacteria genomes, further research to investigate the impact of the putative *petL* and *petE* gene deletions is necessary to elucidate this key physiological process in freshwater picocyanobacteria.

In addition to core photosynthetic electron transport apparatus, the copy number and composition of antennae proteins comprising the light-harvesting PBS displays subtle differences (Table 2.3). *Synechococcus elongatus* strains encode two copies of *apcD*, encoding a key component of the allophycocyanin (AP) central core of PBS, though our newly sequenced sub-cluster 5.2 picocyanobacteria encode solely *apcD1*. The role of ApcD has been shown to slightly vary between *Synechococcus elongatus* PCC 7942 and another cyanobacterial model organism - *Synechocystis* sp. PCC 6803. ApcD is vital for efficient energy transfer from the PBS to PSI in *Synechococcus elongatus* PCC 7942 while the lack of ApcD has no impact on PSI energy transfer in *Synechocystis* sp. PCC 6803, instead inhibiting state transitions in response to unbalanced light conditions [365]. Furthermore, multiple copies of *apcD* have been linked to photoacclimation to far-red light (700 - 750 nm) [366] and low light conditions [367].

Since *Synechococcus elongatus* is not known to be capable of far-red absorption, it's more likely that the additional encoding of *apcD* may result in the enhanced capability of low light photoacclimation. This may suggest greater adaptability in low light conditions for *Synechococcus elongatus* strains, potentially resulting in community shifts in heavily shaded areas.

Table 2.3: **Genes encoding photosynthesis machinery and antennae proteins found in the five sequenced sub-cluster 5.2 *Synechococcus* genomes and five selected *Synechococcus elongatus* strains.** Genes were identified through KEGG annotation. Copy number is indicated by the number of '+' symbols. Absence of the gene indicated by '-'.

KO	Gene Product	CCAP 1479/9	CCAP 1479/10	CCAP 1479/13	CCY 0621	CCY 9618	PCC 7942	UTEX 2973	PCC 6301	FACHB-242	FACHB-1061
Photosynthesis											
PSII											
K02703	PsbA	++	+++	+++	+++	++	+++	+++	+++	+++	++
K02706	PsbD	-	-	-	-	-	++	++	++	-	-
K02705	PsbC	+	+	+	+	+	+	+	+	+	+
K02704	PsbB	+	+	+	+	+	+	+	+	+	+
K02707	PsbE	+	+	+	+	+	+	+	+	+	+
K02708	PsbF	+	+	+	+	+	+	+	+	+	+
K02713	PsbL	+	+	+	+	+	+	+	+	+	+
K02711	PsbJ	+	+	+	+	+	+	+	+	+	+
K02712	PsbK	+	+	+	+	+	+	+	+	+	+
K02714	PsbM	+	+	+	+	+	+	+	+	+	+
K02709	PsbH	+	+	+	+	+	+	+	+	+	+
K02710	PsbI	+	+	+	+	+	+	-	+	-	-
K02716	PsbO	+	+	+	+	+	+	+	+	+	+
K02717	PsbP	+	+	+	+	+	+	+	+	+	+
K08901	PsbQ	-	-	-	-	-	-	-	-	-	-
K03541	PsbR	-	-	-	-	-	-	-	-	-	-
K03542	PsbS	-	-	-	-	-	-	-	-	-	-
K02718	PsbT	+	+	+	+	+	+	-	+	+	+
K02719	PsbU	+	+	+	+	+	+	+	+	+	+
K02720	PsbV	+	+	+	+	+	+	+	+	+	+
K02721	PsbW	-	-	-	-	-	-	-	-	-	-
K02722	PsbX	+	+	+	+	+	+	+	+	+	+
K02723	PsbY	+	+	+	+	+	+	+	+	+	+
K02724	PsbZ	+	+	+	+	++	+	+	+	+	+
K08902	Psb27	+	+	+	+	+	+	+	+	+	+
K08903	Psb28	+	+	+	+	+	+	+	+	+	+
K08904	Psb28-2	-	-	-	-	-	+	+	+	+	+
PSI											
K02689	PsaA	+	+	+	+	+	+	+	+	+	+
K02690	PsaB	+	+	+	+	+	+	+	+	+	+
K02691	PsaC	+	+	+	+	+	+	+	+	+	+

Table 2.3 continued from previous page

KO	Gene Product	CCAP 1479/9	CCAP 1479/10	CCAP 1479/13	CCY 0621	CCY 9618	PCC 7942	UTEX 2973	PCC 6301	FACHB-242	FACHB-1061
K02692	PsaD	+	+	+	+	+	+	+	+	+	+
K02693	PsaE	+	+	+	+	+	+	+	+	+	+
K02694	PsaF	+	+	+	+	+	+	+	+	+	+
K08905	PsaG	-	-	-	-	-	-	-	-	-	-
K02695	PsaH	-	-	-	-	-	-	-	-	-	-
K02696	PsaI	++	++	++	++	++	+	+	+	+	+
K02697	PsaJ	+	+	+	+	+	+	+	+	+	+
K02698	PsaK	+	+	+	+	+	++	++	++	++	++
K02699	PsaL	+	+	+	+	+	+	+	+	+	+
K02700	PsaM	+	+	+	+	+	+	+	+	+	+
K02701	PsaN	-	-	-	-	-	-	-	-	-	-
K14332	PsaO	-	-	-	-	-	-	-	-	-	-
K02702	PsaX	-	-	-	-	-	-	-	-	-	-
Cytochrome b6/f complex											
K02635	PetB	+	+	+	+	+	+	+	+	+	+
K02637	PetD	+	+	+	+	+	+	+	+	+	+
K02634	PetA	+	+	+	+	+	+	+	+	+	+
K02636	PetC	+	+	+	++	+	+	+	+	+	+
K02642	PetL	-	-	-	-	-	+	+	+	+	+
K02643	PetM	+	+	+	+	+	+	+	+	+	+
K03689	PetN	+	+	+	+	+	+	-	+	+	+
K02640	PetG	+	+	+	+	+	+	+	-	+	+
Photosynthetic electron transport											
K02638	PetE	-	-	-	-	-	+	+	+	+	+
K02639	PetF	++++	++++	++++	++++	++++	+++	+++	+++	+++	+++
K02641	PetH	+	+	+	+	+	+	+	+	+	+
K08906	PetJ	+	+	+	++	++	+++	+++	+++	+++	+++
F-type ATPase											
K02112	beta	+	+	+	+	+	+	+	+	+	+
K02111	alpha	+	+	+	+	+	+	+	+	+	+
K02115	gamma	+	+	+	+	+	+	+	+	+	+
K02113	delta	+	+	+	+	+	+	+	+	+	+
K02114	epsilon	+	+	+	+	+	+	+	+	+	+
K02110	c	+	+	+	+	+	+	+	+	+	+
K02108	a	+	+	+	+	+	+	+	+	+	+
K02109	b	++	++	++	++	++	++	++	++	++	++
Antenna Proteins											
Allophycocyanin											
K02092	ApcA	+	+	+	+	+	+	+	+	+	+
K02093	ApcB	+	+	+	+	+	+	+	+	+	+
K02094	ApcC	+	+	+	+	+	+	+	+	+	+
K02095	ApcD	+	+	+	+	+	++	++	+	++	++

Table 2.3 continued from previous page

KO	Gene Product	CCAP 1479/9	CCAP 1479/10	CCAP 1479/13	CCY 0621	CCY 9618	PCC 7942	UTEX 2973	PCC 6301	FACHB-242	FACHB-1061
K02096	ApcE	+	+	+	+	+	+	+	+	+	+
K02097	ApcF	+	+	+	+	+	+	+	+	+	+
Phycocyanin											
K02284	CpcA	++	++	++	+	-	++	++	++	++	++
K02285	CpcB	+++	+++	+++	++	+	++	++	++	++	++
K02286	CpcC	-	-	-	-	-	++	++	++	++	++
K02287	CpcD	+	+	+	+	+	+	+	+	+	+
K02288	CpcE	+	+	+	+	+	+	+	+	+	+
K02289	CpcF	+	+	+	+	+	+	+	+	+	+
K02290	CpcG	++	++	++	++	++	+	+	+	+	+
Phycoerythrin											
K05376	CpeA	-	-	-	-	-	-	-	-	-	-
K05377	CpeB	-	-	-	-	-	-	-	-	-	-
K05378	CpeC	++	++	++	++	++	-	-	-	-	-
K05379	CpeD	-	-	-	-	-	-	-	-	-	-
K05380	CpeE	-	-	-	-	-	-	-	-	-	-
K05381	CpeR	-	-	-	-	-	-	-	-	-	-
K05382	CpeS	-	-	-	-	-	+	+	+	+	+
K05383	CpeT	-	-	-	-	-	-	-	-	-	-
K05384	CpeU	-	-	-	-	-	-	-	-	-	-
K05385	CpeY	-	-	-	-	-	-	-	-	-	-
K05386	CpeZ	-	-	-	-	-	-	-	-	-	-

There are more significant variations in the encoding of phycobiliprotein-rods which radiate out from the PBS core. There are differences in the copy number of phycocyanin (PC) subunits *cpcA* and *cpcB* with *Synechococcus* sp. CCY9618 encoding only *cpcB*. Other newly sequenced genomes encode both subunits with *cpcB* at an increased copy number compared to *Synechococcus elongatus* strains (Table 2.3). Interestingly, *cpcC* is absent from our sub-cluster 5.2 strains. This encodes the LR33 PC-associated linker polypeptide, responsible for stabilising rod substructures [368]. Meanwhile, the same strains encode an additional copy of *cpcG* (encoding a linker protein required for rod attachment to the AP core), with the two copies having distinct roles in PSII (*cpcG1*) and PSI (*cpcG2*) in *Synechocystis* sp. PCC 6803 [369]. The absence of *cpcG2* in *Synechococcus elongatus* strains suggests further differences in photosynthetic machinery between the two groups. Furthermore, while phycoerythrin (PE) is known to be absent in *Synechococcus elongatus* strains, it has been observed in other sequenced sub-cluster 5.2 freshwater picocyanobacteria [108]. However, the strains sequenced in this study are absent of *cpeAB* indicating PBS rods of PC only. Though lacking PE subunits, freshwater *Synechococcus* encode various

PE-associated proteins. Our sequenced *Synechococcus* encode two copies of *cpcC*, a PE-associated rod linker protein, while *Synechococcus elongatus* encode *cpeS*, an S-type lyase essential for mature PE generation [370, 371]. It is unclear if these genes are expressed, and the function they provide for *Synechococcus* lacking PE. Alternatively, this may suggest that the ancestor of *Syn/Pro Synechococcus* and *Synechococcus elongatus* had the capability to produce PE pigments, this trait lost in the subsequent evolutionary history of these strains in their freshwater environments, with remnants of PE-associated genes remaining.

2.4.3.3 *nirA* Operon

The most abundant N source in fresh water is NO_3^- [372], a nutrient which cyanobacteria can access via the *narB-nrtABCD-nirA* operon. This operon encodes the necessary proteins for NO_3^- assimilation, yet the gene neighbourhood of this operon differs between sub-cluster 5.2 freshwater picocyanobacteria and *Synechococcus elongatus.* This operon consists of a NO_3^-/NO_2^- bi-specific ATP-binding cassette-type transporter (*nrtABCD*), nitrate reductase (*narB*), and nitrite reductase (*nirA*). Among our newly sequenced strains (with *Synechococcus* sp. CCY9618 an exception), *nirA* and *nirB* are transcribed in the opposite direction to *nrtABCD* whereas *Synechococcus elongatus* encodes the six core genes contiguously (Figure 2.3). Furthermore, there are unrelated genes flanking *nrtABCD* - anthranilate phosphoribosyltransferase and a hypothetical gene. Contiguous operons are known for rapid gene expression for all proteins of a specific cellular process, however the unassociated genes and two-way transcription may suggest sub-cluster 5.2 freshwater picocyanobacteria respond slower to NO_3^- inducement, though bidirectional promoters may be involved.

Additional genes involved with NO_2^- assimilation are found in *Synechococcus elongatus* strains but absent from our sequenced strains. These include *nirB*, required for maximal nitrite reductase activity, and *ntcB*, a transcription factor involved in NO_2^--induced gene activation [218, 373]. Though NO_3^- is the most abundant traditional N source, it is also the most energetically costly, requiring eight electrons to reduce fully to ammonium (NH_4^+) ($NO_3^- > NO_2^- > NH_4^+$) [374]. Increasing the preference for NO_2^- over NO_3^- can reduce this demand which may result in substantial energy savings. *Synechococcus* sp. CCY9618 encodes a homologous transporter previously only identified in marine picocyanobacteria (*nrtP*) which preferentially takes up NO_3^- over NO_2^- [231]. The differences between sub-cluster 5.2

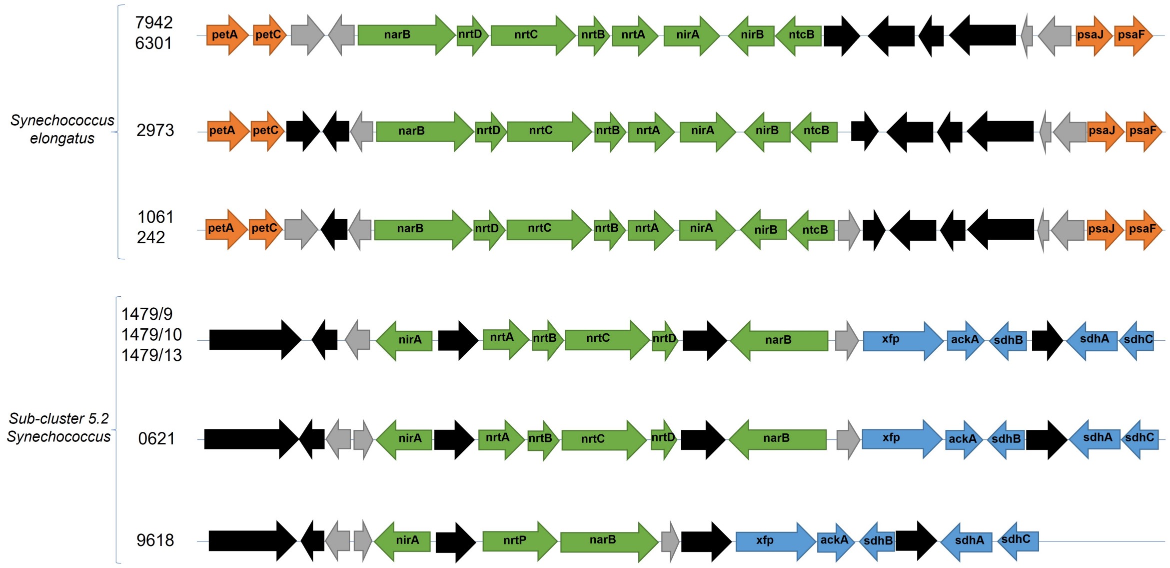

Figure 2.3: **Gene neighbourhood of the *narB-nrtABCD-nirA* operon for nitrate assimilation**. Green arrows are genes involved with NO_3^- assimilation. Orange arrows are genes involved with photosynthesis. Blue arrows are genes involved with C metabolism. Black arrows are other annotated genes while grey arrows indicate hypothetical genes. *petA*: apocytochrome f (K02634). *petC*: cytochrome b6f complex iron-sulphur subunit (K02636), *psaJ*: photosystem I subunit 9 (K02697), *psaF*: Photosystem I subunit 3 K02694), *xfp*: xylulose-5-phosphate/fructose-6-phosphate phosphoketolase (K01621), *ackA*: acetate kinase (K00925), *sdhB*: succinate dehydrogenase/fumarate reductase iron-sulphur subunit (K00240), *sdhA*: succinate dehydrogenase/fumarate reductase flavoprotein subunit (K00239), *sdhC*: succinate dehydrogenase/fumarate reductase cytochrome b subunit (K00241).

freshwater picocyanobacteria and *Synechococcus elongatus* may indicate differing preferences for nutrient growth, influencing the composition of the *Synechococcus* community.

2.5 Conclusions

The five newly sequenced freshwater picocyanobacteria expand the number of genomes available for sub-cluster 5.2 of the *Syn/Pro* clade. Comparative genomic analysis has revealed significant diversity within the genomic capabilities of freshwater picocyanobacteria (*Syn/Pro Synechococcus* spp.) v *Synechococcus elongatus*. The two groups encode different molecular machinery for both photosynthesis (PBS composition) and NO_3^-/NO_2^- assimilation, likely impacting their preferred ecological niches. Variations in encoded antennae proteins suggest that freshwater *Syn/Pro* picocyanobacteria may have reduced capability of low light absorption compared to *Synechococcus elongatus* while the presence of NO_2^--associated genes (*nirB* and *ntcB*) in *Synechococcus elongatus* may be an indication of preferential uptake of this nutrient. Phylogenomic analysis revealed that these newly sequenced strains are closely related in *Syn/Pro* sub-cluster 5.2. Though only two major pathways were investigated in detail in this study, the ecological variety within freshwater *Synechococcus* and within sub-cluster 5.2 of the *Syn/Pro* clade is demonstrated. A greater understanding of the environmental influence and molecular capabilities of keystone freshwater picocyanobacteria populations requires a focus on *Syn/Pro Synechococcus* communities and increased sampling of these strains to reveal additional genomic diversity.

2.6 Limitations

Next generation sequencing techniques have revolutionised genomic research by providing high-throughput sequencing capacity quicker and at reduced cost [375]. This has led to the increase in available sequenced genomes, though completeness has suffered as a result [376]. While closed genomes allow the set of contigs to be converted into a singular genome without gaps, draft genomes contain multiple contigs with unknown regions between contigs harbouring genetic content. The genomes sequenced in this study were not closed and contained gaps between the contigs. The extent of sequences missing can be estimated through assembly data (such as N50) and bioinformatic tools (such as BUSCO [341]), though these are imperfect. The N50 (the smallest contig with which half the genome is sequenced by contigs of size N50 and larger) value is linked to the contiguity of an assembly, though recent research has demonstrated poor correlation between this and genome completeness [377, 378]. BUSCO (Benchmarking Universal Single-Copy Orthologues) is currently the main method for assessing genome completeness [379]. This tool identifies orthologous single-copy genes that are lineage-specific, with the cyanobacteria database comprising 773 BUSCO genes. However, this too has its limitations, over-estimating completeness and reliant on a gene reference database which may not be appropriate [379]. Though there are techniques to infer the presence of missing genes (i.e. *psbD* in this study) such as molecular interrogation (with PsbD identified in the proteomics analysis of Chapter 4), investigating the gene position, and identifying neighbouring clusters, the generation of closed genomes is the most reliable method to accurately determine gene presence or absence.

This work provides a genomic context for *Syn/Pro* sub-cluster 5.2 freshwater picocyanobacteria, yet experimental validation is required to bridge the gap between genotype and phenotype. The presence of a gene alone is not sufficient for expression, as previously demonstrated for the N-fixing *nif* genes of *Vulcanococcus limneticus* LL [33]. The phenotype produced from the differences in molecular machinery in the photosynthesis pathways of *Synechococcus elongatus* and our newly sequenced strains must be characterised to better understand the evolutionary divergence between these two groups. Likewise with NO_2^- assimilation, *Synechococcus elongatus* encodes additional genes for NO_2^- metabolism and regulation though it is unknown the competitive advantage this provides over *Syn/Pro* sub-cluster 5.2 freshwater picocyanobacteria, especially due to the low concentrations of NO_2^- in fresh water [380]. Molecular research focusing on this freshwater group will provide much-needed insight into the diversity within freshwater *Synechococcus* and the ecological responses to multiple environmental conditions.

2.7 Future Directions

This study has sequenced five new freshwater picocyanobacteria from sub-cluster 5.2 of the *Syn/Pro* clade and conducted comparative genomic analyses to compare these genomes to those of *Synechococcus elongatus* strains. In the following chapters, further insights into the molecular machinery of these strains are uncovered with a focus on N assimilation. Utilising similar techniques to this study, the next chapter conducts a large-scale comparative genomic analysis on the N assimilatory capabilities of freshwater *Syn/Pro* clade picocyanobacteria compared to freshwater *Synechococcus elongatus*, marine picocyanobacteria, thermophylic picocyanobacteria, and other larger freshwater cyanobacteria (> 2 μm) (Chapter 3). Once the genomic capabilities of N assimilation are determined, experimental validation is carried out with growth assays and proteomic responses measured for several organic N sources (Chapter 4). This work will provide insights into the response of freshwater picocyanobacteria to a diverse range of N forms.

External to this body of work, further sequencing of sub-cluster 5.2 of the *Syn/Pro* clade is essential to grow the number of genomes available. The global reach of these freshwater picocyanobacteria compared to *Synechococcus elongatus* strains emphasises the requirement for focused research into this sub-cluster, both genomically and ecologically. In addition to sub-cluster 5.2, Cabello-Yeves *et al.* have sequenced genomes from freshwater picocyanobacteria belonging to sub-cluster 5.3, previously only represented by a handful of strains [103]. Further understanding of the evolutionary divergence between these sub-clusters of freshwater picocyanobacteria is essential in developing an holistic outlook towards microbial communities, especially for keystone groups such as the picocyanobacteria.

Chapter 3

Nitrogen Assimilation in Cyanobacteria Varies with Habitat and Cell Morphology

3.1 Contributions and Acknowledgements

This chapter was written with feedback from Stephen Maberly and Patricia Sánchez-Baracaldo, providing advice and contributing to editing. Giorgio Bianchini assisted with phylogenomic tree construction. All other work carried out is my own.

3.2 Introduction

Picocyanobacteria (0.5 – 2 µm) are globally abundant prokaryotes that contribute significantly to aquatic primary production [381, 382]. Marine *Prochlorococcus* and *Synechococcus* genera alone are responsible for 25% of ocean primary production with a total biomass of 1.6 x 10^{14} g C [299, 383]. This has led to widespread research into their ecology, evolution and genomic capabilities [10, 323, 384, 385]. In contrast, less is known about freshwater picocyanobacteria despite their ecological importance [3, 108, 386, 387]. Though picocyanobacteria frequently dominate the freshwater cyanobacterial community (primarily *Synechococcus* spp.), contributing up to 90% of total lake cyanobacteria biomass [4], few studies have utilised molecular ecology techniques [52, 122] and, more recently, genomic data to provide insights into their ecological adaptation [33, 103, 108].

Nitrogen (N) is a limiting or co-limiting nutrient in fresh waters [6, 388]. To overcome this, some cyanobacteria can transform inert atmospheric N_2 into bioavailable ammonium (NH_4^+), though this is an uncommon trait in freshwater picocyanobacteria (only encoded in a single freshwater picocyanobacterium to date (*Vulcanococcus limneticus* LL) [33]). Conversely, the high energetic costs of N fixation favour uptake and assimilation of dissolved forms of N, commonly inorganic forms such as NH_4^+, nitrate (NO_3^-), and nitrite (NO_2^-) [374]. However, the widespread abundance of dissolved organic N (DON) and its recently demonstrated bioavailability represents an under-researched N-pool, consisting of urea, amino acids (AAs), chitin, and glyphosate, among a multitude of humic compounds [154, 184, 252, 322, 389].

The molecular processes behind DON assimilation are not fully understood, with the exception of urea. Urea is primarily imported via an ATP-binding cassette (ABC)-type transporter and catabolised through urease activity, with these systems regulated by NtcA [390]. AA transporters (AATs) have been experimentally identified in model cyanobacteria (*Synechococcus elongatus* PCC 7942, *Synechocystis* sp. PCC 6803, and *Anabaena* sp. PCC 7120), with regulatory mechanisms linked to NtcA [213, 220, 254, 255, 259]. Though AATs are involved in heterocyst function, they also enable environmental AA uptake and provide a highly varied N source [254]. In addition to free AAs, combined AAs in the form of small oligopeptides are further potential sources of N for which transporters have been putatively identified in cyanobacteria [261, 262]. Other components of DON are slowly being uncovered as bioavailable. Among these nutrients with the greatest ecological impact is chitin, one of the most abundant natural compounds [391]. The molecular machinery necessary for chitin assimilation in cyanobacteria is not fully understood, but two catabolic pathways are indicated by the presence of chitinase (ChiA) or chitin deacetylase (ChdA) [389]. ChiA directly catalyses the hydrolysis of chitin into glucosamine monomers that can be assimilated [392], while ChdA catalyses the conversion of chitin to chitosan. This is then hydrolysed by chitosanase (ChoA) to generate glucosamine [393]. The herbicide glyphosate is another novel N (and phosphorus (P)) source which is increasingly found in fresh waters, with recent studies showcasing cyanobacterial growth on glyphosate (including *Synechococcus)* [394–396].

Eutrophication of freshwaters through anthropogenic inputs of inorganic and organic N can lead to community shifts and the formation of cyanobacterial blooms, including the production of cyanotoxins [397]. While most research is focused on *Microcystis* and other traditional bloom-forming species, picocyanobacteria are now being associated with bloom formation with strains of toxin-producing *Synechococcus* having been identified [16, 84]. *Synechococcus* spp. are abundant in both oligotrophic and eutrophic fresh waters. In oligotrophic systems especially, picoplankton can constitute the bulk of the phytoplankton biomass [3], partly as a result of their large surface area to volume ratio that enables them to out-compete larger phytoplankton ($>$ 2 µm) when resources are limiting [5, 19, 398]. *Synechococcus* regularly dominates the picoplankton fraction in aquatic environments, with this high abundance and rapid uptake resulting in picocyanobacteria acting as 'first responders' to natural or anthropogenic nutrient enrichment [35]. In eutrophic conditions, larger freshwater cyanobacteria (e.g. *Microcystis* spp. and *Dolichospermum* spp.) contribute the bulk of biomass to lake systems [43, 399]; however, the abundance of *Synechococcus* remains high and provides an important source of primary production and potential toxin generation [3].

Interest in picocyanobacteria has increased considerably in the last two decades with the use of molecular and genomic approaches underpinning ecological understanding. These studies have traditionally focused on marine environments, elucidating cellular mechanisms behind niche adaptation in the open ocean (such as a large number of subpopulations of *Prochlorococcus* as a response to highly mixed habitats [400] and differential adaptation to low phosphorus concentrations among marine *Synechococcus* eco-types [323]. This has led to a wealth of marine *Prochlorococcus* and *Synechococcus* genomes while knowledge of freshwater picocyanobacteria remains restricted. The scarcity of sequenced freshwater picocyanobacteria genomes has limited genetic approaches to uncover the ecological context of this keystone group. Most molecular research involving freshwater *Synechococcus* has instead utilised *Synechococcus elongatus* PCC 7942 [314, 401, 402], a strain belonging to a sister group of the *Syn/Pro* clade and lacking some of their derived traits (such as pico-size morphology) [108]. The *Syn/Pro* clade contains freshwater and marine picocyanobacteria comprising the genera *Synechococcus*, *Prochlorococcus*, *Cyanobium*, and *Vulcanococcus* [403]. Sparse taxonomic representation has previously limited our understanding of the *Syn/Pro* clade along with their genomic capabilities and evolutionary history. However, recent sequencing of freshwater picocyanobacteria genomes has increased their representation and provided greater accuracy for comparative genomic analyses, uncovering a more flexible genome in freshwater picocyanobacteria which may enable greater adaptability to their environment [34, 103, 108, 301].

In this study, we conducted comparative genomic analyses of cyanobacteria to identify encoded organic and inorganic N assimilation capabilities and mechanisms within a phylogenetic framework. We report the distribution of molecular machinery for N assimilation in freshwater picocyanobacteria, comparing these to picocyanobacteria of different environments (marine and thermal), the non-*Syn/Pro Synechococcus* strains of *Synechococcus elongatus*, and larger freshwater cyanobacteria. Our results have implications to understand how molecular machinery influences cyanobacterial communities across habitats and morphologies.

3.3 Methodology

3.3.1 Taxa Selection and Genome Datasets

Cyanobacterial genomes were obtained from the National Center for Biotechnology Information RefSeq database [404] and JGI IMG/ER [342] in September 2020. Taxa sampling focused on freshwater picocyanobacteria while broad sampling of marine picocyanobacteria, thermophylic picocyanobacteria, *Synechococcus elongatus* strains, and larger ($>$ 2 µm) freshwater cyanobacteria primarily from the Macrocyanobacteria (*senso* [98]) were also selected to provide a comparison between these groups and freshwater picocyanobacteria. This places the capabilities of freshwater picocyanobacteria in a wider context and enables us to identify different approaches to N assimilation between habitats and cell size morphologies.

3.3.1.1 Criteria Methodology

A representative sample of Cyanobacteria taxa were selected based on their habitat, cell size, and sequence availability, with their assembly quality assessed through genome completeness (Appendix Table B.6). This study focuses on the uptake and metabolic capabilities of cyanobacteria in pelagic and limnetic environments, further including benthic picocyanobacteria isolated from hot spring cyanobacterial mats for comprehensive picocyanobacteria coverage. Selected genomes were then categorised based on their isolation source into freshwater and marine habitats, in addition to the benthic hot spring picocyanobacteria. The isolation source was determined through manual checking of the cyanobacterial metadata (Genbank, JGI, scientific literature). In this study, marine habitats are defined as salinities greater than 18 ppt of dissolved salts. This incorporates open ocean environments (salinity of 35 ppt [405]) and coastal/poly-haline brackish environments (salinity of 18 - 32 ppt [406]) into one condition. The salinity of a freshwater habitat is defined as water with less than 0.5 ppt [407]. The genomes were also categorised by cell-size - picocyanobacteria (0.5 - 2 μm) and larger cyanobacteria ($>$ 2 μm). Where possible, cell size data was obtained by experimental observation (i.e. the designation of *Synechococcus elongatus* as non-picocyanobacteria [314]). However, the cell size for the majority of sequenced cyanobacteria is difficult to ascertain due to insufficient experimental analysis. In this case, size data was primarily inferred via taxonomic classification (phylogeny and nomenclature (i.e. *Synechococcus* spp.)). Finally, to facilitate comparisons between freshwater picocyanobacteria of the *Syn/Pro* clade (primarily *Synechococcus* spp.) and basal freshwater *Synechococcus* spp., these latter strains are distinctly categorised and broadly labelled as *Synechococcus elongatus* strains. This group includes the currently sequenced *Synechococcus elongatus* genomes, other basal *Synechococcus* spp. PCC 6312/PCC 7502, and *Thermosynechococcus elongatus* spp. Although this grouping is polyphyletic and contains distantly related strains [408], their historic classification as *Synechococcus* (and *Thermosynechococcus*), cellular size greater than 2 μm [409], yet clear divergence from the *Syn/Pro*, led to their classification as a single group in this analysis. In total, five groupings are compared for N assimilation capabilities:

- F-Pcy (freshwater picocyanobacteria)
- SE-Cy (*Synechococcus elongatus* strains)
- T-Pcy (thermophylic picocyanobacteria)
- M-Pcy (marine picocyanobacteria)
- LF-Cy (larger freshwater cyanobacteria)

Quality control was carried out on all genomes to reduce biased genome representation. The completeness of genomes was assessed using BUSCO [341], utilising a single-copy ortholog (SCO) dataset specific for cyanobacteria. Genomes where SCO presence was < 90 % were excluded. This resulted in a high-quality cyanobacterial dataset of 166 genomes to determine metabolic capabilities reliably across a wide phylogenetic range. This cyanobacterial dataset was split into the five groups previously mentioned consisting of: 17 F-Pcy genomes; nine *Synechococcus elongatus* genomes; eight T-Pcy genomes; 81 M-Pcy genomes; and 51 LF-Cy genomes (of which 16 are from unicellular species, 35 from filamentous species). The recent increase in Cyanobacterial sequencing has resulted in on-going changes to their taxonomic classification [104]. *Anabaena* spp., *Dolichospermum* spp., and *Aphanizomenon* spp. have historically been identified as separate species, but are now grouped into the ADA clade [410]. Furthermore, the taxonomy of *Anabaena* spp. and *Dolichospermum* spp. is undergoing a transition, with previously identified planktonic *Anabaena* spp. being reclassified as *Dolichospermum* [411]. However, in this study the *Anabaena* classification for isolates as used by original authors is retained.

3.3.2 Nitrogen Assimilation Gene Identification

An in-depth search through the scientific literature and maps of metabolic pathways identified 171 genes involved in cyanobacteria N assimilation. To examine literature related to N assimilation, a comprehensive database search of Google Scholar and PubMed was manually carried out with the search terms (”Cyanobacteria” or ”Cyanobacterial) and (”Nitrogen” or ”Dissolved organic nitrogen” or ”nitrate” or ”nitrate” or ”ammonium” or ”ammonia” or ”urea” or ”peptide” or ”amino acid” or ”chitin” or ”glyphosate”) and (”assimilation” or ”uptake” or ”metabolism”). In addition, KEGG pathway mapping [344] was utilised to identify enzymes involved in cyanobacterial AA biosynthesis. These searches identified experimentally characterised proteins involved in the transport of inorganic (NH_4^+, NO_3^-, NO_2^-) and organic (urea, AAs, peptides, chitin, glyphosate) forms of N, in addition to proteins involved in the regulation of N assimilation, N fixation, metabolic enzymes for incorporation of N, and AA biosynthesis enzymes. These target genes were used for comparative genomics analyses with query sequences from multiple species detailed in Appendix Table B.8.

3.3.3 Comparative Genomic Analyses

Target genes in our genome dataset were identified using BLASTP [334], with query sequences (when found) from *Synechococcus* sp. 1G10, *Synechococcus* sp. JA-3-3Ab, *Synechococcus* sp. WH8102, *Synechococcus* sp. MIT9509, *Prochlorococcus marinus* MIT9303, and *Synechococcus elongatus* PCC 7942 (Appendix Table B.8). An E-value threshold of 1×10^{-50} (1×10^{-10} for genes of $<$ 100 AAs) was used to determine the presence of these genes. This E-value was chosen to ensure homology between identified sequences and query sequences (in addition to the generation of gene trees). When identified sequences were matched with more than one target gene, the identity of the sequence was determined by largest bit-score. Identified genes for each target were compiled and then aligned with MAFFT [346] using local pair alignment. For each gene, phylogenetic trees were estimated in IQ-TREE [412] using the LG + F + R7 model, which was determined to be the best-fitting model based on BIC scores. Phylogenies were implemented as an additional check for homology to the target gene, ensuring identified sequences were closely related to the reference sequences. This additional check is necessary to prevent false positive identifications, especially among the many transporters analysed in this study which regularly have high sequence identity to other transporters responsible for different substrates (i.e. NrtABCD for NO_3^- and CmpABCD for bicarbonate). Hypergeometric statistical testing and FDR adjustment was carried out in R [413] to identify statistically significant enrichment and depletion of these genes in the genomes of F-Pcy, SE-Cy, T-Pcy, M-Pcy, and LF-Cy.

3.3.4 Phylogenomic Analysis

Evolutionary relationships of the taxa utilised in this study were estimated using phylogenomic analysis. Our genome dataset consisted of 166 cyanobacteria genomes (further utilised for comparative genomics), Melainabacteria as an outgroup (composed of Gastranaerophilales and Vampirovibrionales), and a further 205 cyanobacteria genomes of varying morphologies and lifestyles to improve the resolution of the phylogenomic analysis. In total, our dataset for phylogenomic analysis included 373 cyanobacteria genomes in addition to ortholog sequences from 143 protein-coding genes (based on previously published studies [94, 97, 98]). To identify these ortholog sequences among our genome dataset, we performed BLAST searches (query sequences found in Appendix Table B.8) using BLASTP v2.11.0+ [345] utilising an E-value cutoff of 10^{-5}. The hit with the highest score was retained and the corresponding protein sequences extracted. These sequences were then aligned using MAFFT v7.511 (-localpair, -maxiterate 1000 parameters) [346] and maximum-likelihood trees constructed with IQ-tree v2.2.0 [347] (*LG* protein evolution model, -fast). These individual gene trees were used to identify 'true' ortholog sequences, determining these based on cluster association with the BLAST query sequences, with these true orthologs re-aligned with MAFFT. Re-alignments were inspected manually and mis-aligned columns were removed, as were alignment positions with a gap content higher than 80%. The model with the lowest BIC score was chosen as the optimal evolutionary model for each gene using IQ-TREE's -m MF option [348]. Subsequently, a maximum-likelihood partitioned phylogenomic analysis was carried out with IQ-TREE, partitioned using the -p and -B 1000 [350] parameters and the previously established evolutionary models with each gene assigned to its own partition (as in Chapter 2). This analysis was repeated twice and the resulting trees compared to confirm there were no appreciable differences between them. Finally, the resultant tree was pruned (to improve visualisation) to include only the 166 cyanobacteria strains utilised in the comparative genomics analyses (plus the Melainabacteria outgroup), with the full tree found in Appendix Figure A.7.

3.4 Results

3.4.1 Phylogenomic Analyses

Phylogenomic analyses find two major clades of Cyanobacteria known as the Macro- and Micro-cyanobacteria, consistent with previous studies [94, 98, 104, 119]. However, in contrast to these previously published studies, this analysis features a different position for the LPP clade (i.e. Nodosilinea, Halomicronema, etc.). This may be due to multiple reasons. Firstly, the over-representation of *Syn/Pro* in this analysis may yield artefacts in the phylogenomic tree, especially in the surrounding taxa of the *Syn/Pro* (i.e. the LPP). Secondly, though the LPP are typically found as a sister group of the *Syn/Pro*, the bootstrap support for this group can be low (as seen in Strunecky et al. [104]), suggesting there remains some uncertainty over the 'true' position of the LPP.

The majority of picocyanobacteria are Microcyanobacteria, belonging to a monophyletic clade of Synechococcales containing *Synechococcus*, *Prochlorococcus*, *Cyanobium*, and *Vulcanococcus* (*Syn/Pro* clade comprising sub-clusters 5.1, 5.2, and 5.3, and *Prochlorococcus* spp.) (Figure 3.1), consistent with previous studies [108, 414]. Freshwater and marine *Synechococcus* are distributed in previously recognised sub-clusters [415]. F-Pcy are found in sub-clusters 5.2 and 5.3, while sub-cluster 5.1 comprises marine *Synechococcus. Prochlorococcus* spp. form a sister sub-cluster to marine *Synechococcus* (Figure 3.1). A monophyletic clade of thermophylic *Synechococcus* (recently renamed *Thermostichales* [104]) isolated from American hot springs [133, 416] are termed the Hot Spring clade (T-Pcy) and are the most divergent picocyanobacteria to the *Syn/Pro* (Figure 3.1). Marine *Synechococcus* spp. are found primarily in sub-cluster 5.1 of the *Syn/Pro* clade, though are also represented in sub-cluster 5.2 and 5.3. M-Pcy are also present in the Macrocyanobacteria with estuarine-dwelling β-cyanobacteria comprising *Synechococcus* spp. forming a monophyletic sister clade to the Chroococales *Geminocystis.* Though in the Macrocyanobacteria, environmental pressures may have played a role in driving a reduction of cell diameter in this group, independently yielding pico-size morphology through convergent evolution. *Synechococcus elongatus* strains appear as a sister group to the *Syn/Pro* clade while *Synechococcus* sp. PCC 6312, PCC 7505, *Thermosynechococcus* spp., alongside *Synechococcus elongatus*, make up the polyphyletic SE-Cy, forming a basal lineage of freshwater *Synechococcus.* LF-Cy are grouped according to their taxonomic classification, with five orders analysed: Oscillatoriales (3 taxa), Nostocales (3 taxa), Chroococales (3 taxa), Pleurocapsales (1 taxon), and Synechococcales (3 taxa).

3.4.2 Cyanobacteria Utilise Inorganic Forms of Nitrogen Through Different Machinery

NH_4^+ is a vital source of N in freshwater and marine environments. Consequently, molecular machinery for its assimilation is abundant across all habitats and present in all genomes studied. Amt1 is the primary NH_4^+ transporter for cyanobacteria and is encoded by almost all genomes studied, though notably absent among the N_2-fixing freshwater ADA clade (*Anabaena* spp., *Aphanizomenon* spp., and *Dolichospermum* spp.) (Figure 3.2). In contrast to the abundant nature of Amt1, the encoded accessory NH_4^+ transporters vary between habitat and cell size. F-Pcy encode *amtB*, producing an accessory transporter which promotes NH_4^+ uptake at lower concentrations [221]. This is the only NH_4^+ accessory transporter encoded by F-Pcy and more frequent in these strains than in both M-Pcy ($q < .001$) and T-Pcy ($q = .001$), fully absent in the latter environment. Among M-Pcy, another NH_4^+ accessory protein, Amt3, is present in the β-*Synechococcus* though the larger clade of sub-cluster 5.1

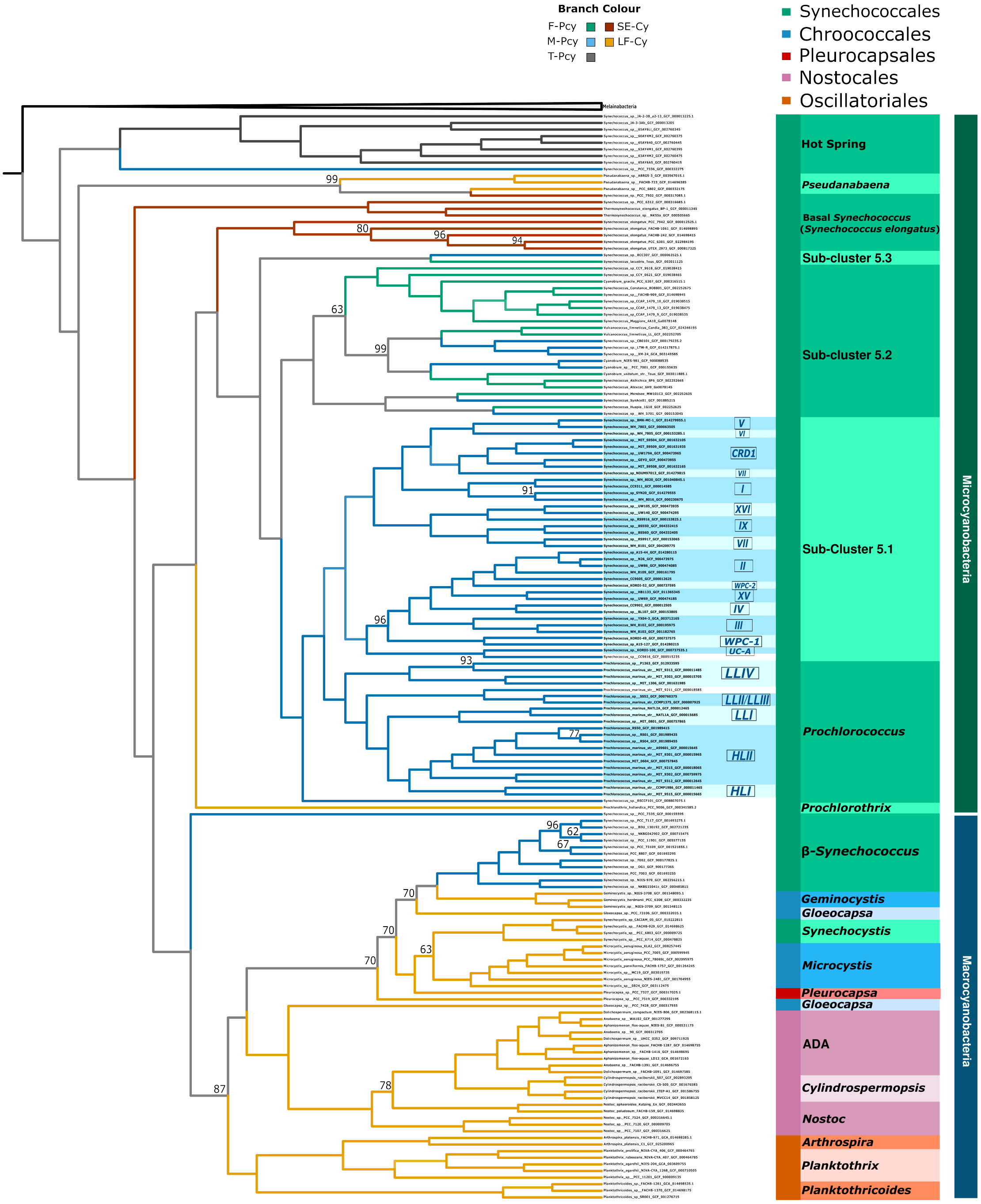

Figure 3.1: **Phylogeny of picocyanobacteria**. This maximal likelihood tree (IQ-tree v2.2.0) is based on 143 orthologous proteins, rooted using Melainabacteria as an outgroup. Node support values were generated with ultrafast bootstrap approximations with bootstrap values less than 100 displayed above branching nodes while blank nodes have support of 100. Taxonomic classification is coloured by order (see inset legend) and sub-groupings. The ADA clade includes *Anabaena*, *Dolichospermum*, and *Aphanizomenon* spp.. Note that some sub-groupings are polyphyletic. Sub-clades are given in boxes adjacent to marine *Synechococcus* sub-cluster 5.1 and *Prochlorococcus* spp..

Synechococcus encode no accessory proteins. AmtB is also the sole NH_4^+ accessory protein found in SE-Cy. Among LF-Cy, eight of the genera studied encode *amtB*, while additional accessory proteins (*amt2*, *amt3*, and *amt4*) are found in seven genera (Figure 3.2).

The mechanism for NO_3^-/NO_2^- uptake differs significantly between habitats. F-Pcy overwhelmingly utilise the NO_3^-/NO_2^- ABC-type NRT transporter system, unlike M-Pcy (Figure 3.2; $q < .001$). This transport system is also the primary method of NO_3^-/NO_2^- uptake in T-Pcy, SE-Cy, and LF-Cy (Figure 3.2). Conversely, M-Pcy utilise two unique transporters for NO_3^- (NrtP) and NO_2^- (FocA) uptake [217]. These are more frequent in M-Pcy than in F-Pcy ($q < .001$) and represent alternative mechanisms for NO_3^-/NO_2^- assimilation. Interestingly, *Synechococcus* sp. CCY9618 is the only F-Pcy that has lost the NRT transporter and encodes the marine *nrtP* (Appendix Table B.9). Additional NO_3^-/NO_2^- assimilatory mechanisms are found in SE-Cy and the M-Pcy *β-Synechococcus*, encoding NrtS and the cyanate transporter CynABD (which can take up NO_2^-) which are not found in F-Pcy.

The ability to fix atmospheric N is relatively common in LF-Cy but rare in picocyanobacteria (Figure 3.2). Of the 15 LF-Cy genera studied, seven (47%) encode components of nitrogenase. On the other hand, T-Pcy are the only picocyanobacteria to predominately have the capability to fix N_2. N-fixation is very rare in F-Pcy, the *nifHDK* operon only found in two strains of sub-cluster 5.2 (*Vulcanococcus limneticus* LL (isolated from Lake Albano, Italy; previously identified [33]) and *Vulcanococcus limneticus* 3B3 (isolated from Lake Candia, Italy - 550 km northeast of Lake Albano).

3.4.3 Amino Acid Transporter Distribution Differs Significantly Among Cyanobacteria

Our analyses reveal statistically significant differences in AAT distribution among cyanobacteria. LF-Cy contain the largest number of different AATs averaging 3.7 AATs per genome, significantly greater than F-Pcy (2.8 AATs), T-Pcy (1.5 AATs), and M-Pcy (1.8 AATs) ($p = .03$, $p < .001$, $p < .001$, respectively; ANOVA + Tukey HSD), owing to their wider genetic diversity and larger genome size (Appendix Figure A.9). SE-Cy encode 3.6 AATs per genome, greater than F-Pcy though not significantly different ($p = .1$; ANOVA + Tukey HSD). F-Pcy contain the most AATs among picocyanobacteria, significantly more than M-Pcy and T-Pcy ($p = .006$; $p < .001$, ANOVA + Tukey HSD).

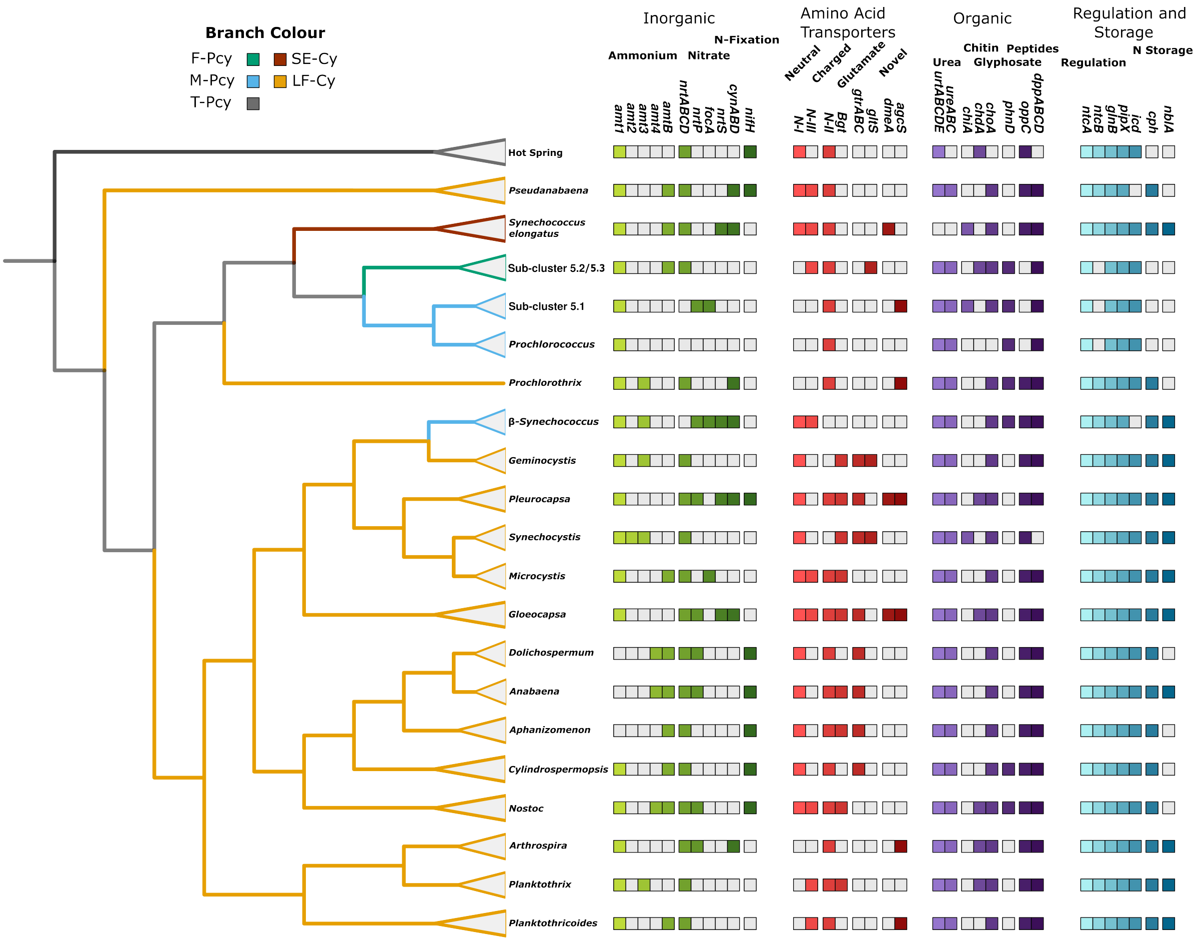

Figure 3.2: **Prevalence of genes encoding transporters and metabolic enzymes involved in assimilation of cyanobacterial nitrogen sources.** A gene is considered present if encoded in ≥ 40% of genomes of that group. Tree groupings and branches are based off the maximum likelihood tree generated in Figure 3.1. The presence of inorganic (green), amino acid (red), other organic (purple), and regulation and storage (blue) genes are shown. *amt1* encodes the primary NH_4^+ transporter; *amt2*, *amt3*, *amt4* and *amtB* encode NH_4^+ uptake accessory proteins. *nrtABCD* encodes the NRT NO_3^-/NO_2^- ABC transport system; *nrtP* encodes a NO_3^- permease; *focA* encodes a NO_2^- transporter; *nrtS* encodes a NO_3^-/NO_2^- transporter; *cynABD* encodes an ABC-type cyanate transporter that can take up NO_2^-; *nifH* encodes the iron protein of nitrogenase; *urtABCDE* encodes a urea ABC-transporter, *ureABC* encodes the major urease complex to catabolise urea; *oppC* encodes an oligopeptide transporter; *dppABCD* encodes a dipeptide transporter; *chiA* encodes chitinase to catabolise chitin; *chdA* encodes chitin deacetylase to convert chitin to chitosan; *choA* encodes chitosanase to catabolise chitosan; *phnD* encodes a phosphonate transporter which uptakes glyphosate; *ntcA* and *ntcB* encode N assimilation transcriptional regulators; *glnB* encodes the PII signal transduction protein; *pipX* encodes the co-activating factor PipX; *icd* encodes isocitrate dehydrogenase which catalyses the production of 2-oxoglutarate; *cph* encodes cyanophycinase which breaks down the N store cyanophycin; *nblA* encodes the phycobilisome degradation protein.

The distribution of AATs differ significantly among habitats within picocyanobacteria (Figure 3.3B; Appendix Table B.10). Among the ABC-type transporters, N-II and N-III are ubiquitously found in F-Pcy; both AATs are found significantly more frequently in F-Pcy than both T-Pcy and M-Pcy (Table 3.1). In contrast, N-I is more prevalent in T-Pcy compared to F-Pcy while N-II is also encoded by 50% of T-Pcy genomes. In M-Pcy, only N-II has high prevalence (73%). Bgt, the only transporter capable of basic AA uptake, is not found in picocyanobacteria. The presence of glutamate-specific transporters also varies. A full complement of Gtr components (encoded by *gtrABC*) is found in only two picocyanobacterial genomes (M-Pcy *Syn/Pro Synechococcus* sp. PCC 7335 and *Synechococcus* sp. PCC 7336 (both basal cyanobacteria, more closely related to thermophylic picocyanobacteria) while genes for the two transmembrane proteins (*gtrA* and *gtrB*) are encoded in 28% of M-Pcy, significantly greater than in F-Pcy (q = .002). Meanwhile, the ATPase component of Gtr (encoded by *gtrC*) is fully prevalent in T-Pcy (100%) though these genomes lack the transmembrane proteins. An additional glutamate transporter, the Na^+-dependent AAT GltS, is more frequent in F-Pcy (70%) than both T-Pcy (0%) and M-Pcy (9%), indicating a potentially greater propensity for glutamate-uptake in F-Pcy. The putative AgcS is the sole fully-encoded AAT which is depleted in F-Pcy (6%) compared to M-Pcy (65%) ($q < .001$).

The variety of AATs found in LF-Cy reflects their greater diversity. Of the 15 taxa analysed, five unique AAT combinations were identified (Figure 3.2). Though all known AATs are found in LF-Cy, the specific AAT distributions differ between LF-Cy and F-Pcy. The N-I AAT is found more frequently in LF-Cy than F-Pcy with 11 genera encoding it though only five LF-Cy genera encode the F-Pcy-dominant N-III. Furthermore, *bgtB* is almost exclusively found in LF-Cy, encoded by nine genera, and all components of the Gtr transporter are encoded by eight genera. *Gloeocapsa* spp. encode the most AATs with seven AATs identified, lacking only GltS.

Freshwater *Synechococcus* also vary in AAT composition (Figure 3.2). The N-II and N-III AATs are prevalent in both the F-Pcy and *Synechococcus elongatus* lineages. However, *Synechococcus elongatus* strains have additional AAT capability, encoding the N-I transporter (100%) which is absent from the *Syn/Pro.* In contrast, GltS found in F-Pcy is significantly absent from the *Synechococcus elongatus* lineage (11%; q = .028). However, like the similarly basal lineage of the T-Pcy, *Synechococcus elongatus* strains encode the ATPase of the Gtr system (*gtrC*) which is absent from F-Pcy. Further differences include the presence of DmeA, an AAT thought to interact with N-II [258]. This is rarely found among cyanobacteria, and totally absent from picocyanobacteria, yet is present in over half of *Synechococcus elongatus* strains (58%). Among LF-Cy, DmeA is only identified in two genera (*Gloeocapsa* and *Pleurocapsa*).

The three clades of M-Pcy display distinct AAT distributions. The β-*Synechococcus* encode N-I (100%) and N-III (100%), a unique set of two neutral AATs, and are the only clade to not encode a charged AAT. Conversely, the only ABC-type AAT encoded by M-Pcy *Syn/Pro Synechococcus* is the charged AA importer N-II (96%), also encoding AgcS (94%). The genomic streamlining of *Prochlorococcus* spp. has resulted in a lack of AAT diversity. Over half of the *Prochlorococcus* spp. studied encode the N-II transporter (57%) while the remaining 43% of genomes encode no known AAT, suggesting no active AA uptake and an inability to utilise AAs as a N source. Of the 57% of *Prochlorococcus* spp. identified with N-II sequences, 58% of these additionally encode AgcS. Both high light and low light clades encode N-II, though those strains encoding N-II + AgcS are exclusively found in low light clades.

Table 3.1: **Statistical significance of differential enrichment of amino acid transporters between F-Pcy and SE-Cy, M-Pcy, T-Pcy, and LF-Cy.** Transporters are divided into their protein-encoding components. *bgtA* is involved in both the N-II and Bgt transport system. Significant differences based on the hypergeometric test and FDR adjustment: NS = not significant. * indicates $q < .05$; ** indicates $q < .01$; *** indicates $q < .001$. Colour shading indicates which cyanobacterial size and habitat that component is enriched in: Green: F-Pcy; Orange: SE-Cy; Grey: T-Pcy; Blue: M-Pcy; Purple: LF-Cy.

Transport System	Transporter Component	F-Pcy vs SE-Cy	F-Pcy vs T-Pcy	F-Pcy vs M-Pcy	F-Pcy vs LF-Cy
N-I	*natA*	***	**	NS	***
	natB	***	***	NS	***
	natC	***	***	NS	***
	natD	***	***	NS	***
	natE	***	***	NS	***
N-III	*natI*	NS	***	***	***
	natJ	NS	***	***	***
	natK	NS	***	***	***
	natL	NS	***	***	***
	natM	NS	***	***	***
N-II	*natF*	NS	*	*	NS
	natG	NS	*	*	NS
	natH	NS	*	*	NS
	bgtA	NS	*	*	NS
Bgt	*bgtB*	NS	NS	NS	***
Gtr	*gtrA*	NS	NS	*	***
	gtrB	NS	NS	*	***
	gtrC	**	***	NS	***
GltS	*gltS*	*	**	***	**
DmeA	*dmeA*	*	NS	NS	NS
AgcS	*agcS*	NS	NS	***	NS

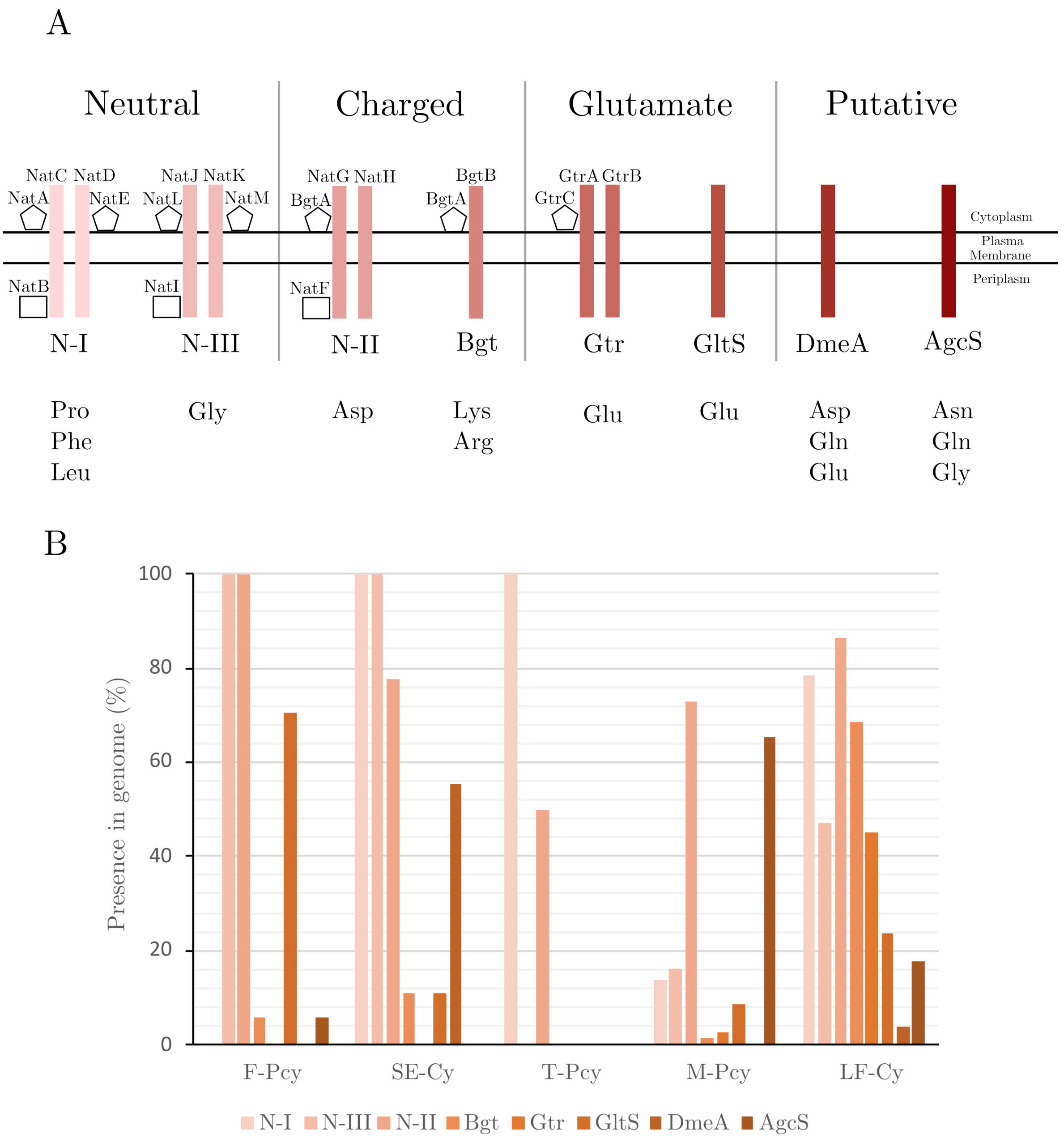

Figure 3.3: **Amino acid transporters and their prevalence in cyanobacteria.** A) AATs identified in cyanobacteria and their favoured AA substrates. N-I, N-II, N-III, and Bgt are ABC-type transporters, responsible for the majority of AA transport. Gtr and GltS are glutamate-specific transporters. DmeA and AgcS are putative AATs which are not fully characterised in cyanobacteria. Pentagons: ATPases; White rectangles: Periplasmic substrate-binding proteins; Red rectangles: Transmembrane proteins. AAT characteristics are displayed above and favoured substrates are displayed below each transporter (for minor substrates see Appendix Table B.16). B) Prevalence of AATs in cyanobacteria genomes based on habitat and cell size. F-Pcy: freshwater picocyanobacteria; SE-Cy: *Synechococcus elongatus* strains; T-Pcy: thermophylic picocyanobacteria; M-Pcy: marine picocyanobacteria; LF-Cy: larger freshwater cyanobacteria.

3.4.4 Picocyanobacteria Encode Pathways to Assimilate Known and Novel Organic Forms of Nitrogen

The ability to take up and metabolise urea is prevalent among picocyanobacteria and LF-Cy, though absent from SE-Cy (Figure 3.2). F-Pcy encode the urea ABC transporter (*urtABCDE*) and all core (*ureABC*) and accessory components (*ureDEFG*) of the urea catabolising enzyme urease. In contrast, the majority of SE-Cy (56%) lack any capability to take up and metabolise urea based on the absence of the *urt* and *ure* operons. In thermal environments, T-Pcy may have urea uptake capability, encoding *urtABCD*, yet the absence of *urtE* (ATPase component) suggests an incomplete urea transporter. Furthermore, only *ureB* of urease is encoded by all T-Pcy (Appendix Table B.11). It is unclear if these strains can import urea, and subsequently metabolise it with an incomplete urease. Among the M-Pcy, the *urt* and *ure* operon are widely encoded, though a minority of *Prochlorococcus* spp. (19%) lack both operons (Appendix Table B.11). Urea bioavailability is also prevalent in LF-Cy with all but four strains fully encoding the *urt* and *ure* operons. Interestingly, the strains deficient of these sequences (*Anabaena* sp. 90, *Dolichospermum* sp. UHCC 0352, *Nostoc* sp. PCC 7107, and *Nostoc* sp. PCC 7524) are capable of N_2 fixation.

Novel sources of N, such as chitin and glyphosate, can be taken up by cyanobacteria to varying degrees. The utilisation of chitin initially requires catabolism of this polymer, with hydrolytic mechanisms varying according to habitat and morphology. In picocyanobacteria, the ability to break down chitin directly through chitinase (encoded by *chiA*) is only found in marine environments. On the other hand, the catabolism of chitin via deacetylation (encoded by *chdA*) and subsequent hydrolysis of chitosan is found in F-Pcy (Figure 3.2). Interestingly, while F-Pcy encode this deacetylation pathway, SE-Cy encode *chiA*, indicating the divergence between these two groups. T-Pcy encode the initial step of the deacetylation pathway (*chdA*), however these strains lack chitosanase, suggesting the absence of chitinolytic activity. Among M-Pcy, *Synechococcus* of the *Syn/Pro* clade possess *chiA* (47%) while all three M-Pcy clades lack *cdhA*. LF-Cy encode both mechanisms for chitin breakdown; while *Synechocystis* spp. are the only LF-Cy to encode *chiA*, chitosan metabolism capabilities are found in five genera (*Pleurocapsa*, *Gloeocapsa*, *Nostoc*, *Arthrospira*, and *Planktothrix*) though significantly less frequently than in *Syn/Pro* F-Pcy (q = .03). Similar variability is found for glyphosate uptake through the phosphonate transporter (encoded by *phnD*). This transporter is significantly more prevalent in F-Pcy compared to SE-Cy, T-Pcy, and LF-Cy (all $q < .001$) though *phnD* is also abundant in M-Pcy.

Small dipeptides and oligopeptides are further potential sources of N present in the DON pool. A dipeptide transporter (encoded by *dppABCD*) is abundant in F-Pcy, M-Pcy, SE-Cy, and LF-Cy yet absent from T-Pcy (Figure 3.2). Meanwhile, the oligopeptide transporter OppC is ubiquitous in LF-Cy and SE-Cy (100%) with a mixed distribution in picocyanobacteria. Oligopeptide transport is absent from F-Pcy and M-Pcy of the *Syn/Pro* clade, yet is encoded in T-Pcy and the β-*Synechococcus* clade of M-Pcy.

3.4.5 Nitrogen Regulation and Nitrogen Metabolism

The regulatory mechanism behind the expression of N assimilatory genes is generally conserved among cyanobacteria. The primary cellular N signal is provided by 2-oxoglutarate, an intermediate in the TCA cycle generated through the oxidative decarboxylation of isocitrate and catalysed by isocitrate dehydrogenase (encoded by *icd*). This enzyme is found abundantly across habitats and morphologies, though is notably absent from the β-*Synechococcus* M-Pcy (Figure 2.2). The essential nature of this enzyme suggests an alternative pathway to 2-oxoglutarate formation for the β-*Synechococcus*. The 2-oxoglutarate sensors P-II and PipX are ubiquitously present as is the global N transcription factor NtcA (Figure 3.2). A further transcription factor, NtcB, is found in all environments, yet not found among picocyanobacteria of the *Syn/Pro* clade. This regulator, associated with NO_2^--induced activation, is present among all LF-Cy, SE-Cy, T-Pcy, and the β-*Synechococcus* M-Pcy.

N storage molecules can enable cells to withstand nutrient starvation yet their utilisation is more prevalent in SE-Cy and LF-Cy than picocyanobacteria. The genes required to encode components that synthesise (*cphA*) and breakdown (*cphB*) cyanophycin, a polymer of aspartate and arginine [191], are abundant in LF-Cy (98%), while present but less widespread in SE-Cy (44%). In contrast, this capability is absent from F-Pcy, though *Synechococcus* sp. 1G10 can synthesise cyanophycin but apparently lacks the ability to degrade it. Similarly, T-Pcy and M-Pcy generally do not encode *cphAB*, with the Macrocyanobacteria lineage of β-*Synechococcus* the only picocyanobacteria to fully encode *cphAB* (100%). Phycobilisomes can also be degraded during N limitation, though this study finds this capability is less common. The PSII light-harvesting antennae are broken down by NblA, which is mostly present in LF-Cy (55%), encoded by nine genera, and SE-Cy (67%) (Figure 3.2). However, NblA is absent from the majority of picocyanobacteria, not present in F-Pcy, T-Pcy, and most M-Pcy, again only found in the β-*Synechococcus* clade, suggesting mechanisms of intracellular N scavenging are not present in picocyanobacteria.

Glutamine synthetase is an essential enzyme in N metabolism that assimilates N (in the form of NH_4^+) into glutamate for subsequent cellular incorporation. Two forms of glutamine synthetase are found in cyanobacteria, encoded by *glnA* (GSI) and *glnN* (GSIII). While *glnA* is highly transcribed in both sufficient- and replete-N conditions, *glnN* is expressed only under N-starvation [226]. Among F-Pcy, both *glnA* and *glnN* are present with *glnN* found significantly more frequently in F-Pcy than M-Pcy, T-Pcy, and LF-Cy (all $q < .001$; Appendix Table B.12). *glnA* is almost universally present, missing only from the LF-Cy *Pseudanabaena* and *Prochlorothrix* (which encode *glnN*) while *glnN* is absent from T-Pcy and the *Syn/Pro* clade of M-Pcy. Among LF-Cy, most genera lack *glnN* with eight of the 15 studied lacking GSIII. Thus the added N assimilation capability gained through GSIII is most relevant for F-Pcy and SE-Cy.

3.5 Discussion

Enhanced genetic and molecular techniques have shed light on the diversity within Cyanobacteria and highlighted the importance of *Synechococcus* as a foundational species in freshwater ecosystems [37]. Oligotrophic freshwater habitats are at risk of rapid nutrient pulses [417, 418], leading to short-lived blooms of picocyanobacteria that can impact the overall health of the ecosystem [84, 419]. Nutrients present in an environment, alongside other physical, chemical, and biological factors, dictate the microbial community, while the transporters and metabolic enzymes encoded provide the blueprint for nutrient bioavailability. In this study, we investigated the assimilation capabilities of the essential nutrient N in F-Pcy, comparing them to SE-Cy, M-Pcy, T-Pcy, and LF-Cy, while placing this in an evolutionary context.

3.5.1 Differential Nitrogen Assimilation Capabilities and Their Ecological Implications

3.5.1.1 Amino Acid Composition in DON Pool May Influence the Cyanobacterial Community

Eight AATs have been identified in cyanobacteria, each exhibiting preferences for specific substrates (neutral/acidic/basic) and varying uptake rates for individual AAs (Figure 3.3A). The N-II transporter is the most prevalent across all genera. This ABC-type transporter takes up acidic and neutral polar AAs, preferentially importing aspartate, and to a lesser extent, glutamate [254]. The widespread presence of acidic and neutral polar AA uptake mechanisms may promote cyanobacteria proliferation in oligotrophic environments considering the abundance of glutamate and aspartate among dissolved free AAs (DFAAs) in DON [173, 277, 278, 280, 290], though dependent on substrate concentrations and N-II kinetic properties. F-Pcy also encode the N-III transporter, responsible for the uptake of hydrophobic AAs, preferentially glycine [255]. This AAT is absent from the majority of M-Pcy, T-Pcy, and LF-Cy, although the presence of the N-I neutral transporter (proline, phenylalanine, leucine preferred) in SE-Cy, T-Pcy, and LF-Cy represents alternative neutral uptake systems in these groups [253]. Neutral AAs are the most abundant AAs in both freshwater and marine environments, and in all strata of the water column, from the pelagic and limnetic zones to the sedimentary layer [280, 290, 420–423]. This prevalence of neutral AAs suggests a competitive advantage in their enhanced transport (as both a N and C source), with glycine regularly at the highest concentration among neutral AAs [171, 290, 423]. With an acidic and a neutral AAT, the bioavailable pool for F-Pcy is large; however, the absence of a basic AAT represents a limitation. In addition to ABC-type transporters, F-Pcy also encode the glutamate-specific transporter GltS. Glutamate is as an important N source as the precursor for N assimilation into macromolecules [228], and the presence of GltS may suggest an increased propensity for glutamate uptake among F-Pcy compared to their basal relatives among SE-Cy.

The AAT variety in LF-Cy is greater than in F-Pcy, a reflection of different evolutionary histories, greater genome sizes, and divergence within the Macrocyanobacteria. While F-Pcy can take up neutral and acidic AAs, these strains appear to lack the capacity to import basic AAs such as lysine and arginine. However, histidine as a weak basic AA, can be taken up by neutral AATs in most environmental conditions [253]. The Bgt transporter, preferentially taking up lysine and arginine, is the only basic AAT to be characterised in cyanobacteria and is found only in LF-Cy [254]. The differential presence of N-III primarily in F-Pcy and Bgt in LF-Cy may have implications for the composition of freshwater cyanobacterial communities. Lysine is one of the most abundant AAs on the lake sediment floor [421], while arginine is relatively abundant at the water surface [171, 278]. Disturbance of lake sediment and nutrient resuspension can occur through various physical mechanisms, including wind and storm-events [424]. Upon disturbance, the expression of Bgt may provide selective advantages for LF-Cy in the utilisation of basic AAs, while the abundance of neutral AAs in the limnetic may promote picocyanobacteria growth. As such, both DON composition and wider physical factors could influence the cyanobacterial community.

The diversity of AATs in M-Pcy are lower than in their freshwater counterparts. The *Syn/Pro* M-Pcy encode a charged AAT (N-II), while the β-*Synechococcus* encode two neutral AATs, suggesting a limited range of bioavailable AAs. These differences are likely due to their evolutionary histories - the β-*Synechococcus* derived from Macrocyanobacteria while *Syn/Pro* M-Pcy are derived from Microcyanobacteria - and subsequent specialisation into radically different habitats. β-*Synechococcus* are generally isolated from estuarine environments while *Syn/Pro* M-Pcy are typically found in the open ocean [425]. The concentration of DON increases significantly towards coastal areas, while the variability of DON composition also increases as a result of varying land use and water runoff [426]. As neutral AAs are widespread, the presence of N-I and N-III in β-*Synechococcus* suggests a large pool of AA N sources for these strains. On the other hand, glutamate and aspartate are some of the most abundant AAs in the pelagic [422], providing a large source of N for the N-II carrying *Syn/Pro.* Since M-Pcy have a high uptake rate of DON [427], the lower diversity of encoded AATs may be countered by the broad uptake capabilities of prevalent AAs facilitated by the N-II transporter. This may be linked to greater diversity of DON in fresh water, providing a larger pool of potential N sources and requiring more diverse AATs for F-Pcy [428, 429]. Alternatively, extracellular AA oxidation into NH_4^+ has been identified in marine environments with an abundance of the diazotroph *Trichodesmuim*, suggesting other mechanisms for AA assimilation for M-Pcy; however, the concentration of found AA oxidase activity is low, in addition to general DFAA concentrations. This casts doubts on the capability of extracellular AA oxidation in sustaining the large communities of *Prochlorococcus* and *Synechococcus* in marine environments, though this method of uptake may partially contribute to N requirements.

Among picocyanobacteria, *Prochlorococcus* spp. belonging to the high light sub-clades (HL-I and HL-II; Figure 3.1) are unique in that they do not encode any identified AATs (in addition to lacking transporters for NO_3^-/NO_2^-). These strains inhabit the epipelagic of tropical ocean systems, an environment where DON is present in greater abundance than inorganic N sources [430]. The absence of AATs suggests that these strains cannot utilise AA as a N source, in contrast to other *Prochlorococcus* ecotypes [431]. Meanwhile, their encoding of urea importers and metabolic enzymes indicates an alternative uptake focus. Furthermore, although NH_4^+ is present at low concentrations compared to DON in the marine environment, it is the most abundant form of dissolved inorganic N [432]. This suggests that the genomic streamlining of HL *Prochlorococcus* results in a limited pool of bioavailable N (namely NH_4^+ and urea) due to specialisation to low concentrations of nutrients. However, recent research has putatively identified HL *Prochlorococcus* strains with PAAT family AATs [385], suggesting the presence of novel AATs and additional AA uptake capabilities in marine environments than previously thought.

3.5.1.2 The Bioavailability of the Total DON Pool is Expanded Due to Putative Dissolved Combined Amino Acid and Chitin Uptake

DON is a substantial nutrient source in fluvial and limnic systems, often comprising more than 50% of the total N [433]. It encompasses a wide range of naturally occurring compounds, with dissolved combined AAs (DCAAs) seasonally constituting a considerable portion of total DON [434]. Concentrations of DCAAs ranging from 1 - 5 µM have been observed in eutrophic lakes, while oligotrophic environments typically exhibit DCAA concentrations between 0.3 - 2.5 µM [435]. Although the putative dipeptide and oligopeptide transporters have yet to be experimentally characterised in cyanobacteria, their expression (and expected function) would mitigate the requirement for extracellular enzymatic degradation of DCAA to DFAAs. This not only reduces the energy costs associated with excessive enzyme production and export, but is competitively advantageous by reducing potential uptake time. The dipeptide transporter, DppABCD, has been characterised in *Pseudomonas aeruginosa* PA14, demonstrating the uptake and assimilation of dipeptides and tripeptides as N sources [262]. As the function of DppABCD is inferred from *Pseudomonas*, further investigation is needed to characterise DppABCD in cyanobacteria; however, the presence of genetic homology suggests similar Dpp activity which will significantly increase DON bioavailability.

Chitin has emerged as a novel N source which has attracted recent attention due to its potential bioavailability. This study has identified two mechanisms of chitin assimilation in picocyanobacteria which are generally absent from LF-Cy, indicating the presence of a significant pool of differentially bioavailable N. It has been demonstrated that chitin can serve as a sole N source for cyanobacteria, though studies have provided mixed results. Growth of *Synechococcus* sp. WH7803 (encoding *chiA*) on chitin was not observed [389], however other studies have demonstrated chitinase activity in multiple marine *Synechococcus* and *Prochlorococcus* strains [436]. Interestingly, recent studies have linked chitin utilisation to the spread of M-Pcy in the open ocean, forming 'chitin rafts' to enable a mat-like environment for M-Pcy ancestral pioneers, supporting a key role of chitin in *Syn/Pro* speciation [436]. The enrichment of *chiA* in M-Pcy compared to F-Pcy may reflect the abundance of chitin in marine environments compared to fresh waters [437–439], and subsequent loss of chitinase utilisation following the radiation of picocyanobacteria back to freshwater environments [108]. In fresh waters, the diversity of N sources may have diminished metabolic benefits of direct chitin utilisation [426]. On the other hand, pathways to assimilate chitin via initial conversion to chitosan is present in freshwater environments, predominantly in F-Pcy. This mechanism may be more widespread than anticipated, also found in diatoms [440], yet the advantage of indirect chitosan utilisation over chitin remains to be clarified. Should chitin be bioavailable through chitosan conversion and hydrolysis for F-Pcy, this would provide access to an abundant nutrient source which is widespread throughout trophic status, potentially contributing to picocyanobacteria dominance in oligotrophic environments [438].

Another novel source of N is glyphosate [441]. This herbicide poses a distinct ecological risk due to extensive use and subsequent leaching into rivers, yet cyanobacteria have been shown to utilise it as a N source [395]. Picocyanobacteria encode a phosphonate transporter, which is absent from LF-Cy and SE-Cy, suggesting a divergent response to glyphosate infiltration. As a potential source of co-limiting N and P, the ability to assimilate low doses of glyphosate is highly advantageous. However, recent research has demonstrated that the addition of glyphosate results in *Microcystis* growth [442], a genus lacking a phosphonate transporter, suggesting alternative pathways of glyphosate assimilation are present. While the mechanism for glyphosate uptake remains unclear, this pervasive molecule may substantially impact microbial communities and contribute to the proliferation of harmful cyanobacterial blooms [396].

3.5.1.3 Mechanisms for the Assimilation of Traditional Organic and Inorganic Forms of Nitrogen Are Consistent With Known Previous Research

Urea has become one of the most common forms of fertiliser applied in the UK, second only behind ammonium nitrate [143]. However, urea itself is now being identified as highly bioavailable, with transporters which substantially enhance its uptake to biologically significant levels and prevalent catabolic machinery for its metabolism [12, 184]. The absence of the *urt* and *ure* operons from 10 cyanobacteria in this study, primarily from the SE-Cy and low-light *Prochlorococcus* spp., is particularly notable. While the streamlined genomes of *Prochlorococcus* spp. may explain the lack of urea assimilation in these strains, the genomes of SE-Cy are sizable. Urea offers an energetically advantageous source of N compared to inorganic sources, breaking down into two NH_4^+ molecules, thereby yielding a two-fold increase of N while providing CO_2 as a by-product [443]. It is unclear why SE-Cy lack the *urt* and *ure* operons; however, their inability to assimilate urea may be compensated by their additional variety of encoded AATs compared to F-Pcy. This work and previous research has identified genomes encoding the *urt* operon but lacking *ure*, primarily T-Pcy [184]. While there is the possibility that urease-activity may be present in an as yet unidentified enzyme, the inability to catabolise urea (which can be toxic in large amounts if accumulated inside the cell [444]) suggests that UrtABCDE may be associated with the transport of additional compounds.

N_2 fixation is rare in picocyanobacteria. The presence of the *nifHDK* operon has previously been identified in the volcanic lake-dwelling F-Pcy strain *Vulcanococcus limneticus* LL and two T-Pcy isolated from Octopus Spring in Yellowstone [33, 133, 445]. While nitrogenase was not expressed under N starvation in *Vulcanococcus limneticus* LL, the two T-Pcy express nitrogenase and fix N_2 during dark conditions [446]. The majority of N_2-fixing picocyanobacteria are found in thermal environments, suggesting large selection pressures for N_2 fixation capabilities in these habitats. This is supported by previous research demonstrating increasing temperatures leading to community shifts towards N_2-fixing bacteria and low concentrations of NO_3^- and NH_4^+ in hydrothermal environments [447, 448]. In freshwater environments, this study has identified an additional picocyanobacterium with the genetic capability for N_2 fixation (*Vulcanococcus limneticus* 3B3) alongside *Vulcanococcus limneticus* LL; interestingly, the presence of N_2-fixing machinery in *Vulcanococcus limneticus* 3B3, isolated from a eutrophic glacial lake in Northern Italy, is not linked to thermal water. *Vulcanococcus limneticus* strains may have obtained the *nif*-encoding sequences through HGT from N_2-fixing strains though the competitive benefit of the heavily energy-demanding nitrogenase in a glacial eutrophic system is unclear [33]. High ATP and electron requirements for nitrogenase suggest that N_2-fixation is primarily present in environments which lack other forms of N. This may have resulted in the loss of other forms of N acquisition capabilities, however this is generally not seen, with the exception of the absence of the Amt1 transporter in several N_2-fixing LF-Cy.

NH_4^+ is the most reduced and energetically favourable form of inorganic N. As such, its uptake capabilities are found in all habitats and morphologies, though the presence of NH_4^+ uptake accessory proteins are environment dependent. F-Pcy and LF-Cy encode various accessory proteins that facilitate NH_4^+ uptake at lower concentrations. For instance, AmtB (predominantly found in F-Pcy) increases NH_4^+ uptake by 55% [221] while *amt2*, *amt3*, and *amt4* (found in LF-Cy) have a limited impact (5%) on NH_4^+ uptake capabilities [220, 222]. The absence of these accessory proteins in M-Pcy is interesting; lower concentrations of NH_4^+ are found in marine environments yet these *Syn/Pro* strains lack a mechanism to boost NH_4^+ uptake when limited - a feature found in F-Pcy. As M-Pcy dominate oligotrophic regions of the open ocean, it suggests Amt1 alone is sufficient for NH_4^+ uptake in this environment. However, the presence of AmtB in F-Pcy indicates these picocyanobacteria can take up NH_4^+ at lower concentrations than M-Pcy, potentially supporting growth in ultra-oligotrophic freshwater environments, such as high-altitude and large volume lakes [449, 450].

The distribution of NO_3^- and NO_2^- transporters in picocyanobacteria is consistent with previous research [11, 217]. The NRT system is present in freshwater, while M-Pcy generally utilise NrtP and FocA. On the other hand, *Prochlorococcus* spp. exhibit interesting distribution patterns [217]. High light strains lack NO_3^- and NO_2^- transporters, relying on NH_4^+ and urea as N sources, which is likely a response to reduced levels of NO_3^- and NO_2^- at the ocean surface [451]. Meanwhile, low light strains encode the *focA* gene, allowing them to import NO_2^- which accumulates in deeper layers of the ocean due to the activity of nitrifying bacteria [452].

3.5.1.4 The Inducement of Nitrogen Assimilation Regulatory Mechanisms and the Ability to Store Nitrogen Can Alter Freshwater *Synechococcus* Predominance

The regulation of N assimilation, involving the activation of NtcA through PII and PipX, is conserved across habitats and morphologies. NtcA activation during periods of N limitation leads to the expression of various N assimilatory genes. These include the *nir* operon, responsible for NO_3^- and NO_2^- uptake and metabolism; *amt1*, involved in NH_4^+ uptake; *glnA* for N incorporation; and the *urt* and *ure* operon, involved in urea assimilation, among others [453]. NtcA binding sites are also found in AAT and AA biosynthesis genes, ensuring N-regulated control over the delicate balance of N uptake and incorporation [213]. An additional transcription factor (NtcB), involved in N assimilation regulation, has been identified in LF-Cy, SE-Cy and non-*Syn/Pro* picocyanobacteria. NtcB binding motifs have been identified upstream of nitrite reductase (*nirA*) and *nrtP* in the marine β-cyanobacterium *Synechococcus* sp. UTEX 2380 [454]. NtcB is associated with the NO_2^--responsive activation of NO_3^- and NO_2^- assimilatory genes [216]. The absence of *ntcB* from the *Syn/Pro* clade is surprising considering the competitive benefits of NtcB in the utilisation of N in NO_3^--limited water and the oligotrophic environments in which F-Pcy thrive. The differential presence of *ntcB* in *Synechococcus* (F-Pcy vs. SE-Cy) may influence cyanobacterial communities when NO_2^- is present but limited, potentially favouring the growth of SE-Cy over F-Pcy.

The ability of picocyanobacteria to produce and utilise N stores is significantly different from LF-Cy and SE-Cy. These strains are capable of synthesizing cyanophycin, an N-storage polymer of arginine and aspartate, while F-Pcy, *Syn/Pro* M-Pcy, and T-Pcy lack this capability. The distribution pattern of the capability of phycobilisome degradation via NblA [455] mirrors that of cyanophycin synthetase, suggesting redundancy in N storage capabilities and the absence of N storage utilisation in the majority of picocyanobacteria. This may be due to size constraints within the cellular interior, with the presence of cyanophycin synthetase in the β-*Synechococcus* M-Pcy potentially suggesting reduced spatial limitations in this pico-convergent clade [456]. Alternatively, the utilization of diverse N sources to prevent starvation, coupled with a rapid response to nutrient influx and a high surface area to volume ratio, may diminish the necessity of N storage in picocyanobacteria. Meanwhile, the multiple mechanisms for N storage in LF-Cy may be a result of higher energy requirements and a slower response to nutrient input, necessitating N storage to ensure survival under fluctuating environmental conditions.

3.5.2 Phylogenetic Affinity Largely Determines Nitrogen Assimilation Capabilities

The divergence of *Syn/Pro* picocyanobacteria during the Neoproterozoic ($\sim$ 599 Mya), along with the earlier divergence of T-Pcy around 2,500 Mya [94, 108], represent significant milestones in picocyanobacterial evolution. While T-Pcy adapted to exploit resources around hot springs, the *Syn/Pro* are thought to derive from a marine ancestor, later radiating back into freshwater environments [108]. The molecular machinery for N assimilation encoded in these groups is the result of millions of years of history. Different strategies have evolved across the cyanobacterial tree of life to utilise environmental N and some patterns can be identified when comparing the *Syn/Pro* to Macrocyanobacteria. Notably, the distribution of NtcB suggests this important mechanism for N assimilatory activation evolved following *Syn/Pro* divergence, a distribution also found in *oppC*. Additionally, differences in N assimilation capabilities between M-Pcy from the *Syn/Pro* and M-Pcy from the Macrocyanobacteria (*β-Synechococcus*) represent distinct evolutionary histories of these two phylogenetically-distant groups, and highlights the convergent evolution of 'pico' morphology while encoding differing assimilation mechanisms. Interestingly, no gene involved with N assimilation is found solely encoded by the *Syn/Pro*. This may suggest that the combination of surface area to volume ratio and variety of bioavailable N sources represent the primary advantages of *Syn/Pro* proliferation in the context of N, yet another explanation may be due to the scarce research towards F-Pcy and novel assimilatory methods in this clade. On the other hand, marked patterns of N assimilation are found within closely related groups, for instance comparing marine and freshwater picocyanobacteria from the *Syn/Pro* clade. These groups differ in their encoding capabilities of NH_4^+, NO_3^-, NO_2^-, AAs, and chitin. Due to freshwater *Syn/Pro* deriving from their marine relatives, it is unclear if the genes currently present in F-Pcy were subsequently lost in M-Pcy following F-Pcy speciation, or if horizontal gene transfer resulted in the acquisition of these genes from the Macrocyanobacteria.

3.6 Conclusions

This study has increased our understanding of the N assimilation capabilities of F-Pcy and how they compare to SE-Cy, LF-Cy, and picocyanobacteria from different habitats. It sheds light on how molecular machinery, alongside morphological factors, influences the cyanobacterial community and may contribute to the dominance of F-Pcy in N-limited oligotrophic and mesotrophic environments. The differential encoding of AATs represent a major source of N assimilation diversity among freshwater cyanobacteria, with AAs taken up via combinations of four ABC-type transporters and two glutamate-specific transporters. Though each transporter can import a range of AAs based on their charge and hydrophobicity, transporter uptake preferences favour the import of specific AAs, potentially influencing growth rates and cyanobacterial dominance depending on DFAA composition. Freshwater DON composition is highly diverse and heterogeneous [426], potentially driving community composition in low inorganic N environments. This selection process may favour the dominance of either F-Pcy or LF-Cy depending on dominant DON sources. Additional bioavailable DON sources include DCAA, chitin, and glyphosate, with the mechanisms for their assimilation varying between F-Pcy and LF-Cy.

In addition to the differences between F-Pcy and LF-Cy, notable variations are found in the molecular mechanisms of inorganic and organic N assimilation among picocyanobacteria. This displays the variety within picocyanobacterial mechanisms of N assimilation but also illustrates the major contrasts between environments, mirroring nutrient availability and competition. The consistency and homogeneity of N sources in marine environments has enabled niche adaptation in *Prochlorococcus*, while *Synechococcus* spp. encode a greater range of assimilatory mechanisms. However, the freshwater N pool is less consistent and more dependant on external factors, resulting in a more diverse pattern of potentially bioavailable N for F-Pcy. In contrast to these, the extreme environment of T-Pcy, and lack of other suitable N sources, has resulted in the energetically demanding capability of N_2-fixation for a reliable source of N. Therefore, the molecular mechanisms of N assimilation of picocyanobacteria represent adaptations to open ocean, estuarine, thermal, or freshwater conditions, and demonstrates their ability to dominate these environments.

3.7 Limitations

Comparative genomic analyses provide valuable insights into the molecular machinery encoded in the genome. However, it is essential to validate the activity of these genes through experimental characterisation to ensure these genes are functional and not pseudogenes. For example, in the case of the freshwater picocyanobacterium *Vulcanococcus limneticus* LL, which encodes a complete set of the *nifHDK* operon, these genes were not transcribed under both N limitation and N replete conditions, suggesting that they are not functional in this species (though it should be noted that Di Cesare *et al* only tested *nif* activation under one set of specific conditions) [33]. Similarly, in the marine *Synechococcus* sp. WH7803, although it encodes *chiA* which is associated with chitin degradation, it was unable to grow on chitin, indicating either non-functioning of the encoded enzyme or the lack of glucosamine uptake and assimilation in this strain [389]. These examples emphasize the importance of experimental characterization to establish the link between genomic content, gene functionality, and the ecological response to proposed assimilation pathways.

Genes screened in this study were selected based on proteins which have been previously experimentally characterised. While there are a numerous studies investigating nutrient assimilation proteins in marine picocyanobacteria, the majority of freshwater research has been carried out in non-picocyanobacteria model organisms, commonly *Synechocystis*, *Anabaena*, and *Synechococcus elongatus* [374, 457–459]. This limits our knowledge of potentially novel assimilation mechanisms available in the *Syn/Pro*, though other studies have identified putative AATs in *Prochlorococcus* inferred from an intraprotein PBPd domain [385]. Experimentally unknown putative transporters were not within the scope of this study, thus potentially under-estimating the complete picocyanobacteria N diet.

The primary finding in this study is the diversity of AATs in cyanobacteria and the influence this may exert on their community composition. However, AAT diversity may have evolved in response to the uptake of leaked AAs rather than as a result of allochthonous external nutrient sources. In heterocystous N-fixing cyanobacteria, N is transferred from heterocysts to vegetative cells in the form of AAs, mostly arginine, glutamate, glutamine, and aspartate, with AATs thereby fulfilling the important role of taking up the newly-fixed N [294]. Furthermore, in both heterocystous and non-heterocystous cyanobacteria, AAT knockout mutants accumulate extracellular hydrophobic AAs [253].

Thus, AATs have an important role in preventing excessive AA release in cyanobacteria. Additionally, AAs may also be utilised through extracellular hydrolysis. Though the environmental concentration of DFAA indicates active uptake of this substrate is required, alternative mechanisms include the secretion or cell surface presence of amino oxidases, capable of hydrolysing AAs into NH_4^+ which can be subsequently imported via Amt1 [131]. However, the reduction of AA uptake in AAT knockouts suggests that this import mechanism is inefficient and unlikely to result in sustainable AA growth.

P is a co-limiting nutrient alongside N in fresh water. Therefore, the assimilation of sufficient P sources from the environment is also essential to support cyanobacterial growth [6]. It has been demonstrated that the phosphonate transporter (*phnD*) is constitutively expressed in F-Pcy, though the diversity of P sources and other mechanisms for P assimilation have not been extensively studied [460]. To achieve a holistic understanding of the impact of different nutrient sources on the composition of the cyanobacterial community, further research into synergistic co-limitation is required.

3.8 Future Directions

This study has identified the N assimilation capabilities encoded by F-Pcy, with a focus on organic sources of N. The diversity of AATs may suggest that growth on different AAs results in different growth rates among cyanobacteria, influencing dominant genera and the composition of the cyanobacterial community. This work solely utilised comparative genomics to identify potential differences. Further research carried out in this thesis used the results from this chapter to investigate how F-Pcy respond to different AA N-sources. A proteomic approach is utilised in chapter 4 to identify differential protein expression when grown on various AA sources. This expands our knowledge of the *Syn/Pro* response to DON, in addition to beginning to elucidate how DON composition may influence primary producers.

To understand how nutrient composition influences cyanobacterial community, experiments in natural settings must be carried out in addition to the laboratory. Mesocosm spiking of a mixed community with various organic N sources would provide some real-world application for the response to DON composition. Community change, detected through metagenomic analysis, will aid in revealing nutrient influence, while screening for cyanotoxins may provide insight into harmful bloom production in DON-heavy waters, enabling accurate predictions of toxin generation.

Chapter 4

The Picocyanobacterial Response to Growth on Amino Acids

4.1 Contributions and Acknowledgements

This chapter was written with feedback from Patricia Sánchez-Baracaldo and Stephen Maberly. Fluorescence-activated cell sorting was carried out by Andrew Herman at the Flow Cytometry Facility, University of Bristol. TMT-10plex quantitative proteomics was carried out at the Proteomics Facility, University of Bristol with the aid of Kate Heesom. All other work is my own.

4.2 Introduction

Synechococcus spp. of the *Syn/Pro* clade are major primary producers in freshwater, especially oligotrophic environments where these picocyanobacteria constitute a large proportion of total biomass [70]. In both oligo- and eu-trophic environments, nitrogen (N) can be a limiting or co-limiting nutrient, determining the extent of *Synechococcus* productivity and distribution [9, 388]. Sources of N can be divided into organic and inorganic forms. Prevalent inorganic sources of N consist of NO_3^-, NO_2^-, and NH_4^+ with these N forms commonly monitored in littoral and fluvial systems [461]. In contrast, though organic sources of N (DON) constitute the largest pool of fixed N in many aquatic ecosystems (on top of their carbon (C) content), these forms are often overlooked [131]. DON consists of characterised components such as urea and amino acids (AAs) (combined and free), and a large proportion of humic substances of unknown chemical structure [14]. Of the characterised components, urea has recently attracted attention due to its increased role in fertiliser applications and links to cyanobacterial blooms [462]. Meanwhile, though previous research has revealed the prevalence and diversity of AAs in the DON pool [434, 463], the molecular mechanisms behind the assimilation capabilities of AAs are less well understood.

Knowledge about the utilisation of AAs as cyanobacterial N and C sources has a long history, yet the AA assimilation capabilities of freshwater picocyanobacteria remain unstudied. AA growth assays have been conducted with filamentous cyanobacteria [164–166], unicellular freshwater macrocyanobacteria [165, 464], and marine picocyanobacteria [164, 465]. These studies have displayed a mixed picture of cyanobacterial AA assimilation capabilities, with growth on AAs as sole N sources differing between genera and species [252], as seen in Table 1.1. Similar investigations focusing on freshwater picocyanobacteria have been hindered by the lack of a model freshwater picocyanobacterium and few studies on the *Syn/Pro* clade [100, 108]. Similarly, our understanding of cyanobacterial metabolism is primarily based on the model cyanobacteria *Synechocystis* sp. PCC6803 [189, 266], *Anabaena*/*Nostoc* sp. PCC7120 [191, 203], and *Synechococcus elongatus* PCC7942 [189, 203]. However, our current knowledge is limited with regards to AA assimilation and catabolic pathways. Many arginine degradation pathways have been identified in *Synechocystis*, yet none are fully encoded in *Synechococcus elongatus* [266], while cyanobacteria lack the histidine utilisation system for histidine degradation [295]. Likewise, the biosynthesis and degradation pathways of aromatic AAs are not fully understood, while the methionine biosynthetic precursor is yet to be determined [189, 466].

The capacity of cyanobacteria to utilise AAs as sole N substrates is primarily dependent on their capability to import these molecules across the plasma membrane. ATP-binding cassette (ABC) transporters are the primary transport system for AA import in cyanobacteria, with four AA transporters (AATs) characterised, two of which are found in freshwater picocyanobacteria (Chapter 3). These are the N-II and the N-III transporters. Furthermore, cyanobacteria also encode glutamate-specific transporters, with GltS present in freshwater picocyanobacteria (Chapter 3, Figure 3.3). In the previous chapter, diversity in the AATs encoded within cyanobacteria was identified. Differences were found within picocyanobacteria between habitats, differences between *Synechococcus* and the non-picocyanobacteria *Synechococcus elongatus* strains, and differences between freshwater picocyanobacteria and larger freshwater cyanobacteria. Cyanobacterial AATs are specific for certain characteristics, e.g. neutral AAs, charged AAs, in addition to having preferred substrates for uptake [255]. This variety in encoded AATs may represent different capabilities in assimilating AAs, with the specific complement of encoded AATs resulting in the enhanced or reduced ability to grow on specific AA substrates. In addition to active uptake, other methods not based on direct import, such as secretion of amino acid oxidases, have been proposed for AA assimilation though this has only been identified to a limited extent during cyanobacterial blooms in marine environments [467]. The possibility of extracellular AA catabolism and subsequent import of inorganic or more accessible organic N forms in oligotrophic freshwater environments must not be ruled out, though the typically low concentrations of DFAA (and extracellular AA oxidase activity) in aquatic environments suggest active and direct import is the primary method of AA assimilation [468].

Genomics and proteomics are parallel methods to understand the molecular potential of a cell. Genomics utilises the genetic code as its foundation, inferring the capabilities of cellular action from the gene products (i.e. proteins). In contrast, proteomics focuses on the expressed protein and uses cellular context (e.g. environmental conditions) to determine the genetic response. Utilising genomics and proteomics reveals the full molecular capacity encoded by an organism and identifies expression patterns of these gene products. The rise of isobaric tandem mass tag (TMT) labelling, a technique that labels peptides with molecules of identical mass but differing distribution of heavy isotopes (yielding distinct reporter ions after fragmentation), has improved multiplexed proteomics, enabling globally quantified proteins across multiple samples in a single experiment [469]. Furthermore, while genomics can identify encoded genes, it is important to verify these genes can be expressed and are not pseudogenes - the N-fixing *nifHDK* operon is encoded in *Vulcanococcus limneticus* LL yet no expression of this operon has been detected [33], whether due to an inactive genes, not reproducing necessary environmental conditions, or lacking an appropriate microbial consortium [470]. The use of proteomic quantification to investigate cyanobacterial environmental response has increased in recent years, with the proteomic response to elevated CO_2 levels [471], rising temperatures [472], and exposure to glyphosate and antibiotics [473, 474], among others, explored. However, the proteomic response to varying forms of N have focused on NO_3^-, NH_4^+, and urea, ignoring potential AA N-sources [472, 475].

This study aims to investigate the AA assimilation capabilities of freshwater picocyanobacteria. Due to the differential encoding of AATs among cyanobacteria, we hypothesised that the preferred substrates of AATs (aspartate, glutamate, glycine) would result in greater growth rates compared to other AAs. We further hypothesised that growth on basic AAs, especially arginine and lysine, would not be viable due to the lack of uptake capability of these AAs among the N-II and N-III AATs. To accomplish this, axenic cultures of the freshwater picocyanobacteria *Synechococcus* sp. CCY9618 and *Synechococcus* sp. CCAP1479/10 were generated. Growth assays were then carried out with these cultures utilising 20 proteinogenic AAs as sole N sources, along with other novel and traditional sources of N. Following this, proteomic analysis was carried out on *Synechococcus* sp. CCAP1479/10 to conduct a novel investigation of the proteomic response of freshwater picocyanobacteria to growth on four AAs as sole N sources - arginine, asparagine, proline, and glutamate - compared to NO_3^-. We hypothesised that the N-II and N-III AATs would display differential expression between AA substrates, in addition to differential up-regulation of varying AA biosynthetic and degradative pathways to incorporate the assimilated AAs. Our results display widespread AA assimilation bioavailability in freshwater picocyanobacteria, with interspecific variation present among freshwater picocyanobacteria in regards to adjustment time for various nutrients, i.e lag time. Proteomic analysis putatively revealed a mild stress response under all AA conditions and potential mechanisms for arginine assimilation.

4.3 Methodology

4.3.1 Generation of Axenic Cultures

Five freshwater (salinity < 0.5 ppt) picocyanobacteria strains were obtained from Culture Collection Yerseke (The Netherlands; *Synechococcus* sp. CCY0621 and *Synechococcus* sp. CCY9618) and the Culture Collection of Algae and Protozoa (UK; *Synechococcus* sp. CCAP1479/9; *Synechococcus* sp. CCAP1479/10, and *Synechococcus* sp. CCAP1479/13). These strains were mono-phototrophic yet non-axenic. Strains were maintained in BG-11 media [332] under 10 - 20 µmol m^{-2} s^{-1} of white LED light (in the spectral band of 400 - 700 nm) with a 16 h: 8 h light:dark cycle at 20 °C. For information of failed methods used to generate axenic cultures, see Appendix C.1.

4.3.1.1 Fluorescent-Activated Cell Sorting

Fluorescent-activated cell sorting (FACS) was carried out on the five *Synechococcus* strains to purify them from contaminating bacteria, based on previous research [476]. Samples were sent to the Flow Cytometry Facility in the School of Biomedical Sciences, University of Bristol. FACS sorting was performed using a BD Influx instrument (BD Biosciences, USA). The optical configuration, in brief, employed 200 mW 488 nm blue (light scatter), 50 mW 640 nm red (chlorophyll detection), and 100 mW 552 nm yellow (chlorophyll excitation and detection) lasers. A small-particle light detector provided high sensitivity in detecting forward scatter and a 0.45 threshold on a logarithmic scale was used. A 4 mm obscuration bar was used for optimal forward scatter light detection with a 100 µm diameter ceramic nozzle tip and 21 PSI sheath pressure. Cells of interest were gated on using forward scatter v. side scatter, and singlets by light forward scatter v. trigger pulse width to exclude aggregated cells. Red chlorophyll positive events were identified by fluorescence emission from the 552 nm and 640 nm lasers using 670/30 nm bandpass filters. Cell sorting was achieved using single cell cloning precision into 96 well polystyrene plates.

After FACS, the resultant plates were incubated in 200 µL BG-11 medium per well under 10 - 20 µmol m^{-2} s^{-1} of white LED light with a 16 h: 8 h light:dark cycle at 20 °C. Once picocyanobacteria growth was visible (well turning green), each culture was streaked onto BG-11 agar plates and incubated for 15 days. Following this, the culture was transferred to 10 mL BG-11 medium and incubated for a further 15 days. Strains were then streaked onto LB agar plates and incubated for three days in darkness to assess bacterial growth. Plates without bacterial growth contained axenic cultures of *Synechococcus*. Axenic cultures were produced for all five *Synechococcus* strains, with *Synechococcus* sp. CCY9618 and *Synechococcus* sp. CCAP1479/10 selected for further experimental use. These strains were selected as they encode differential AATs (CCAP1479/10 encodes N-II, N-III, and GltS; CCY9618 encodes N-II and N-III, but lacks the glutamate-specific transporter GltS) and were isolated from different environments with different trophic states (mesotrophic v eutrophic).

4.3.2 Growth Rate Measurements

Picocyanobacteria growth on various sources of organic and inorganic N was determined indirectly through optical density (OD) measurements. *Synechococcus* sp. CCY9618 (Vinkeveen, The Netherlands) and *Synechococcus* sp. CCAP1479/10 (Windermere, UK) were grown in 150 cm^2 vented flat-sided cell culture flasks containing 400 mL BG-11 media at 20 °C. The initial inoculum of each picocyanobacterium was 50 mL of culture with an OD_{750} of 0.5. Strains were grown for four days under 10 - 20 µmol m^{-2} s^{-1} of white LED light with a 16 h: 8 h light:dark cycle at 20 °C. Each culture was centrifuged for five minutes at 2,500 rpm, with the resultant pellet three times washed with N-free BG-11 medium [389]. The cultures were then cultivated for a further 24 hours to remove residual N. Triplicate 25 cm^2 vented flat-sided cell culture flasks were prepared for each strain and N source with 11 mL of N-free BG-11 medium and 1 mL of culture inoculum, supplemented with a sole N source. Sole N sources included organic (20 proteinogenic AAs, urea, chitin, glyphosate) and inorganic (NH_4^+ and NO_3^-) N sources (all of high-grade purity) (Table 4.1; Appendix Table B.17). Two N concentrations were utilised – one high concentration based on BG-11 media (250 mg N L^{-1}) and a lower, more environmentally relevant concentration (1 mg N L^{-1} [173, 477–479]). In total, two freshwater picocyanobacteria strains were grown in triplicate under 25 sole N sources at two concentrations, resulting in 312 individual samples (including N starvation negative control). Each flask was incubated for fourteen days under the conditions described above. Daily 200 µL aliquots were taken to determine OD at 750 nm, measured using a Multiskan SkyHigh Microplate Spectrophotometer (Thermo Scientific, USA). Growth on high concentrations of chitin did not yield accurate OD measurements due to the insolubility of chitin particulates interfering with photon transmittance. This resulted in OD measurements of only 24 substrates at high concentration. Furthermore, poor solubility of tyrosine in water necessitated a reduced high N concentration of 25 mg N L^{-1} for this condition. Growth rates and lag phase duration were determined using Growthcurver v3.0.1 [480]. Statistical analysis was carried out using a two-tailed t test with FDR-adjusted p-values.

Table 4.1: **Substrates utilised as sole nitrogen sources**

Organic			
Amino Acids (L)			
Alanine	Glutamate	Leucine	Serine
Arginine	Glutamine	Lysine	Threonine
Asparagine	Glycine	Methionine	Tryptophan
Aspartate	Histidine	Phenylalanine	Tyrosine
Cysteine	Isoleucine	Proline	Valine
Novel			
Chitin	Glyphosate		
Traditional			
Urea			
Inorganic			
Ammonium	Nitrate		

4.3.3 Protein Extraction and Quantitative Proteomics

4.3.3.1 Growth Conditions

Synechococcus sp. CCAP1479/10 was grown in triplicate in 400 mL BG-11 media at 20 °C under 10 - 20 µmol m^{-2} s^{-1} of white LED light in a 16 h light: 8 h dark cycle. The culture was incubated for four days before being harvested by centrifugation for five minutes at 2,500 rpm, and washed three times with N-free BG-11 medium. The culture was then resuspended in 400 mL N-free BG-11 media and incubated for 24 hours to deplete any residual N. Following N-free incubation, 1 mL of culture was inoculated into triplicate flasks containing 11 mL N-free BG-11 supplemented with 250 mg N L^{-1} of a sole N source (NO_3^-, arginine, asparagine, glutamate, proline) or no N for a total of six conditions and 18 flasks. Cultures were then incubated for two to five days, until exponential phase was reached (48 hours incubation for N-starvation condition). 2 mL aliquots were subsequently collected for protein extraction.

The four AA substrates were selected for proteomic analysis as they represent an array of chemical properties and transporter preferences identified in Chapter 3 (Figure 3.3). Glutamate is a known N source and directly incorporates into N assimilation pathways. *Synechococcus* sp. CCAP1479/10 encodes one glutamate-specific transporter (*gltS*), in addition to its uptake by N-II and N-III. The two neutral AAs selected are taken up by only one of the two general AATs - asparagine (neutral polar) by N-II

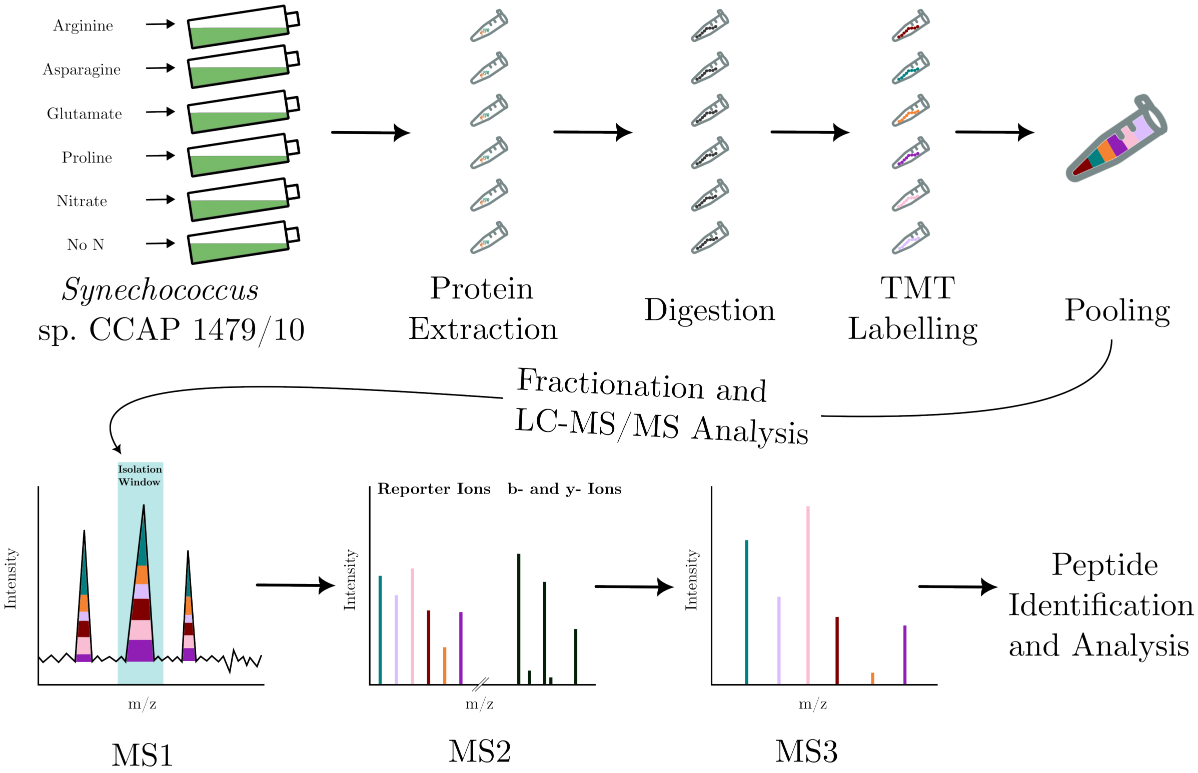

Figure 4.1: **Workflow of TMT-based quantitative proteomics to investigate the picocyanobacteria response to growth on selected amino acids.** Triplicates of *Synechococcus* sp. CCAP1479/10 were grown under a sole N source with protein content subsequently extracted. Samples were denatured, reduced, alkylated, and digested with resultant peptides labelled with TMTs. The pooled sample was then fractionated using high pH reverse-phase chromatography before being analysed through an Orbitrap Fusion Tribrid mass spectrometer under an SPS-MS3 workflow.

and proline (neutral non-polar) by N-III. Lastly, growth on a basic N source (arginine) was surprising owing to the lack of a positively charged AAT (absence of Bgt). The proteomic response to growth on these substrates was compared to growth on a traditional N source (NO_3^-) and to N-starvation.

4.3.3.2 Protein Extraction and TMT Labelling

Protein content was extracted from each sample using a NoviPure Microbial Protein Kit (Qiagen, Germany) according to the manufacturer's instructions. Protein concentration was determined using a Nanodrop Spectrophotometer 2000 (Thermo Scientific, USA) and sent to the Proteomics Facility at the University of Bristol for proteomic analysis according to the workflow in Figure 4.1. 30 µg of each sample was reduced with tris(2-carboxyethyl)phosphine (10 mM) and incubated for one hour at 55 °C. Alkylation of cysteine residues was carried out with iodoacetamide (17 mM) and incubated for 30 minutes in darkness. Samples were finally digested with trypsin (1.25 µg) at 37 °C overnight. Samples were labelled with TMT 10Plex reagents according to the manufacturer's protocol (Thermo Fisher Scientific, UK) and the labelled samples were pooled together.

4.3.3.3 High pH Reverse-Phase Chromatography

An aliquot of the pooled sample (100 µg) was evaporated to dryness and resuspended in 5% formic acid. A SepPak cartridge was used to desalt the solution according to the manufacturer's instructions (Waters, USA). The resultant eluate was prepared for fractionation by again evaporating to dryness before being resuspended in 20 mM ammonium hydroxide. Fractionation was carried out by high pH reverse-phase chromatography on an Ultimate 3000 liquid chromatography system (Thermo Scientific, USA). The sample was loaded onto an XBridge BEH C18 Column (130 Å, 3.5 µm, 2.1 mm x 150 mm; Waters, UK) with peptides eluted over a steadily increasing gradient of 20 mM ammonium hydroxide in acetonitrile over one hour. The fractions produced were again evaporated to dryness and resuspended in formic acid (1%) to prepare for further fractionation.

4.3.3.4 Nano-Liquid Chromatography Mass Spectrometry

Further fractionation was carried out with nano-liquid chromatography mass spectrometry using an Orbitrap Fusion Tribrid mass spectrometer (Thermo Scientific, USA). Prepared peptides were loaded onto an Acclaim PepMap C18 nano-trap column (Thermo Scientific, USA) and washed with 0.1% (vol/vol) formic acid in 0.5% (vol/vol) acetonitrile. Peptides were then resolved on an Acclaim PepMap C18 reverse phase analytical column (250 mm x 75 µm; Thermo Scientific, USA) with a 150 min organic gradient. Seven gradient segments were utilised alongside Solvent A (0.1% formic acid) and Solvent B (SB; aqueous 80% acetonitrile in 0.1% formic acid) with a flow rate of 300 nL min^{-1}. The seven gradient segments included: 1-6% SB for 1 min; 6-15% SB over 58 min; 15-32% SB for 58 min; 32-40% SB for 5 min; 40-90% SB over 1 min then held at 90% SB for 6 min followed by reduction to 1% SB over 1 min.

Nano-electrospray ionisation (2.0 kV) produced peptide ions using a stainless-steel emitter with a capillary temperature of 275 °C and an internal diameter of 30 µm (Thermo Scientific, USA). An Orbitrap Fusion Tribrid mass spectrometer was used to obtain all spectra, controlled by Xcalibur 2.1 (Thermo Scientific, USA) and ran under an SPS-MS3 workflow in data-dependent acquisition mode. FTMS1 spectra were collected at a resolution of 120,000 with a maximum injection time of 50 ms and an automatic gain control target (AGC) of 200,000. Precursor ions were filtered with an intensity threshold of 5,000 and monoisotopic peak determination was set to peptide. A dynamic window (60 s +/- 10 ppm) removed previously analysed precursors. A quadrupole isolation window (1.2 m/z) was used to isolate MS2 precursors. Resultant ITMS2 spectra were collected with a maximum injection time of 70 ms, AGC target of 10,000 and CID collision energy of 35%.

FTMS3 analysis was carried out at a resolution of 50,000, a maximum injection time of 105 ms, and an AGC target of 50,000. High energy collision dissociation at a normalised collision energy of 60% was used to fragment precursor ions and ensure optimal TMT reporter ion yield. The FTMS3 scan was able to include up to 10 MS2 fragment ions by enabling Synchronous Precursor Selection.

4.3.3.5 Data Analysis

Raw data files were quantified using Proteome Discoverer software v2.1 (Thermo Scientific, USA) and searched against the *Synechococcus* sp. CCAP1479/10 protein database (3,441 sequences) using the SEQUEST HT algorithm [481]. Precursor mass tolerance was set to 10 ppm and MS/MS tolerance set to 0.6 Da. Variable modifications (oxidation of methionine (+15.995 Da), acetylation of protein N-terminus (+42.011 Da), methionine loss + N-terminus acetylation (-89.03 Da)) and fixed modifications (carbamidomethylation of cysteine (+57.021 Da) and addition of TMT mass tag (+229.163 Da)) were included as search criteria. A threshold of 2 missed cleavages was set and full tryptic digestion searches were performed with reverse database search enabled. The data was filtered using a 5% false discovery rate cut-off and contaminants were identified against a ‘common contaminants’ database. Peptide data was normalised to the total peptide in each sample and scaled using a pooled sample common to all runs. Only proteins detected in all replicates were used for further analysis. ANOVA was used to determine significant enrichment among proteins, followed by Tukey’s Post-Hoc test (FDR-adjusted) to determine significance between conditions. Volcano plots were generated by transforming P-values (-$\log_{10}$) and fold change ($\log_2$). Differentially expressed proteins (DEPs) were deemed statistically significant with a Q-value less than 0.05 and a $\log_2$ fold change greater than 0.5/less than -0.5. Proteins were functionally annotated using eggNOG [343] and pathway enrichment analysis was carried out using KEGG [344] and hypergeometric distribution tests.

4.4 Results

4.4.1 Most Amino Acids are Bioavailable with Limited Differences in Maximum Specific Growth Rate

The bioavailability of organic N sources was investigated using two freshwater picocyanobacteria strains (*Synechococcus* sp. CCY9618 and *Synechococcus* sp. CCAP1479/10) under a high (250 mg N L^{-1}) and a low (1 mg N L^{-1}) N concentration. Twenty-five substrates were tested as sole N sources, supplemented into N-free BG-11 growth medium. The growth curves for each substrate at high N concentrations are shown in Figure 4.2 while growth curves for both concentrations are displayed in Appendix Figure A.10. Most of the tested substrates exhibited some degree of bioavailability and supported the growth of both picocyanobacteria strains under both high and low N concentrations. However, two AAs, cysteine and threonine, prevented growth entirely. Interestingly, limited tyrosine bioavailability was demonstrated only for *Synechococcus* sp. CCAP1479/10 at a high concentration, whereas methionine was able to be utilised, to some extent, only by *Synechococcus* sp. CCY9618.

The two novel sources of organic N tested in this study (glyphosate and chitin) proved unavailable (or unquantifiable) in picocyanobacteria. No growth was measured under glyphosate; while this was to be expected at high concentration due to its toxic nature, no growth under a lower nutrient load indicates high sensitivity to glyphosate or the requirements of even lower dosages for effective bioavailability. Chitin at low concentration did not result in measured picocyanobacteria growth. At high concentrations of chitin, OD_{750} measurements could not be accurately quantified due to chitin particulate occlusion to the spectrophotometer, though some greening of the culture was seen. However, due to the unquantifiable nature of this, growth rates and duration of lag phase on chitin growth are not within the scope of this study.

The use of OD to quantify picocyanobacteria growth is an indirect method of determining cyanobacterial yield, without providing a direct measure of cell count. The greatest picocyanobacteria yields were achieved under high concentrations of N, with *Synechococcus* sp. CCY9618 using aspartate as a substrate (0.1489 ± 0.0026) and *Synechococcus* sp. CCAP1479/10 utilising proline (0.2121 ± 0.0024) (Tables 4.2 and 4.4). When grown under a reduced concentration of N, both strains displayed similar OD

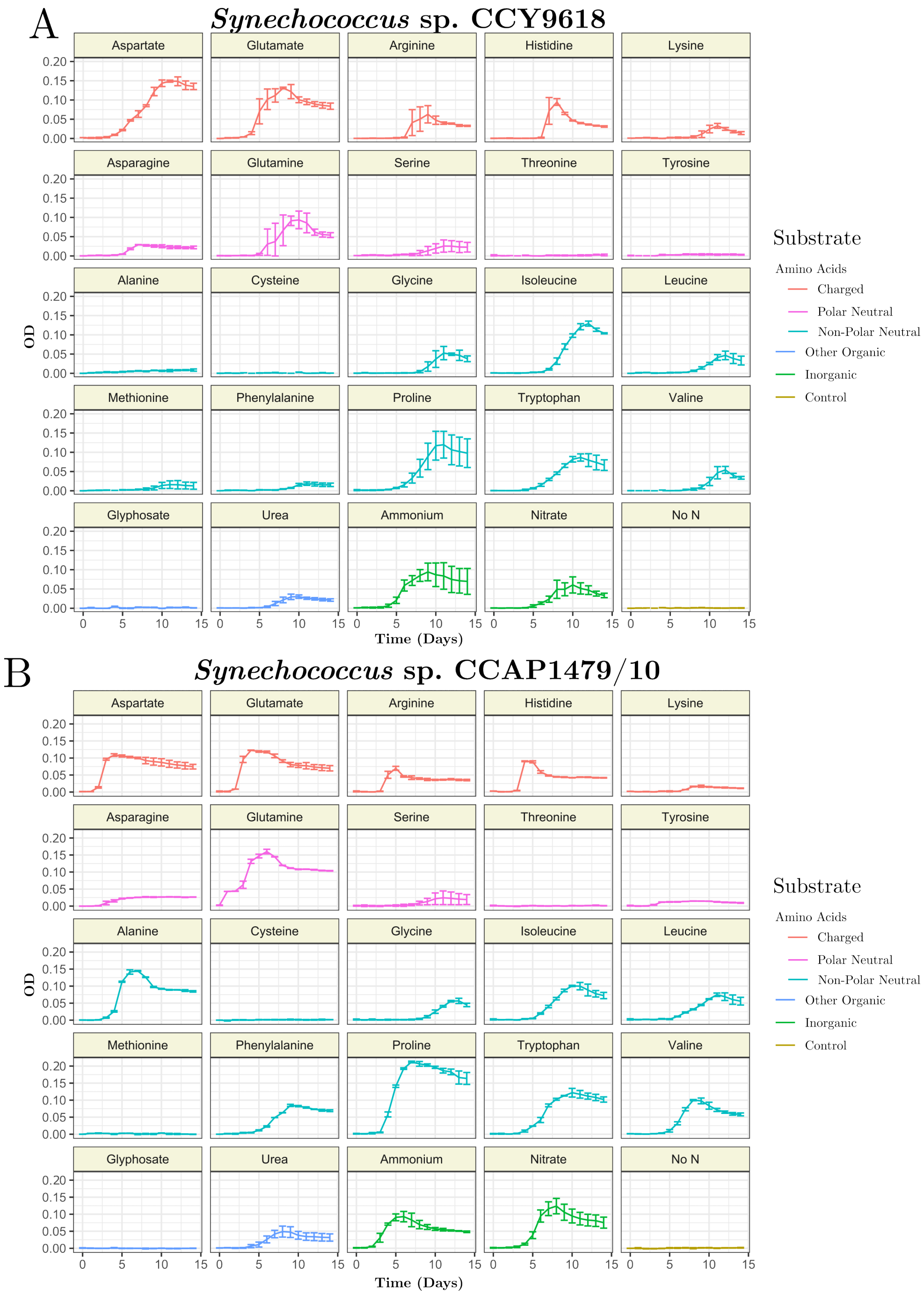

Figure 4.2: **Growth curves of *Synechococcus* sp. CCY9618 and *Synechococcus* sp. CCAP1479/10 cultivated under different nitrogen substrates.** A) *Synechococcus* sp. CCY9618; B) *Synechococcus* sp. CCAP1479/10. Substrates are divided into AAs (subdivided into properties), other organic sources, and inorganic sources (see figure inlay). A high concentration of N (250 mg N L^{-1}) was supplemented into BG-11 media with OD$_{750}$ measurements taken daily over 14 days. Error bars represent the standard deviation from three biological replicates. High concentration of chitin led to particulate occlusion resulting in inaccurate OD measurements. High concentration of tyrosine was reduced to 25 mg N L^{-1} due to poor solubility. For low N concentration growth curves (1 mg N L^{-1}) see Appendix Figure A.10.

Table 4.2: **Summary of growth characteristics for *Synechococcus* sp. CCY9618 supplemented with nitrogen substrates at high concentration (250 mg N L^{-1}).** Note that the high concentration of tyrosine was reduced to 25 mg N L^{-1} due to poor solubility. ND - not detected due to spectrophotometer distortion. Missing values are due to lack of growth. * indicates a significant difference ($q < .05$) between species at high concentration. † indicates a significant difference ($q < .05$) within species between concentrations. Maximum growth rate and duration of lag phase was calculated using Growthcurver v3.0.1 [480]

***Synechococcus* sp. CCY 9618**		**High Concentration**	
Nitrogen Source	**Maximum OD**	**Maximum growth rate (μ_{max} hour^{-1})**	**Duration of lag phase (days)**
Alanine	0.0085 ± 0.0029	0.0163 ± 0.0030 *	5.28 ± 1.10
Arginine	0.0628 ± 0.0225	0.1709 ± 0.0685	6.45 ± 0.71
Asparagine	0.0292 ± 0.0003	0.1088 ± 0.0209	5.55 ± 0.14
Aspartate	0.1489 ± 0.0026	0.0353 ± 0.0040	7.24 ± 0.24 *
Cysteine	0.0028 ± 0.0006	-	-
Glutamate	0.1312 ± 0.0021	0.1029 ± 0.0386	4.70 ± 0.61
Glutamine	0.0935 ± 0.0229	0.0478 ± 0.0084	6.59 ± 1.26
Glycine	0.0526 ± 0.0172	0.0895 ± 0.0230	9.25 ± 1.00 *
Histidine	0.0949 ± 0.0087	0.2095 ± 0.1069	6.32 ± 0.35 *†
Isoleucine	0.1292 ± 0.0065	0.0606 ± 0.0091	8.72 ± 0.20
Leucine	0.0466 ± 0.0111	0.0549 ± 0.0104	9.39 ± 0.26
Lysine	0.0329 ± 0.0062	0.0844 ± 0.0263	8.96 ± 0.50
Methionine	0.0159 ± 0.0099	0.0506 ± 0.0144	8.63 ± 0.61
Phenylalanine	0.0186 ± 0.0044	0.0612 ± 0.0168	8.86 ± 0.21 *
Proline	0.1197 ± 0.0347	0.0545 ± 0.0096	7.75 ± 0.26 *
Serine	0.0256 ± 0.0159	0.0461 ± 0.0082	8.72 ± 0.58
Threonine	0.0025 ± 0.0022	-	-
Tryptophan	0.0865 ± 0.0093	0.0403 ± 0.0036	7.49 ± 0.40
Tyrosine	0.0042 ± 0.0012	-	-
Valine	0.0543 ± 0.0090	0.0781 ± 0.0213	9.72 ± 0.57 *
Chitin	ND	-	-
Glyphosate	0.0036 ± 0.0010	-	-
Nitrate	0.0605 ± 0.0210	0.0634 ± 0.0144	6.75 ± 0.92
Ammonium	0.0942 ± 0.0229	0.0714 ± 0.0127	5.54 ± 0.66 †
Urea	0.0301 ± 0.0044	0.0749 ± 0.0158	7.03 ± 0.45
No N	0.0014 ± 0.0002	-	-

Table 4.3: **Summary of growth characteristics for *Synechococcus* sp. CCY9618 supplemented with nitrogen substrates at low concentration (1 mg N L^{-1}).** Missing values are due to lack of growth. * indicates a significant difference ($q < .05$) between species at low concentration. † indicates a significant difference ($q < .05$) within species between concentrations. Maximum growth rate and duration of lag phase was calculated using Growthcurver v3.0.1 [480]

.

***Synechococcus* sp. CCY 9618**		**Low Concentration**	
Nitrogen Source	**Maximum OD**	**Maximum growth rate (μ_{max} hour^{-1})**	**Duration of lag phase (days)**
Alanine	0.0176 ± 0.0001	0.0529 ± 0.0071	4.75 ± 0.37
Arginine	0.0264 ± 0.0079	0.0228 ± 0.0025	6.79 ± 1.82
Asparagine	0.0224 ± 0.0045	0.0225 ± 0.0031	7.38 ± 1.05
Aspartate	0.0226 ± 0.0009	0.0309 ± 0.0042	6.10 ± 0.41
Cysteine	0.0015 ± 0.0010	-	-
Glutamate	0.0259 ± 0.0034	0.0201 ± 0.0034	8.38 ± 0.94
Glutamine	0.0265 ± 0.0037	0.0255 ± 0.0028	8.59 ± 0.73
Glycine	0.0246 ± 0.0017	0.0351 ± 0.0052	9.77 ± 0.07
Histidine	0.0254 ± 0.0036	0.0499 ± 0.0079	9.81 ± 0.31 †
Isoleucine	0.0171 ± 0.0059	0.0360 ± 0.0044	8.21 ± 0.70
Leucine	0.0150 ± 0.0004	0.0229 ± 0.0077	8.56 ± 0.18
Lysine	0.0200 ± 0.0079	0.0227 ± 0.0072	8.86 ± 1.28
Methionine	0.0275 ± 0.0007	0.0441 ± 0.0099	9.39 ± 0.04
Phenylalanine	0.0253 ± 0.0014	0.0449 ± 0.0081	9.34 ± 0.01
Proline	0.0313 ± 0.0022	0.0235 ± 0.0035	8.84 ± 0.06 *
Serine	0.0269 ± 0.0013	0.0320 ± 0.0050	7.84 ± 0.26
Threonine	0.0014 ± 0.0016	-	-
Tryptophan	0.0235 ± 0.0013	0.0279 ± 0.0064	8.68 ± 0.12 *
Tyrosine	0.0023 ± 0.0005	-	-
Valine	0.0241 ± 0.0033	0.0691 ± 0.0151	10.20 ± 0.19
Chitin	0.0034 ± 0.0009	-	-
Glyphosate	0.0018 ± 0.0005	-	-
Nitrate	0.0634 ± 0.0145	0.0486 ± 0.0081	8.45 ± 0.75
Ammonium	0.0361 ± 0.0067	0.0498 ± 0.0095	9.59 ± 0.07 *†
Urea	0.0364 ± 0.0030	0.0515 ± 0.0116	8.73 ± 0.71
No N	0.0015 ± 0.0008	-	-

patterns (Tables 4.3 and 4.5). Growth on NO_3^- resulted in the largest OD_{750} measurements (CCY9618: 0.0634 ± 0.0145; CCAP1479/10: 0.0750 ± 0.0143) while the largest OD_{750} from the non-traditional sources of N was obtained through proline as a N substrate for both strains (CCY9618: 0.0313 ± 0.0022; CCAP1479/10: 0.0407 ± 0.0014).

Growth rates are largely consistent for each substrate between species and N concentrations (Tables 4.2 - 4.5; Appendix Figures A.11 and A.12). No significant differences in growth rate are found for most substrates with two exceptions. *Synechococcus* sp. CCY9618 grew poorly when supplemented with a high concentration of alanine ($r = 0.0163 \pm 0.003$ hr^{-1}), whereas *Synechococcus* sp. CCAP1479/10 growth under the same conditions was significantly greater (r = 0.1235 ± 0.019 hr^{-1}; q = .045). Growth on low concentrations of alanine did not result in significant differences in growth rate between the species. Meanwhile, growth on glutamate in *Synechococcus* sp. CCAP1479/10 was the only substrate to result in a significantly greater growth rate at higher compared to lower concentrations (r = 0.1737 ± 0.0137 hr^{-1} v. r = 0.0257 ± 0.004 hr^{-1}; q = .011). Interestingly, though the maximum OD in *Synechococcus* sp. CCY9618 was obtained with aspartate, this was with a low growth rate (and long lag phase) (Table 4.2). Phenylalanine also seems to display differing extents of bioavailability between the two strains, however no significance difference is found between the growth rates. Furthermore, no significant differences in growth rate were identified within species between substrates of the same concentration.

The fastest growth rates for both picocyanobacteria strains were achieved with higher concentrations of N, especially among basic AAs. Histidine cultivated the greatest growth rates for both *Synechococcus* sp. CCY9618 (r = 0.2095 ± 0.1069 hr^{-1}) and *Synechococcus* sp. CCAP1479/10 (r = 0.2668 ± 0.5012 hr^{-1}), however the large error values limit the reliability of these growth rates. The greatest reliable growth rates utilised arginine as a N substrate for *Synechococcus* sp. CCY9618 (r = 0.1709 ± 0.0685 hr^{-1}) and aspartate for *Synechococcus* sp. CCAP1479/10 (r = 0.2011 ± 0.0484 hr^{-1}). At low concentrations, basic AAs continued to yield high growth rates, yet the fastest rates were achieved utilising valine for both strains (CCY9618: r = 0.0691 ± 0.0151 hr^{-1}; CCAP1479/10: r = 0.1335 ± 0.0809 hr^{-1}).

4.4.2 The Duration of Lag Phase Varies Significantly Between Picocyanobacteria Strains and Nitrogen Concentration

Lag phase measures the time taken for cells to adjust to their environmental conditions and enter an exponential period of growth. A shorter lag phase indicates a quicker translational response in producing the required assimilatory apparatus for nutrient uptake and metabolism. There are significant differences in the duration of lag phase between *Synechococcus* sp. CCY9618 and *Synechococcus* sp. CCAP1479/10 when grown under high and low concentrations of N, suggesting variety in species-specific response to nutrient influx. At high concentrations, *Synechococcus* sp. CCAP1479/10 predominantly achieved shorter lag phases than *Synechococcus* sp. CCY9618 (Tables 4.2 and 4.4; Appendix Table B.19), with growth on five AA substrates resulting in significantly reduced lag phases (aspartate (q = .0049), histidine (q = .014), valine (q = .031), phenylalanine (q = .0097), and proline

Table 4.4: **Summary of growth characteristics for *Synechococcus* sp. CCAP1479/10 supplemented with nitrogen substrates at high concentration (250 mg N L^{-1}).** Note that the high concentration of tyrosine was reduced to 25 mg N L^{-1} due to poor solubility. ND - not detected due to spectrophotometer distortion. Missing values are due to lack of growth. * indicates a significant difference ($q < .05$) between species at high concentration. † indicates a significant difference ($q < .05$) within species between concentrations. Maximum growth rate and duration of lag phase was calculated using Growthcurver v3.0.1 [480].

***Synechococcus* sp. CCAP 1479/10**	**High Concentration**		
Nitrogen Source	**Maximum OD**	**Maximum growth rate (μ_{max} hour^{-1})**	**Duration of lag phase (days)**
Alanine	0.1446 ± 0.0012	0.1235 ± 0.0190 *	4.19 ± 0.03
Arginine	0.0697 ± 0.0057	0.2180 ± 0.2556	3.38 ± 0.20
Asparagine	0.0271 ± 0.0005	0.0527 ± 0.0043	3.71 ± 0.46 †
Aspartate	0.1087 ± 0.0038	0.2011 ± 0.0484	2.16 ± 0.09 *†
Cysteine	0.0018 ± 0.0013	-	-
Glutamate	0.1227 ± 0.0006	0.1737 ± 0.0137 †	2.21 ± 0.12 †
Glutamine	0.1601 ± 0.0068	0.0484 ± 0.0183	2.38 ± 0.12 †
Glycine	0.0566 ± 0.0077	0.0646 ± 0.0125	10.06 ± 0.39 *
Histidine	0.0904 ± 0.0012	0.2668 ± 0.5012	3.31 ± 0.03 *†
Isoleucine	0.1005 ± 0.0017	0.0493 ± 0.0058	7.12 ± 0.34
Leucine	0.0746 ± 0.0050	0.0400 ± 0.0069	7.94 ± 0.48
Lysine	0.0171 ± 0.0031	0.1353 ± 0.0649	7.00 ± 0.01
Methionine	0.0034 ± 0.0002	-	-
Phenylalanine	0.0846 ± 0.0032	0.0561 ± 0.0102	6.56 ± 0.07 *†
Proline	0.2121 ± 0.0024	0.0788 ± 0.0063	4.44 ± 0.10 *†
Serine	0.0244 ± 0.0201	0.0489 ± 0.0088	8.34 ± 0.67
Threonine	0.0021 ± 0.0009	-	-
Tryptophan	0.1218 ± 0.0127	0.0547 ± 0.0070	6.25 ± 0.13 †
Tyrosine	0.0148 ± 0.0004	0.1072 ± 0.0242	3.28 ± 0.08
Valine	0.1000 ± 0.0017	0.0836 ± 0.0159	6.08 ± 0.05 *†
Chitin	ND	-	-
Glyphosate	0.0011 ± 0.0011	-	-
Nitrate	0.1238 ± 0.0226	0.0879 ± 0.0143	5.11 ± 0.08
Ammonium	0.0929 ± 0.0152	0.0907 ± 0.0133	3.03 ± 0.24 †
Urea	0.0490 ± 0.0161	0.0586 ± 0.0056	5.50 ± 0.35
No N	0.0015 ± 0.0009	-	-

Table 4.5: **Summary of growth characteristics for *Synechococcus* sp. CCAP1479/10 supplemented with nitrogen substrates at low concentration (1 mg N L^{-1}).** Missing values are due to lack of growth. * indicates a significant difference ($q < .05$) between species at low concentration. † indicates a significant difference ($q < .05$) within species between concentrations. Maximum growth rate and duration of lag phase was calculated using Growthcurver v3.0.1 [480].

***Synechococcus* sp. CCAP 1479/10**	**Low Concentration**		
Nitrogen Source	**Maximum OD**	**Maximum growth rate (μ_{max} hour^{-1})**	**Duration of lag phase (days)**
Alanine	0.0232 ± 0.0020	0.0344 ± 0.0059	4.90 ± 0.58
Arginine	0.0260 ± 0.0043	0.0354 ± 0.0053	4.90 ± 0.61
Asparagine	0.0233 ± 0.0012	0.0237 ± 0.0046	7.07 ± 0.19 †
Aspartate	0.0243 ± 0.0039	0.0421 ± 0.0055	5.90 ± 0.46 †
Cysteine	0.0025 ± 0.0005	-	-
Glutamate	0.0297 ± 0.0010	0.0257 ± 0.0040 †	8.45 ± 0.24 †
Glutamine	0.0270 ± 0.0015	0.0342 ± 0.0063	8.50 ± 0.47 †
Glycine	0.0330 ± 0.0031	0.0292 ± 0.0041	9.42 ± 0.51
Histidine	0.0266 ± 0.0075	0.0621 ± 0.0129	10.72 ± 0.75 †
Isoleucine	0.0214 ± 0.0051	0.0397 ± 0.0104	9.90 ± 0.57
Leucine	0.0226 ± 0.0034	0.0176 ± 0.0052	11.75 ± 1.30
Lysine	0.0166 ± 0.0017	0.0291 ± 0.0060	6.82 ± 0.50
Methionine	0.0033 ± 0.0003	-	-
Phenylalanine	0.0268 ± 0.0023	0.0365 ± 0.0059	10.09 ± 0.46 †
Proline	0.0407 ± 0.0014	0.0383 ± 0.0060	8.34 ± 0.07 *†
Serine	0.0224 ± 0.0037	0.0339 ± 0.0066	7.18 ± 0.42
Threonine	0.0045 ± 0.0044	-	-
Tryptophan	0.0293 ± 0.0003	0.0252 ± 0.0069	10.13 ± 0.09 *†
Tyrosine	0.0034 ± 0.0010	-	-
Valine	0.0242 ± 0.0051	0.1335 ± 0.0809	10.04 ± 0.08 †
Chitin	0.0026 ± 0.0009	-	-
Glyphosate	0.0011 ± 0.0008	-	-
Nitrate	0.0750 ± 0.0143	0.0570 ± 0.0109	8.34 ± 0.75
Ammonium	0.0406 ± 0.0065	0.0539 ± 0.0160	8.26 ± 0.13 *†
Urea	0.0419 ± 0.0054	0.0561 ± 0.0212	7.79 ± 0.92
No N	0.0019 ± 0.0009	-	-

($q = .0073$)). On the other hand, growth on glycine resulted in a significantly reduced lag phase for *Synechococcus* sp. CCY9618 ($q = .022$). At low N concentrations, significant differences in the duration of lag phase were less prevalent. Among AAs, only growth on proline resulted in a shorter lag phase at low concentrations for *Synechococcus* sp. CCAP1479/10 ($q = .045$), while growth on tryptophan resulted in a significantly reduced lag phase for *Synechococcus* sp. CCY9618 ($q = .0097$). Interestingly, growth on NH_4^+ led to a reduced lag phase in *Synechococcus* sp. CCAP1479/10 ($q = .012$) - the only non-AA substrate to significantly differ in growth rate or duration of lag phase between the two species, further demonstrating species-specific variation in response to common N sources.

The lag phase response between substrate concentrations in the same species presents differing results.

Synechococcus sp. CCAP1479/10 was able to adapt in a shorter time for 40% of N substrates studied when exposed to higher concentrations (Tables 4.2 and 4.4). These mostly include charged (aspartate (q = .018), glutamate (q = .0055), histidine (q = .011)) and non-polar neutral (phenylalanine (q = .019), proline (q = .0032), tryptophan (q = .0097), valine (q = .0065)) AAs, also including NH_4^+ (q = .019). On the other hand, the duration of lag phase in *Synechococcus* sp. CCY9618 is generally consistent between high and low N concentrations (Tables 4.3 and 4.5). Growth on histidine (q = .021) and NH_4^+ (q = .034) were the only substrates which resulted in significantly reduced lag phases at higher concentrations in both species.

When grown under a high N load, the shortest lag phases were found in *Synechococcus* sp. CCAP1479/10, on substrates which can be immediately incorporated into N metabolic pathways - glutamate (2.21 ± 0.12 days) and glutamine (2.38 ± 0.12 days). Growth on aspartate also contributed to short lag phases (2.16 ± 0.09 days), suggesting acidic AAs may require the minimum adaptation time.

4.4.3 Proteomic Response to Growth on Amino Acids

We carried out TMT quantitative proteomic analysis on *Synechococcus* sp. CCAP1479/10 with a subset of AA-substrates to understand the impact that different AA characteristics have on the proteome and elucidate the molecular basis of differential AA assimilation. The proteomic response to growth on arginine, asparagine, glutamate, and proline was compared to picocyanobacteria growth on NO_3^- and N-starvation, with substrates chosen based on encoded AATs and their preferred substrates in *Synechococcus* sp. CCAP1479/10 (see section 4.3.3.1). TMT proteomics resulted in the identification of 5,720 unique peptides and 1,167 proteins. Of these, only proteins which were detected in all three biological replicates were selected for further analysis, resulting in a total of 5,134 peptides and 836 proteins (Figure 4.3A and Appendix Tables B.20 and B.21). Among these, 92% were identified with at least two unique peptides and 80% had sequence coverage greater than 10%. FDR-adjusted ANOVA and Tukey Test analyses identified 223 differentially expressed proteins (DEPs) in total (Figure 4.3A). The 836 proteins detected in triplicate in this study corresponds to 24.3% of the predicted 3,441 proteins encoded in the *Synechococcus* sp. CCAP1479/10 genome (Table 2.1). The range of represented proteomes detected in other cyanobacterial studies varies considerably, from 18% [471] to 90% [482], though the majority of studies identify between 35 and 70% of the total predicted proteins [472, 483–487].

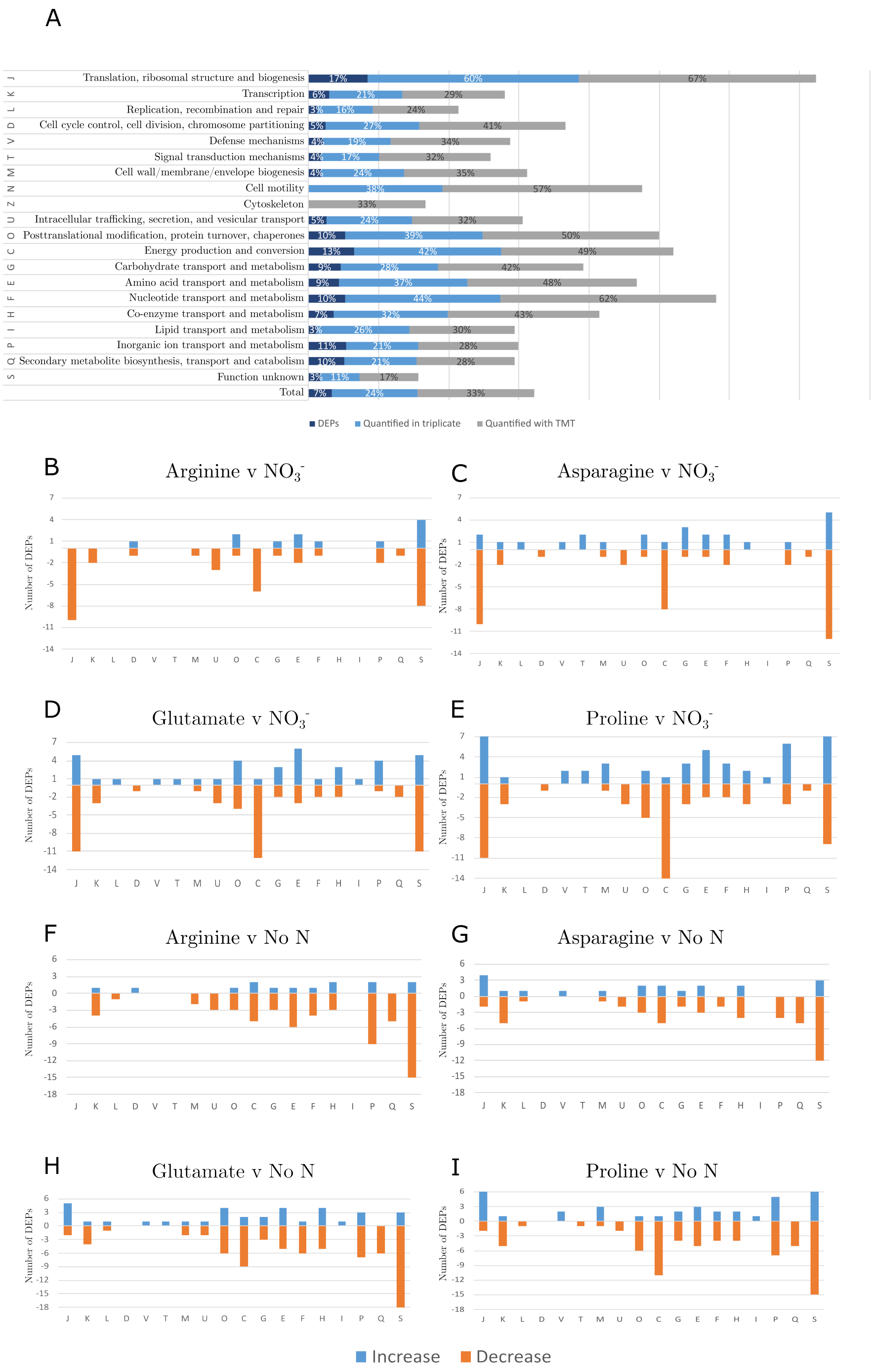

Figure 4.3: **Functional annotation of *Synechococcus* sp. CCAP1479/10 DEPs when grown on various amino acids compared to NO_3^- and nitrogen-starvation.** A) Breakdown of TMT-quantified peptides in *Synechococcus* sp. CCAP1479/10 as a percentage of predicted number of total proteins for each COG category. Proteins identified with TMT proteomics (grey); proteins identified in triplicate (light blue); proteins identified as DEPs (dark blue). For total number of proteins for each category, see Table 2.2). B - E) Distribution of DEPs for sole AA-N growth compared to NO_3^-; F - I) Distribution of DEPs for sole AA-N growth compared to N starvation. Blue columns indicate up-regulated proteins, orange columns indicate down-regulated proteins, when grown on the specified AA.

Thus, the number of detected proteins here is roughly consistent with other studies, albeit towards the bottom of the expected range.

The number of DEPs varied considerably between conditions. Of the total 223 DEPs identified, 160 DEPs were present in *Synechococcus* sp. CCAP1479/10 grown across the four AA N-substrate conditions compared to NO_3^- (Appendix Tables B.24 - B.27) while 172 DEPs were identified compared to N-starvation (Appendix Tables B.28 - B.31). However, the number of DEPs between AA N-substrates compared to NO_3^- and N-starvation noticeably differed. Growth on arginine and asparagine consistently yielded fewer DEPs to growth on glutamate and proline (51 and 69 v. 103 and 112, respectively (NO_3^-); 83 and 90 v. 122 and 116, respectively (N-starvation)). Meanwhile, within each condition, the number of down-regulated DEPs was greater than the number of up-regulated DEPs, with the greatest differences found when comparing AA N-substrate growth to N-starvation. Of particular interest is the overlap of DEPs between AA substrate conditions. As would be expected, the majority of down-regulated DEPs for *Synechococcus* sp. CCAP1479/10 growth on AA N-substrates compared to NO_3^- displayed considerable overlap, with 79 proteins down-regulated among all four substrate-conditions (Figure 4.4B). On the other hand, the overlap of up-regulated DEPs reveals differences in the cellular response to growth on different substrates (Figure 4.4A). Only four up-regulated DEPs are shared between the four conditions, while approximately half of all DEPs identified under proline- (42%) and glutamate- (53%) growth conditions are only over-expressed in that singular condition. Similar patterns are shown when compared to N-starvation (Appendix Figure A.13). This demonstrates that the molecular response varies to growth on specific AA N-substrates, exposing possibilities for niche adaptation to environmental AA compositions. In turn, this may influence the composition of the cyanobacteria (and wider) community based on encoded N assimilatory mechanisms.

Comparing the *Synechococcus* sp. CCAP1479/10 proteomic response between the four AA N-substrates reveals further differences. In total, 80 unique DEPs were identified within AA N-substrate conditions (Table 4.6 and Appendix Tables B.32 - B.37). The two conditions which resulted in the greatest proteomic differences are proline and arginine with 35 and 13 up-regulated proteins respectively. Meanwhile, growth on arginine and asparagine yielded only seven DEPs (Table 4.6).

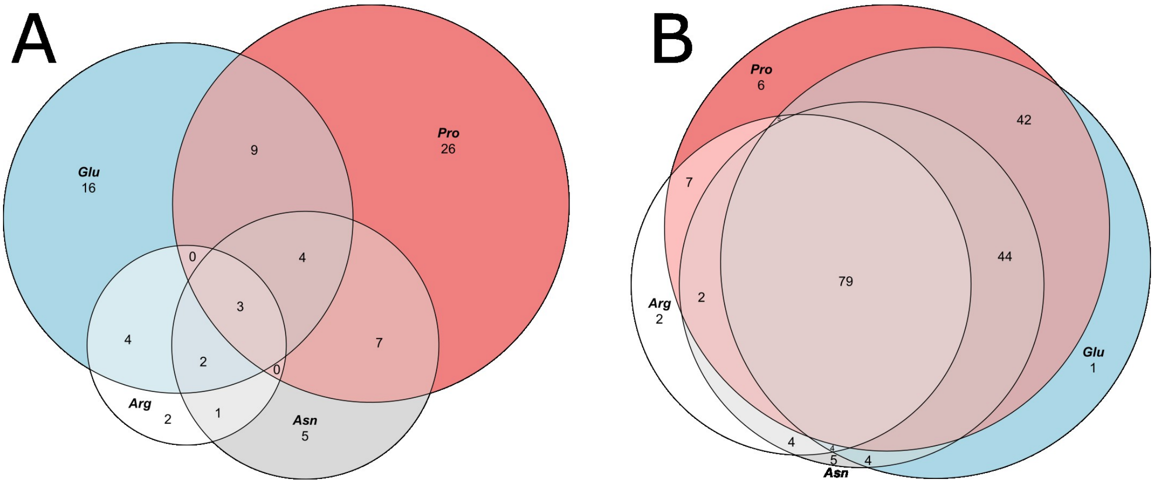

Figure 4.4: **Area-proportional Euler diagrams for amino acid nitrogen-substrate DEP overlaps compared to NO_3^-.** A) Overlap of proteins up-regulated in *Synechococcus* sp. CCAP149/10 when grown under four AA conditions; B) Overlap of proteins down-regulated in *Synechococcus* sp. CCAP149/10 when grown under four AA conditions. For overlap of DEPs compared to N starvation, see Figure A.13.

Table 4.6: **Comparisons of amino acid nitrogen-substrate DEPs**. The number of DEPs up-regulated in each AA N-substrate condition compared to each other AA growth condition is displayed. In each comparison, the left value indicates up-regulated DEP number under the substrate to the left, while the right value indicates DEP number under the substrate above.

	Arginine	Asparagine	Glutamate	Proline
Arginine		4 / 3	9 / 15	13 / 35
Asparagine	3 / 4		7 / 9	10 / 15
Glutamate	15 / 9	9 / 7		19 / 21
Proline	35 / 13	15 / 10	21 / 19	

The physiological processes associated with the greatest number of DEPs were determined with COG functional annotation. Proteins associated with 'AA metabolism and transport', 'translation', and 'inorganic ion transport and metabolism' were the most common up-regulated DEPs across all four AA N-substrate growth conditions compared to NO_3^- (Figure 4.3B - E). Meanwhile, proteins involved with 'energy production and conversion' and 'translation' contributed to over half of down-regulated DEPs of known function compared to NO_3^- (Figure 4.3B - E). When comparing AA-grown *Synechococcus* sp. CCAP1479/10 to a N-starvation state, similar categorisation of up-regulated DEPs are found (Figure 4.3F - I) while the primary categories for down-regulated DEPs are 'energy production and conversion', 'inorganic ion transport and metabolism', and 'transcription'. Between AA substrates, the proportions of DEP functional categorisation is mostly consistent (both compared to NO_3^- and N-starvation), though no DEPs associated with 'translation' are identified as up-regulated when grown on arginine, with the largest percentage of DEPs instead associated with 'energy production and conversion' (N-starvation: 20%) and 'post translational modification and protein turnover' (NO_3^-: 15%).

4.4.3.1 Pathway-Enrichment Analysis

KEGG pathway-enrichment analysis identified multiple pathways with differential expression between growth on AA N-substrates and both NO_3^- and N-starvation. Growth on AAs compared to NO_3^- resulted in the over-expression of 21 KEGG pathways in total, however only five pathways are up-regulated under two or more AA-substrate growth conditions (including 'RNA polymerase' and 'aminoacyl t-RNA biosynthesis'), indicating a large degree of variation in nutrient response (Figure 4.5A - D). The four AA growth conditions display varying degrees of pathway enrichment, with arginine only significantly up-regulated in one pathway ('cytoskeleton proteins') whereas growth on proline resulted in the significant up-regulation of 11 pathways (the most relevant including 'lysine biosynthesis', 'arginine biosynthesis', and 'transporters'). Pathways involved with AA metabolism and transporters were expected to be up-regulated in AA-grown *Synechococcus* sp. CCAP1479/10 compared to growth on NO_3^-, yet this was found only under glutamate and proline growth conditions. As described above, growth on arginine yielded a single enriched pathway, while the eight pathways enriched under asparagine growth included nucleotide biosynthesis and metabolism among others (Figure 4.5B).

Down-regulated pathways on growth under AA N-substrates compared to NO_3^- were more consistent between conditions as expected (Figure 4.5A - D). In total, 19 down-regulated pathways were identified among the four AA conditions, with 12 of these shared among at least two substrates. Four pathways were depleted in all four conditions compared to growth on NO_3^-, including 'N metabolism', 'ribosome', and 'photosynthesis proteins'.

KEGG pathway enrichment between AA-grown *Synechococcus* sp. CCAP1479/10 and N-starvation displays similar patterns compared to growth on NO_3^- (Appendix Figure A.14A - D). Of the 23 KEGG pathways enriched under AA N-substrate conditions compared to N-starvation, seven of these are shared between conditions with only the 'RNA polymerase' pathway up-regulated under all four AA substrates. Compared with NO_3^-, the enrichment patterns of AA-associated pathways are similar (found in glutamate and proline, not enriched in arginine and asparagine). The majority of pathways up-

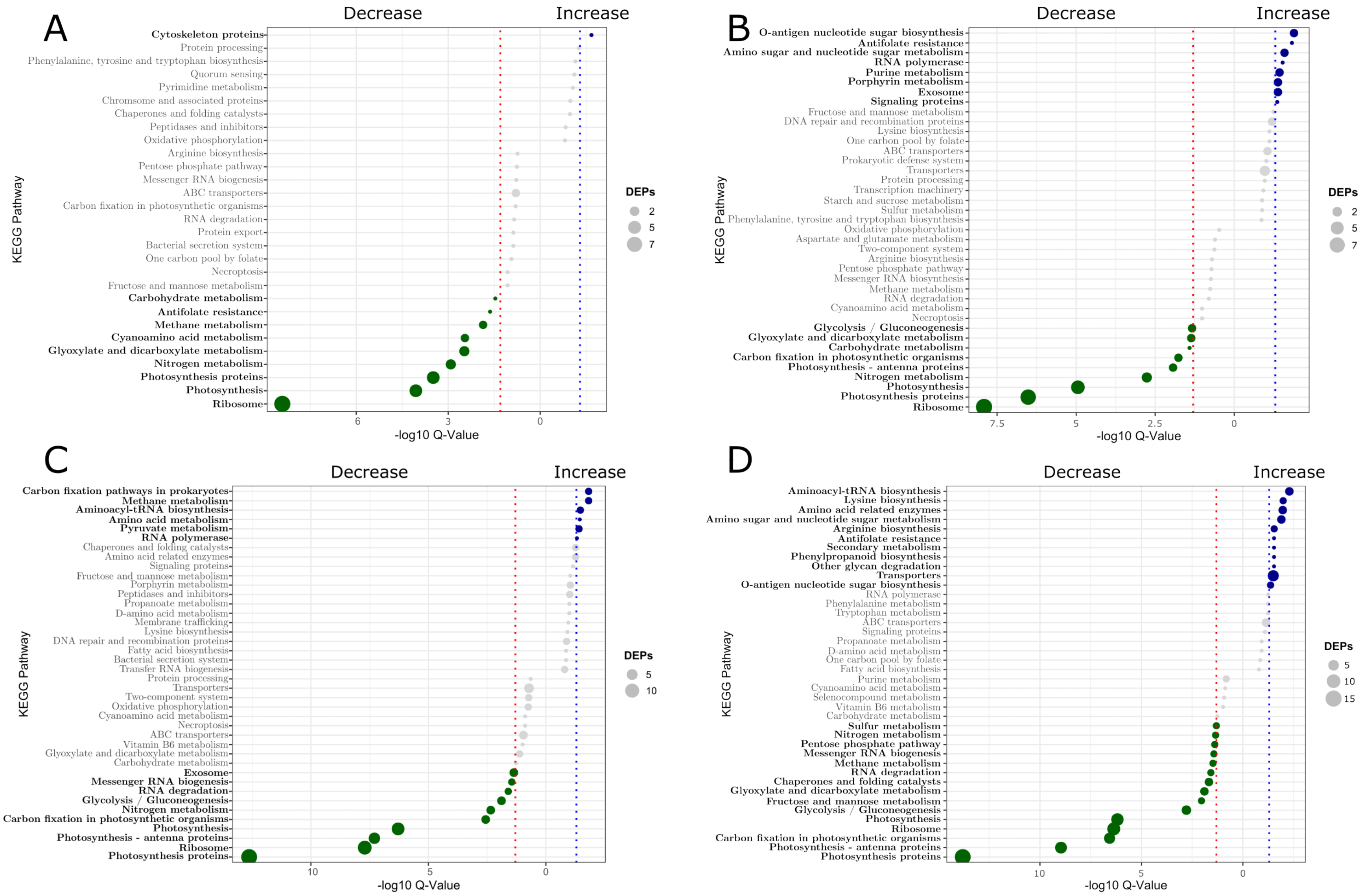

Figure 4.5: **KEGG pathway enrichment analysis of growth on amino acid nitrogen-substrate v. NO_3^-.** A) Arginine v. NO_3^-; B) Asparagine v. NO_3^-; C) Glutamate v. NO_3^-; D) Proline v. NO_3^-. Top 20 enriched pathways are shown with non-significant pathways in grey, significant pathways are in bold. Blue indicates up-regulated pathways while green indicates down-regulated pathways compared to NO_3^-.

regulated under arginine and asparagine growth are associated with photosynthesis (Appendix Figure A.14A - B) while enriched pathways under glutamate and proline include amino acid biosynthesis and metabolism (Appendix Figure A.14C - D).

4.4.3.2 Nitrogen Assimilation and Amino Acid Associated DEPs

Several N-associated proteins investigated in Chapter 3 displayed differential expression in AA grown *Synechococcus* sp. CCAP1479/10 compared to NO_3^- (Table 4.7). AA metabolism / biosynthesis proteins involved with lysine (DapB, DapF, DapL), arginine (ArgJ, ArgB), and asparagine (GatC) were up-regulated in at least one AA-grown condition. All except DapB were up-regulated under proline growth, with ArgJ and DapF up-regulated under both the proline and glutamate conditions. Conversely, the only N-assimilation associated protein up-regulated under asparagine growth is DapB, catalysing an earlier step in lysine biosynthesis than DapF. Growth on arginine did not result in the up-regulation of any proteins associated with N assimilation, consistent with the lack of pathway enrichment. Of further interest is the abundance of Amt1, the primary cyanobacterial NH_4^+ transporter, under proline growth, perhaps indicating external proline degradation by extracellular/periplasmic oxidase and subsequent deaminase activity to form NH_4^+. Further DEPs categorised as involved with AA metabolism include AroG (involved in the shikimate pathway of aromatic AA biosynthesis), up-regulated across all four conditions.

Down-regulated DEPs provide further insights into the molecular basis of picocyanobacterial growth on AAs (Table 4.7). Six proteins associated with N assimilation are less abundant in AA-grown *Synechococcus* sp. CCAP1479/10 than when grown utilising NO_3^-. These include the substrate-binding subunits of the urea and NO_3^-/NO_2^- ABC transporter complexes, UrtA and NrtA respectively. These two subunits are significantly less abundant among all four conditions of AA-based growth compared to NO_3^-, however the remaining detected subunits of these transporters (UrtD, UrtE; NrtC, NrtD) are not differentially expressed between conditions. Additionally, glutamine synthetase (GlnN) is down-regulated in *Synechococcus* sp. CCAP1479/10 when grown on arginine, asparagine, and glutamate. The periplasmic substrate-binding component of the AAT N-III (NatI) is the only differentially expressed AAT subunit, down-regulated in *Synechococcus* sp. CCAP1479/10 grown on glutamate compared to NO_3^-.

Table 4.7: **Amino acid-grown *Synechococcus* sp. CCAP1479/10 up-regulated DEPs associated with nitrogen assimilation and amino acid biosynthesis compared to NO_3^-.** Tick indicates DEP in that AA N-substrate condition. R: arginine; N: asparagine; E: glutamate; P: proline.

DEP	Protein	N Source or AA Biosynthesis Pathway	R	N	E	P
Increase	DapB	Lysine		✓		
Increase	DapF	Lysine			✓	✓
Increase	DapL	Lysine				✓
Increase	ArgB	Arginine				✓
Increase	ArgJ	Arginine			✓	✓
Increase	Amt1	Ammonium				✓
Increase	GatC	Asparagine				✓
Increase	AroG	Aromatic AAs	✓	✓	✓	✓
Decrease	GlyA	Serine/Glycine	✓			
Decrease	GlnN	Nitrate	✓	✓	✓	
Decrease	UrtA	Urea	✓	✓	✓	✓
Decrease	NrtA	Nitrate	✓	✓	✓	✓
Decrease	ThrC	Threonine			✓	✓
Decrease	NatI	Amino Acids			✓	

The proteomic response of *Synechococcus* sp. CCAP1479/10 when comparing growth on AA N-substrates to each other reveals substantial up-regulation in AA biosynthesis and N transporter proteins when grown under proline (Appendix Tables B.34, B.36, and B.37). The NH_4^+ transporter Amt1 is up-regulated with proline as N-substrate compared to both glutamate and arginine, while the NO_3^-/NO_2^- transporter ATP-binding subunit (NrtC) is up-regulated under proline growth compared to arginine. Meanwhile, among AA biosynthesis proteins, ArgB (arginine) and GatC (Asparagine) are up-regulated in proline compared to the other three AA N-substrate conditions while also up-regulated are DapF (enriched in proline v. arginine) and DapL (enriched in proline v. arginine and asparagine), both involved in lysine biosynthesis. On the other hand, the only AA-associated up-regulated DEP when grown on glutamate is DapF (compared to asparagine growth), while asparagine as a N-source resulted in the up-regulation of ThrC (threonine biosynthesis) compared to arginine. These results indicate that growth on proline resulted in the most significant alterations to AA-biosynthesis. This may potentially represent a favoured source of N, consistent with growth patterns summarised earlier (section 4.4.1), or be due to extracellular degradation into NH_4^+ for Amt1 uptake and immediate GS/GOGAT incorporation.

Table 4.8: **Amino acid-grown *Synechococcus* sp. CCAP1479/10 up-regulated DEPs associated with non-nitrogen transporters compared to NO_3^-.** Tick indicates DEP in that AA N-substrate condition. Uncharacterised proteins were identified using BLASTP with highest homology based on *Synechococcus* sp. FACHB-909. R: arginine; N: asparagine; E: glutamate; P: proline.

Protein	**Name**	**R**	**N**	**E**	**P**
SbpA	Sulfate-binding protein		✓		✓
CmpA	Bicarbonate-binding protein				✓
CmpC	Bicarbonate transport ATP-binding protein			✓	
Ga0436389_004_46165_47994	ABC transporter ATP-binding protein		✓	✓	✓
Ga0436389_026_116581_117309	LPS export ABC transporter ATP-binding protein		✓		✓
Ga0436389_030_38087_39844	ABC-F family ATP-binding cassette domain-containing protein				✓

4.4.3.3 Transporters

Differential expression of transporters can assist in elucidating the cellular response to AA assimilation. As previously stated, only NatI of the N-III system was identified as an AAT DEP, down-regulated in glutamate-growth compared to NO_3^-, in addition to Amt1 and NrtC. However, multiple other transporters, characterised and novel, display increased abundance in under AA N-substrates (Table 4.8). Among the non-N associated differentially expressed transporters, subunits of two systems are found. These include the substrate-binding protein of the sulphate ABC transporter (SbpA) (asparagine and proline v. NO_3^-) and subunits of the high-affinity bicarbonate ABC transporter - CmpC (ATPase; glutamate v. NO_3^-) and CmpA (substrate-binding protein; proline v. NO_3^-). Uncharacterised proteins associated with ABC transporters were also identified with Ga0436389_004_46165_47994 of greatest interest. This protein is up-regulated under asparagine, glutamate, and proline growth compared to NO_3^- and is an unknown substrate ABC transporter ATP-binding protein. BLAST analysis reveals an MdlB domain superfamily involved in multidrug transport, primarily efflux, of small hydrophobic molecules [488]. This may suggest associations to hydrophobic AA export due to build-up of intracellular AAs, though proline is the only hydrophobic AA growth condition tested.

4.4.3.4 Other Proteins

Expression of proteins involved in multiple physiological processes were up-regulated between AA-grown and NO_3^--grown *Synechococcus* sp. CCAP1479/10, including those involved with translation, photosynthesis, and the stress response (Table 4.9). Translation-associated DEPs were identified in the asparagine, glutamate, and proline conditions - noticeably absent when grown on arginine. These proteins include tRNA ligases in addition to several core components of the 50S ribosomal subunit (Table 4.9). However, while the differential expression of tRNA ligases are limited to up-regulation, a substantial number of both 30S and 50S ribosomal proteins are down-regulated on AA substrates (Appendix Tables B.24 - B.27). This pattern is also found with DEPs associated with photosynthesis. While proteins involved with pigment biosynthesis (AcsF and CpcF) are up-regulated under asparagine, glutamate, and proline conditions, protein subunits of the PSI and PSII complexes are consistently down-regulated in AA N-substrate conditions compared to NO_3^-. Of note is the identification of PsbD in *Synechococcus* sp. CCAP1479/10; encoded by *psbD*, this gene was unidentified in the comparative genomics in Chapter 2, indicating multiple copies of this gene with paralogous sequences. Furthermore, FtsH1, linked with nutrient stress response in cyanobacteria, was up-regulated in the arginine, asparagine, and glutamate conditions, but not when grown on proline. Notably however, FtsH1 was also up-regulated in the same AA conditions when compared to N-starvation (Appendix Tables B.28 - B.31).

DEPs involved with nucleotide biosynthesis and carbohydrate metabolism were also identified between growth on AA N-substrates and NO_3^- (Table 4.9). Two proteins involved with nucleotide biosynthesis were differentially expressed with a notable pattern. CTP synthase (PyrG) catalyses the last committed step in pyrimidine biosynthesis - the amination of UTP to CTP utilising glutamine as a N source [489]. This is up-regulated under arginine and glutamate growth but not differentially expressed when grown on asparagine or proline. Conversely, purine biosynthesis proteins are up-regulated under asparagine and proline growth (compared to NO_3^-), with the abundance of PurH catalysing the last two steps of purine formation. Further differences are identified with carbohydrate metabolism, including the up-regulation of the glycogen catabolising GlyP only under proline growth (Table 4.9).

4.5 Discussion

The dominance of freshwater picocyanobacteria in aquatic environments is associated with rapid nutrient uptake due to their reduced cell size, however genetic contributions to this dominance are often

Table 4.9: **Selected DEPs in *Synechococcus* sp. CCAP1479/10 grown on amino acid nitrogen-substrate compared to NO_3^-**. Tick indicates DEP in that AA N-substrate condition. R: arginine; N: asparagine; E: glutamate; P: proline.

Function	**Protein**	**Name**	**R**	**N**	**E**	P
Translation	GltX	Glutamate tRNA ligase		✓		✓
Translation	AlaS	Alanine tRNA ligase			✓	✓
Translation	LysS	Lysine tRNA ligase			✓	
Translation	Rpl2	50S ribosomal protein L2				✓
Translation	Rpl6	50S ribosomal protein L6				✓
Translation	RplQ	50S ribosomal protein L17				✓
Translation	RplR	50S ribosomal protein L18		✓	✓	
Translation	RpsQ	30S ribosomal protein S17				✓
Stress Response	FtsH1	ATP-dependent metalloprotease FtsH1	✓	✓	✓	
Nucleotide Biosynthesis	PyrG	CTP synthase	✓		✓	
Nucleotide Biosynthesis	PurH	Bifunctional purine biosynthesis protein purH		✓		✓
Respiration	NdhF1	Proton-translocating NADH-quinone dehydrogenase subunit F1 NdhF1	✓		✓	
Photosynthesis	AcsF	Magnesium-protoporphyrin IX monomethyl ester [oxidative] cyclase		✓	✓	
Photosynthesis	CpcF	Phycocyanin alpha phycocyanobilin lyase CpcF			✓	✓
Transcription	RpoC1	DNA-directed RNA polymerase subunit beta’		✓	✓	✓
Carbohydrate Metabolism	ManC	Mannose-1-phosphate guanylyltransferase		✓		✓
Carbohydrate Metabolism	PpsA	Phosphoenolpyruvate synthase			✓	
Carbohydrate Metabolism	GlyP	Alpha-1,4 glucan phosphorylase				✓
Cell Cycle	MinD	Septum site-determining protein	✓			

overlooked [16]. The encoded N assimilatory mechanisms of freshwater picocyanobacteria propose metabolic variability, especially among organic sources of N such as AAs (Chapter 3). The cyanobacterial AAT diversity, featuring varied preferred substrates and uptake rates, suggest differential growth responses to different AAs. Here, we conducted a broad study to determine AA bioavailability as a N source in freshwater picocyanobacteria and identify differences between two strains. We subsequently investigated the proteomic response of *Synechococcus* sp. CCAP1479/10 to growth on a variety of AA N-sources in comparison to growth on NO_3^-.

4.5.1 The Diversity of Bioavailable Amino Acids

4.5.1.1 Preferred Substrates for Amino Acid Transporters Are Not Specifically Favoured

Our growth assays indicate that the majority of AAs can act as a sole N source for freshwater picocyanobacteria. Previous research demonstrates mixed results of the utilisation of AAs as N sources (Table 1.1). *Synechococcus elongatus* PCC 6301, a commonly used model organism, has only been successfully grown on glutamine [165], though *Synechococcus* sp. PCC 7002, a marine β-*Synechococcus* can utilise a range of neutral and acidic AAs [164]. Heterocystous cyanobacteria also display a variety of capabilities, with *Pseudanabaena* spp. exhibiting growth on acidic AAs [165] and *Anabaena* sp. PCC 7122 utilising neutral AAs [166]. Interestingly, only *Spirulina platensis* displays similar uptake capabilities to those found in this study (Table 1.1). The widespread ability of the tested picocyanobacteria to grow on AAs as their sole N source is significant, and presents a broad range of N substrates which may not be accessed by other cyanobacteria (Table 1.1). Based on encoded AATs in *Synechococcus* sp. CCY9618 and *Synechococcus* sp. CCAP1479/10, the preferred substrates for AA assimilation were expected to be aspartate (N-II), glycine (N-III), and glutamate (GltS). High growth rates and yields were identified under glutamate and aspartate growth, though glycine was not an optimal N source with an increased lag phase and reduced yield. This may be linked to the role of these molecules in the wider context of N metabolism. The ease of cellular incorporation of these molecules differs, yet each has short catabolic pathways to yield other AAs, NH_4^+, or other metabolic intermediates. Glutamate is involved in many biosynthetic processes including its role as the direct and indirect amine donor for all AAs [188]. Meanwhile, aspartate is the precursor molecule for lysine, threonine, and isoleucine biosynthesis; it can also be rapidly transaminated by aspartate aminotransferase (AspC) to yield glutamate or deaminated (aspartate-ammonia lyase (AspA)) to NH_4^+ [268]. On the other hand, the simple molecular structure of glycine results in two primary catabolic mechanisms; the oxidative deamination of glycine yields NH_4^+ and glyoxylate, while the glycine cleavage system catalyses decarboxylation to produce NH_4^+ and one-C units of 5,10-methylenetetrahydrofolate. As such, whereas aspartate and glutamate can directly inter-convert to other AAs while also catabolising to NH_4^+, glycine catabolism only yields NH_4^+, potentially limiting growth on this substrate due to indirect AA biosynthesis compared to glutamate and aspartate.

Furthermore, picocyanobacterial cells may be unable to effectively process elevated amounts of glycine. High concentrations of intracellular glycine is associated with a toxicity response, reducing Mg^{2+} ions and subsequently reducing the available Mg pool for essential cellular mechanisms (including photosynthesis) [490]. The high concentration utilised in this study is accordant with growth inhibition seen elsewhere, and while lower concentrations of glycine weakened this toxicity, the long lag phase was retained which is consistent with other bacteria [490, 491]. Therefore, the advantages of N-III preferential glycine uptake remain unclear, with a function for preferential uptake of leaked AAs unlikely due to the low permeability of glycine limiting its extracellular leakage [492].

4.5.1.2 Non-Bioavailable Substrates

The mechanisms behind the inability to utilise cysteine and threonine as sole N-substrates underlines the limited understanding of cyanobacterial AA degradation. Though cyanobacterial growth on cysteine and threonine has not been previously investigated, analyses of other bacteria suggest mixed responses to both cysteine and threonine uptake [291]. Cysteine cytotoxicity is well-documented, even at low concentrations, resulting in enzyme inhibition and oxidative stress, necessitating strict cellular control systems [493–495]. In *E.coli*, the primary method of cysteine control is via the master regulator CysB regulating cysteine biosynthesis, a protein which may be homologous to the cyanobacterial repressor NdhR [496, 497]. However, cysteine degradation is less understood. Cysteine can be catabolised through cysteine desulfurase, yielding alanine and a thiol, while cysteine desulfhydrase activity (to yield pyruvate, NH_4^+, and sulfide) is not known to be present in cyanobacteria [495]. Therefore, the inability to utilise cysteine suggests insufficient catabolic pathways are present to effectively mitigate elevated cysteine toxicity. Meanwhile, threonine was also unable to be assimilated as a sole N source.

Accumulation of threonine is not associated with toxicity and two major catabolic pathways (via propanoate and glycine) are known. Though potentially homologous proteins involved with the degradation of threonine are found in cyanobacteria (data not shown), the efficacy of this pathway in cyanobacteria is not well-documented. Furthermore, the ability to uptake threonine has been demonstrated in *Anabaena* sp. PCC 7120, suggesting the cause of growth inhibition is intracellular [256]. Further research is required to elucidate the impact of threonine on cyanobacteria growth.

4.5.1.3 Unexpected Bioavailability of Basic Amino Acids

Assimilation capabilities of basic AAs reveal the complexity of picocyanobacteria AA uptake. Freshwater picocyanobacteria lack a basic-specific AAT, however the ability to utilise arginine and lysine as N sources indicates uptake capabilities for these AAs. Although basic, import of histidine has been demonstrated from the N-II transporter due to the low pKa of its side chain, hence acting as a neutral substrate in physiological pH [254]. Meanwhile, import of arginine and lysine from N-II and N-III have not been demonstrated in deletion studies [254, 255]. However, all previous AAT characterisation research has been carried out in *Anabaena* and *Synechocystis*, taxa which encode the Bgt basic transporter and have been shown to effectively utilise arginine as a N source [165]. There are three possible mechanisms for basic AA uptake without a dedicated transporter. The first is a broader specificity for the charged N-II AAT than expected for *Synechococcus*. Lysine and arginine may be able to bind to the N-II substrate-binding domain, albeit with low affinity. However, the excess of AAs in this study may have overcome this and allow active uptake catalysed by BgtA (N-II ATPase component shared with the Bgt transporter). Research has shown a limited reduction (up to 33%) in basic AA uptake in *Anabaena* sp. PCC 7120 ΔN-II mutants while basic AA uptake of ΔBgt mutants is reduced by 95% [254]. While this pathway may allow limited uptake of basic AAs, the high growth rates achieved by *Synechococcus* spp. on these substrates at both high and low concentrations would suggest greater uptake of AA is required. A second potential cause of basic AA assimilation is extracellular degradation preceding transport of NH_4^+ or NO_3^- into the cell. Lastly, a novel transporter may be responsible for basic AA uptake yet to be identified. Firm conclusions could not be drawn from the proteomic response to arginine growth compared to NO_3^-. This resulted in few DEPs, potentially supporting the pathway of extracellular degradation or conversion of arginine to NO_3^- via nitric oxide, a pathway previously identified in *Synechococcus* sp. PCC7335 [267].

A further complication for basic AA degradation is the lack of understanding over their catabolic pathways in *Synechococcus*. Histidine catabolism to glutamate and formate/formamide is carried out by the highly conserved Hut pathway, however these enzymes are absent in cyanobacteria with alternative pathways not identified [295]. Arginine catabolic enzymes are encoded by the majority of cyanobacteria yet no complete pathways have been identified in freshwater *Synechococcus elongatus* [266]. However, an arginine oxidase (AoxA) is found in freshwater *Synechococcus* spp. (Appendix Table B.12), potentially revealing a pathway to NH_4^+ via the oxidative deamination of arginine to 2-ketoarginine [266]. Additionally, potential homologues for the lysine cadaverine pathway are identified in cyanobacteria (data not shown). To understand the mechanisms behind basic AA growth of Bgt-lacking freshwater *Synechococcus*, a greater understanding of uptake and catabolism is required.

4.5.1.4 Species-Specific Variability

Species-specific variability can have a large impact on the composition of a cyanobacterial community. Although the two freshwater picocyanobacteria utilised in this study are closely-related, lag times for *Synechococcus* sp. CCAP1479/10 growth on several N-substrates was significantly reduced, mostly at high concentrations. Lag phase is primarily preparative, with cells adapting to stress and assembling components for assimilation and division [498]. The lag phase heterogeneity observed indicates the reduced adaptation period required for *Synechococcus* sp. CCAP1479/10 for a number of substrates compared to *Synechococcus* sp. CCY9618. This could be linked to the evolutionary history of the two strains. With *Synechococcus* sp. CCAP1479/10 isolated from the mesotrophic Windermere [499], the response to N influx may be different to that of *Synechococcus* sp. CCY9618 which was isolated from a waterbody in Leiden (Chapter 2). Though nutrient data for this waterbody is not known, its proximity to urban features suggest that this pond is eutrophic. This may contribute towards the diversity found in lag time duration and other differential bioavailable AAs, supporting the species-specific variety of growth on diverse N sources and concentrations. Further research to investigate the competition of these strains may provide valuable insights into the role of substrate composition on the bacteria community.

In addition to lag time variation, growth profiles for methionine and alanine altered between the two species. Growth on methionine was expected due to its uptake by the N-II AAT [254], suggesting the different growth response may be a result of different capabilities of intracellular methionine degradation. While methionine biosynthesis and catabolism is not well-established in cyanobacteria, putative L-aspartate semialdehyde sulfurtransferases may provide a pathway between aspartate-4-semialdehyde and homocysteine (and subsequent methylation to methionine) [189]. With potential degradation pathways in both species, causes of the differential growth response remain to be elucidated. This is similarly seen with alanine utilisation yielding poor growth of *Synechococcus* sp. CCY9618. Transporters (N-II and N-III) and metabolic enzymes (*asd*) for alanine uptake and catabolism are encoded by both strains studied, suggesting indirect causes for the observed growth inhibition.

4.5.2 Differential Proteomic Response Between Amino Acid Nitrogen-Substrates

4.5.2.1 Amino Acid-Associated Pathways and a Potential Stress Response

A limited number of DEPs associated with AA biosynthesis were identified in this study, primarily under asparagine, glutamate, and proline N-substrate growth. The lack of variation in AA pathway up-regulation may be due to the central role these three AAs have in N metabolism. Glutamate itself is a key precursor for all AAs, while asparagine is hydrolysed by AsnB (asparaginase) to aspartate, itself a precursor to lysine, and further transaminated to glutamate [188, 189]. Meanwhile, proline can be converted to glutamate by PutA through two oxidative steps, though the activity of this pathway in picocyanobacteria is undetermined [265]. Thus, incorporation of these three AAs in the wider N framework may follow similar pathways.

The up-regulation of lysine biosynthesis under AA N-substrates compared to NO_3^- is notable (Table 4.7). Lysine accumulation has been linked to environmental stress response throughout the biosphere, from drought response [500], protection against oxidative stress [501–503], and osmotic stress mitigation [504]. Though mechanisms of the lysine-mediated stress response are not yet fully known, they are thought to involve an increase in lysine biosynthesis and subsequent conversion to various metabolites including saccharopine [502], cadaverine [505], and the compatible solute pipecolate [504]. Therefore, the up-regulation of lysine biosynthesis pathways in *Synechococcus* CCAP1479/10 when grown on asparagine, glutamate, and proline (compared to NO_3^-) may be an adaptive response to a stress stimulus, though results from the growth assay do not indicate a stress response, especially under glutamate and proline growth. Interestingly, an additional stress response protein (FtsH1) is up-regulated under arginine, asparagine, and glutamate-growth conditions compared to NO_3^-. FtsH1 is involved in the cyanobacterial nutrient stress response, forming a FtsH1/3 complex protease which digests transcription factors repressing activation of Fe, P, N, and inorganic C assimilation proteins [506]. The conditions in this study provide an excess of nutrient, thus the up-regulation of nutrient stress responses compared to NO_3^- is striking. The enrichment of lysine biosynthesis and up-regulation of FtsH1 may indicate that growth on AAs is not as efficient as growth on NO_3^-, though growth seemed to improve under some AAs in the growth assay, potentially suggesting short-term growth is favourable but long-term growth may have negative consequences. The additional catabolism involved in AA N-substrate growth may present difficulties in N incorporation, especially for substrates such as arginine without well-established degradation pathways. An additional source of stress may be the high concentration of AAs utilised in this study. This is much higher than would be expected in a natural environment, and the imbalance of the intracellular AA pool may induce a stress response though the consistent expression of NtcA along all conditions indicates N stress is not contributing to this response (Appendix Table B.21, Ga0436389_022_103393_104133).

In addition to lysine, arginine (glutamate- and proline-growth) and aromatic AA biosynthesis pathways (all conditions) contain up-regulated proteins. Arginine, along with aspartate, can form cyanophycin, a N storage molecule found in cyanobacteria. However, genes necessary for cyanophycin formation are not encoded in picocyanobacteria (Chapter 3). Thus, transfer of N into arginine is unable to be explained via the generation of cyanophycin, though increased abundance of PBS compared to N-starvation may indicate the utilisation of these light-harvesting antennae as N storage molecules. The additional doubt over arginine catabolism in *Synechococcus* spp. [266] leads to uncertainty for the cause of arginine biosynthesis flux. In addition, to arginine protein up-regulation, AroG of the shikimate pathway is up-regulated under all conditions compared to NO_3^-. The shikimate pathway yields chorismate, an aromatic compound which is an essential precursor to aromatic AA biosynthesis in addition to metabolites involved in electron transport and signalling [507]. The up-regulation of AroG, the first step of the shikimate pathway, may suggest greater flux into aromatic metabolites when grown on AAs. Further research to reveal aromatic metabolite yields may elucidate the functional purpose behind AroG up-regulation.

4.5.2.2 Nitrogen Assimilation DEPs Are Limited

N assimilation-associated DEPs offer limited insights into the response to AA N-substrates. The down-regulation of NtcA-dependent NrtA, UrtA, and GlnN in the majority of AA conditions compared to NO_3^- indicates differential expression for NtcA-regulated genes, though NtcA itself is not differentially expressed. On the other hand, the up-regulation of Amt1 under proline-growth is also controlled by NtcA expression. The purpose of an NH_4^+ transporter for *Synechococcus* sp. CCAP1479/10 grown under proline N-substrate may suggest the extracellular degradation of proline, releasing its amine group and subsequent ammonification, however if this mechanism was contributing to growth on proline it would be expected to enhance other AA N-substrate growth yet no accompanying Amt1 up-regulation is found in other conditions. Thus, the presence of Amt1 up-regulation is unclear, and may instead provide an outlet to transport excess NH_4^+ out of the cell. Of further interest is the expression patterns of the AATs identified in Chapter 3. The only differentially expressed AAT component was NatI, a substrate-binding protein, down-regulated under glutamate-growth compared to NO_3^-. This may be due to the presence of glutamate-specific transporters encoded by *Synechococcus* sp. CCAP1479/10, reducing the importance of ABC-type AATs for glutamate uptake.

4.5.2.3 Potential Implications of Photosynthesis and Translation-Associated DEPs

Consequences of growth on AA N-substrates or NO_3^- are revealed under proteomic analysis, and may indicate NO_3^- is a preferred source at high concentrations or the utilisation of C from the AAs. The down-regulation of photosynthesis and ribosomal proteins in *Synechococcus* sp. CCAP1479/10 when grown under AA N-substrates compared to NO_3^- is substantial (Figure 4.5). Down-regulation of photosynthesis and ribosomal proteins have previously been identified in cyanobacteria in response to a range of abiotic stressors (temperature, oxidative stress, nutrients) [472, 508]. The decrease in photosynthesis proteins may be caused by ROS-induced bleaching of the PS complexes through excess metabolites, limiting the photosynthesis reaction and reducing the generation of reduced NADPH and ATP [508]. This would subsequently reduce the energy capacity for respiration and other cellular processes, resulting in a down-regulation of ribosomal proteins. The up-regulation of tRNA ligases further supports the role of AA N-substrate contributing to a stress response, with phenylalanine-tRNA ligase previously identified as a common up-regulated protein in a heavily antibiotic-dosed cyanobacterial culture [474]. As such, high concentrations of the AA may be insufficiently regulated by cyanobacteria. Other bacteria are capable of converting glutamate to glutathione to mitigate oxidative stress [509], however picocyanobacteria lack glutamate dehydrogenase (*gdhA*) (Chapter 3), necessitating the use of alternative mitigation strategies [510]. Alternatively, photosynthesis proteins may be down-regulated due to intracellular AA degradation. The need for C fixation may be bypassed by utilisation of C from the imported AAs. Sufficient C would reduce the requirements for photosynthesis output, potentially resulting in the down-regulation of photosynthesis proteins to avoid unnecessary C fixation. The Redfield ratio of C and N requirements (6.6:1) against those in glutamate and proline (5:1) are similar, suggesting some alleviation of photosynthesis may be present. On the other hand, up-regulation of RNA polymerase may indicate adaptation to specific organic substrates. Previous long-term evolution studies have demonstrated that RNA polymerase is capable of acquiring mutations that change patterns of global expression, facilitating greater nutrient assimilation [511, 512]. Thus, changes in abundance of RNA polymerase may be linked to changes in the patterns of expression, further resulting in myriad metabolic changes, however the limited number of DEPs may suggest this is not the case.

4.6 Conclusions

Our investigation into the AA assimilation capabilities of freshwater picocyanobacteria have revealed a contrasting picture and ultimately has generated more questions than it has answered. Growth assays demonstrated a vast diversity in the bioavailability of AAs, revealing widespread growth capability while utilising AAs as sole N sources. However, the hypothesis of preferred substrates of N-II and N-III resulting in greater growth rates was not met. Furthermore, though lacking a basic AAT, picocyanobacteria exhibited growth on arginine and lysine, potentially indicating novel uptake capabilities or altered function of N-II (currently uncharacterised in picocyanobacteria). Degradation pathways in picocyanobacteria are not fully elucidated, especially for basic and aromatic AAs. The high N content of arginine has yielded multiple catabolism pathways in cyanobacteria, however none of these are fully encoded in *Synechococcus* spp. [266]. The lack of understanding over picocyanobacteria AA degradation results in unknown assimilation pathways where further experimental characterisation is required. Additionally, the variation in duration of lag phase displayed between *Synechococcus* sp. CCAP1479/10 and *Synechococcus* sp. CCY9618 demonstrates the variety within freshwater picocyanobacteria. Due to different AA N-substrates resulting in significant differences in lag time between the two strains, the nutrient composition of the freshwater environment (and the DON composition of oligotrophic environments) may heavily influence the picocyanobacteria community. The widespread ability to grow on AAs was greater than expected and indicates the utilisation of a substrate pool which has previously been relatively ignored. With other cyanobacteria displaying a decreased capability of AA bioavailability (Table 1.1), this extensive capacity to take advantage of the majority of the AA pool may contribute to the dominance of freshwater picocyanobacteria in oligotrophic environments.

Subsequent proteomic analysis of freshwater picocyanobacteria growth on AAs compared to NO_3^- resulted in the identification of a subtle stress response. Enrichment of lysine biosynthesis proteins and down-regulation of photosynthesis and ribosomal proteins when grown on arginine, asparagine, glutamate, and proline as the sole N source suggests inefficient growth on these substrates compared to NO_3^-. The cause is unclear though the buildup of intracellular AAs and subsequent accumulation of metabolite intermediates may induce this stress response. However, the limited stress response detected was insufficient to hinder picocyanobacteria growth on the AAs. The number of DEPs identified between the four AA conditions also provide insight into differential response to different AA N-substrates. The largest number of changes in the proteome were found under glutamate and proline growth compared to NO_3^-, with significant enrichment of AA-associated enzymes. Meanwhile, growth on the basic AA arginine revealed minimal changes in the proteome compared to NO_3^-, not enabling us to reveal basic AA assimilation mechanisms, instead suggesting extracellular arginine digestion or conversion to NO_3^- via nitric oxide. The varied response of picocyanobacteria to different AA substrates was demonstrated by the limited overlap of up-regulated DEPs between AA conditions, further supporting the role of DON composition in influencing the picocyanobacteria community. To conclude, the findings of this study reveal widespread ability to assimilate AAs among picocyanobacteria and reveals several insights into the mechanisms of AA assimilation and its cellular response.

4.7 Limitations

The work in this chapter was carried out to investigate freshwater picocyanobacteria AA assimilation over a 14-day period under two AA concentrations and comparing two picocyanobacteria strains. To enable this scale of measurements, the methods involved in cell proliferation detection and the growth assay itself had limitations. Firstly, measurements of OD_{750} in a spectrophotometer were utilised to detect cell growth upon the addition of the AA N-substrate. This turbidity-based method enables high throughput for measurements, though is not a direct measure of cell count [513]. The detection of dead cells and the potential of exopolysaccharide production [514] result in OD measurements which do not always correlate to the abundance of living cells. Alternative methods include the use of flow cytometry to detect single cells via forward and side scatter, yielding accurate quantification of cell counts. However, increased preparation time and reduced throughput rendered flow cytometry unsuitable for this large-scale investigation. Secondly, no measurement was taken to quantify the amount of AA imported into the cells. The inclusion of radioactively-labelled AAs would allow intracellular AA uptake to be quantified, as seen in previous studies [256]. This further raises the question of how the AAs are assimilated. Are they imported whole into the cell through the AATs, or are AA substrates degraded through cell-surface or excreted AA oxidases to yield NH_4^+? Previous research demonstrates AA uptake [255, 259] however the import of basic AAs remain unclear. Further research to track imported AAs through catabolism and subsequent incorporation may reveal novel metabolic pathways associated with AA assimilation. The implications of AA assimilation based on extracellularly degraded AAs represent additional genomic influence into the bioavailable DON pool, supporting the importance of the community composition in accessing various DON forms.

The high concentrations of AAs (250 mg N L^{-1}) used in the above growth assays and proteomic analyses are not found in the natural world nor environments under anthropogenic nutrient stress. This concentration was selected to match the N input of BG-11 media, utilising NO_3^- as its N source. Previous studies of bacterial AA-based growth have utilised similar concentrations [291, 515], though these concentrations have been associated with slight toxicity and stress responses [490, 494, 516]. In this study, while two AA concentrations were measured in the growth assay to determine growth under a standard medium concentration of N and a more environmentally-relevant concentration, proteomic analysis constraints limited growth to only one concentration. A longer lag phase and reduced bacterial yield under the low AA concentration restricted the down-stream applications for further picocyanobacteria analysis for this condition, necessitating the utilisation of high concentrations of AA to investigate the proteomic response to different AA N-substrates. The aims of this study were broad and act as a general insight into the molecular mechanisms of AA assimilation in picocyanobacteria. In oligotrophic environments, concentrations of DFAA may be substantially lower than 1 mg N L^{-1} [423]; however, this work indicates bioavailable mechanisms are present to assimilate (or extracellularly degrade) a variety of AAs. Further studies to investigate picocyanobacteria AA assimilation capabilities and the influence of the AA-pool should utilise further reduced AA concentrations to mimic natural conditions and narrow the focus to expressed AATs and the mechanism of AA import (whole or extracellularly-degraded).

A limitation with the proteomic analysis in this study is the low proportion of peptides identified compared to predicted proteins encoded in the genome. Though the proportion of peptides identified is consistent with other proteomic studies (see Section 4.4.3), 75% of the proteome is still missed. While the total protein content will not be constitutively expressed, and linked to other conditions such as P, high light, low temperature, etc., it does yield a large pool of potentially expressed protein with unknown activity. A consequence of this low representation is the limited picture of proteome change it generates. A major example of this is *Synechococcus* sp. CCAP1479/10 growth on arginine compared to NO_3^-. This comparison resulted in only seven proteins up-regulated under arginine-growth, while growth on asparagine, glutamate, and proline generated DEP numbers comparable with other studies [473, 485]. It is unclear why growth on arginine yielded so few DEPs compared to the other substrates, yet the cellular response to basic AAs may be hidden in the 75% of proteins not detected in this study. Further possibilities of DEP reduction may be due to the mechanism of arginine uptake and assimilation. Extracellular degradation of arginine and conversion of arginine to NO_3^- through nitric oxide are both mechanisms which may not affect the proteome significantly, though some indication of these mechanisms would still be expected to be seen.

4.8 Future Directions

The work in this chapter is a broad investigation of the AA assimilation capabilities of picocyanobacteria and their proteomic response to different AA N-substrates. The widespread bioavailability yet putative stress response suggests future work should focus on the role of the AATs and elucidation of AA degradation pathways. To further investigate the expression of AATs under varying AA compositions, targeted quantification of these membrane complexes should be carried out. Though proteomic analysis can provide a general overview of the intracellular protein content, it has limitations in regards to the quantification of membrane proteins [517, 518]. Of the 10 AAT components encoded by *Synechococcus* sp. CCAP1479/10, only three were detected by TMT-quantitation. AAT expression could be targeted through qPCR and will aid to determine the importance of encoded AAT variety in the picocyanobacteria response to AA sources. Furthermore, elucidation of the pathways and mechanisms involved in cyanobacterial AA degradation is required, with an emphasis on basic AAs. ^{13}C, ^{15}N-AA spiking to trace degradation and subsequent incorporation of N and C will expand our understanding of AA assimilation and how these substrates are utilised as N (and C) sources.

This study provides evidence for intraspecific variation within freshwater picocyanobacteria. As the duration of lag phase differs significantly between the two species of *Synechococcus* studied, this suggests that the composition of AAs in oligotrophic environments may influence the picocyanobacteria community. Further experiments to investigate a mixed community of picocyanobacteria and measure its composition shift as a response to varying AAs will reveal the impact of anthropogenic organic pollution on keystone primary producers. Greater understanding of the association between nutrient input and community composition in conjunction with widespread nutrient measurements will enable accurate predictions of hazardous aquatic conditions, including the formation of blooms and cyanotoxin generation.

Chapter 5

Conclusions

5.1 Contributions and Acknowledgements

This chapter was written with feedback from Patricia Sánchez-Baracaldo and Stephen Maberly.

5.2 Summary

Picocyanobacteria are a significant source of global primary production, and their ability to dominate their bacterial community is well known [296]. As such, they have attracted large amounts of research interest albeit primarily focusing on marine environments. Many gaps remain in our understanding of the picocyanobacteria of freshwater environments; though their importance in biogeochemical cycles is known, a lack of molecular and genomic approaches to their study has limited our knowledge of their internal physiological processes. Of particular interest is the picocyanobacterial ability to assimilate nitrogen (N), an essential nutrient that is often (co)limiting in the oligotrophic environments where picocyanobacteria thrive [8]. The overall aim of this thesis was to increase our understanding of freshwater picocyanobacteria, especially in relation to their capability to assimilate various species of N with a focus on organic sources. This was carried out through an ecological genomics framework, utilising genomic data to explore ecological behaviour and identify the versatility in picocyanobacteria N uptake.

To achieve the objectives of this thesis, outlined in Chapter 1, the first step was to increase genomic representation of freshwater picocyanobacteria. This involved the sequencing of five non-axenic *Syn/Pro* sub-cluster 5.2 strains, generating draft assemblies, and comparing the genomic features of these to the freshwater *Synechococcus* model of *Synechococcus elongatus*. Following this, large-scale comparative genomics analysis was carried out on 166 cyanobacteria, comparing the capacity to encode molecular machinery associated with both inorganic and organic N assimilation. This study compared the N assimilation capabilities of freshwater picocyanobacteria to picocyanobacteria of other environments and larger freshwater cyanobacteria. Finally, the encoding patterns gleamed from the comparative genomics analysis was used to investigate the ability of freshwater picocyanobacteria to utilise amino acids (AAs) as sole N sources and to explore the proteomic response of growth on four selected substrates. An overview of the findings of this thesis are provided below:

Objective 1: Increase the representation of freshwater picocyanobacteria by sequencing five strains and comparing these genomes to *Synechococcus elongatus*

Five freshwater picocyanobacteria were obtained, three isolated from Windermere, UK (*Synechococcus* sp. CCAP1479/9; *Synechococcus* sp. CCAP1479/10; *Synechococcus* sp. CCAP1479/13) and two isolated from lakes in the Netherlands (*Synechococcus* sp. CCY0621 and *Synechococcus* sp. CCY9618). The DNA of these strains were extracted, sequenced, and assembled, resulting in five draft genomes ranging from 2.9 Mbps to 3.3 Mbps with completeness of 98% to 99%. Maximum likelihood phylogenomic analysis placed these strains within sub-cluster 5.2 of the *Syn/Pro* clade.

With the genome pool of freshwater picocyanobacteria expanded, genomic features of these strains were compared to that of *Synechococcus elongatus*, an often-used model organism of freshwater cyanobacteria. EggNOG functional annotation revealed significant genomic diversity between the newly sequenced strains and *Synechococcus elongatus*, while in-depth comparative analysis of genes associated with photosynthesis and NO_3^-/NO_2^- assimilation revealed further interesting differences. For example, the antennae protein allophycocyanin is encoded in greater copy number by *Synechococcus elongatus* - multiple copies of *apcD* are associated with the ability to absorb reduced levels of light, potentially favouring *Synechococcus elongatus* growth in shaded conditions. Furthermore, the presence of NO_2^--associated *nirB* and *ntcB* in *Synechococcus elongatus*, yet absent from the sequenced freshwater picocyanobacteria, indicate an enhanced response to environmental NO_2^- in *Synechococcus elongatus*, suggesting varying nutrient preferences between the two taxa. Finally, the molecular machinery for NO_3^- assimilation displays variety within the *Syn/Pro* clade, with *Synechococcus* sp. CCY9618 encoding an alternative method of NO_3^- uptake than the remaining four newly sequenced strains. This work showcases the differences between *Synechococcus elongatus* and freshwater picocyanobacteria, highlighting the need for model freshwater picocyanobacteria and greater understanding of the diversity within the *Syn/Pro*.

Objective 2: Investigate the encoded N assimilation capabilities of freshwater picocyanobacteria, comparing these to marine picocyanobacteria, thermophylic picocyanobacteria, *Synechococcus elongatus*, and larger freshwater cyanobacteria

Comparative genomic analysis was carried out investigating the presence of 171 genes associated with N assimilation in freshwater picocyanobacteria, and comparing this to their presence in picocyanobacteria of other habitats and larger freshwater cyanobacteria. This increased our understanding of N assimilation capabilities present in freshwater picocyanobacteria and how these mechanisms are a product of adaptation to their environment (whether freshwater, marine, or thermal).

We find that amino acid transporters (AATs) provide the greatest source of diversity among N assimilation genes in freshwater environments. This may suggest that the composition of free amino acids in the dissolved organic nitrogen (DON) nutrient fraction may influence the cyanobacteria community based on uptake preferences for individual AATs. We further find differential encoding of uptake mechanisms of other organic N sources, including dissolved combined amino acids (DCAA), chitin, and glyphosate. The broad uptake capacity of AAs and other organic N sources in freshwater picocyanobacteria, the potentially preferred uptake of glycine (and its high concentration in freshwater environments [171]), along with a large surface area to volume ratio, may provoke a quick uptake response by freshwater picocyanobacteria to minimal but varied nutrient input and promote their dominance in these environments.

Objective 3: Investigate the capability of freshwater picocyanobacteria to utilise AAs as sole N sources and explore the proteomic response to growth on glutamate, proline, arginine, and asparagine

The results of Chapter 3 prompted an investigation into AA uptake capabilities of two freshwater picocyanobacteria strains, *Synechococcus* sp. CCAP1479/10 and *Synechococcus* sp. CCY9618. A growth assay of 20 proteinogenic AAs revealed widespread AA assimilation in freshwater picocyanobacteria at both high and low AA concentrations, with 18 AA substrates promoting growth as a sole N source. On the other hand, growth on novel organic N species - chitin and glyphosate - was either not identified or non-quantifiable (in the case of high concentrations of chitin), indicating these nitrogenous compounds may not be effective N sources for freshwater picocyanobacteria though further analysis is required. Furthermore, the hypothesised preferences of AA assimilation based on AATs (glycine, glutamate, and aspartate) did not yield significantly greater growth rates than other AA substrates. The growth of the two strains on the basic AAs arginine and lysine represents an interesting finding. As freshwater picocyanobacteria lack a basic AAT, the assimilation of arginine and lysine may indicate a novel uptake mechanism, an increased range of AA uptake by encoded AATs, or secretion of AA oxidase enzymes to catalyse extra-cellular degradation of these AAs. This suggests that solely utilising genomics to predict cyanobacterial growth patterns is ineffective, instead requiring a holistic overview of the microbial community and potential degradative pathways. Furthermore, the variety of lag phase between the two *Synechococcus* spp. indicates variety within the *Syn/Pro* clade, providing evidence for the nutrient source influencing the composition of the picocyanobacterial community.

An exploration of the proteomic response to growth on glutamate, proline, arginine, and asparagine revealed a mixed response. Although growth on these AAs were comparable to growth on NO_3^- and NH_4^+ in the growth assay, proteomics revealed a potential stress response for these substrates (compared to NO_3^-); enrichment of lysine biosynthesis enzymes indicates an increase in lysine production, a common physiological stress response potentially caused by an accumulation of AA metabolic intermediates, in addition to decreased ribosomal proteins and photosynthesis (though this may also be linked to increased AA-carbon uptake). The finding of few up-regulated proteins when grown on arginine compared to NO_3^- is also surprising. While these findings were not able to accurately elucidate arginine catabolic mechanisms in picocyanobacteria, it may suggest the extracellular degradation of arginine to NO_3^- via a recently discovered pathway [519]. Finally, the up-regulated proteins of the four AA substrates compared to NO_3^- displayed a varied response, with a large number of proteins up-regulated under only one substrate. This suggests that a distinct metabolic response is required when grown on specific AAs, in turn promoting the idea of DON speciation influencing the microbial community.

5.3 Implications

The research conducted in this thesis has improved our understanding of the N assimilatory capabilities of freshwater picocyanobacteria with an emphasis on organic sources of N. This work has also highlighted the importance of studying the *Syn/Pro* clade, a group of which it is becoming increasingly clear have an integral role as primary producers in N limiting and co-limiting environments worldwide. While genomic investigations into this keystone clade are increasing in scope [33, 103], experimental molecular characterisation of freshwater picocyanobacteria is still lacking. Furthermore, the utilisation of genomic data to address environmental questions is a relatively novel approach, studying encoded molecular machinery to elucidate ecological behaviours from the organism perspective [10, 520]. This thesis reveals the molecular diversity of N assimilation within picocyanobacteria and freshwater cyanobacteria, and combined with recent studies which have investigated the diversity of nutrient metabolism within the *Syn/Pro* [103], provides a genomic foundation from which picocyanobacterial dominance can be investigated.

A major theme of this thesis is the utilisation of organic sources of N. These N sources are especially prevalent in oligotrophic environments [152, 153] - environments where picocyanobacteria are integral sources of biomass. With AAs a large component of bioavailable DON, the diversity of encoded AATs between freshwater picocyanobacteria and larger cyanobacteria may be crucial to understanding cyanobacterial communities in DON-dominated environments. While NO_3^- and NH_4^+ are routinely measured in water bodies, the concentration of DON (including urea and AA) is usually ignored. This may present DON as an undetected driver of microbial communities in these environments, potentially resulting in bloom and cyanotoxin generation as a product of the DON composition. In contrast, the experimental work on freshwater picocyanobacteria conducted in this thesis (Chapter 4) found widespread growth on AA substrates, yet uptake of basic AAs was unexpected based on the anticipated substrates of the encoded AATs. This indicates the complexity in identifying potential bioavailable nutrients. The implication of freshwater picocyanobacteria growth on arginine suggests an unknown factor involved in its assimilation, whether this be involved in uptake of arginine or extracellular arginine catabolism. A greater understanding of the AA catabolic pathways, especially for aromatic and basic AAs, are required in freshwater picocyanobacteria. While genomic data can provide insights into putative assimilatory pathways and mechanisms, the experimental characterisation of these is essential to match an expected phenotype to a genotype.

Lastly, it is difficult to accurately extrapolate natural microbial behaviours from mono-cultural laboratory experiments. Freshwater picocyanobacteria live amongst a community, with each member of that community potentially encoding differing mechanisms for nutrient assimilation and breakdown. This thesis identifies subtle encoded differences within the *Syn/Pro* clade, whether related to NO_2^- metabolism (Chapter 2) or glutamate uptake (Chapter 3). The impact of differentially encoded AATs and differing AA preferences on a microbial community must be investigated with a microbial community, a task which was not within the scope of this thesis. While this thesis has begun to uncover the diversity within freshwater picocyanobacteria, more extensive sampling of this group must be carried out. Furthermore, predictive patterns of DON composition can only be elucidated through community experiments. An holistic view of primary producer communities, incorporating species-level physiological variance and genomic diversity, will enable us to better understand the bacterial dynamics of this essential biomass.

5.4 Critical Analysis and Future Directions

5.4.1 The Dominance of Picocyanobacteria

This thesis has improved our knowledge of the N assimilatory capabilities of freshwater picocyanobacteria through comparative genomics and increased genomic representation (Chapters 2 and 3) followed by experimental validation of AA assimilation in Chapter 4. The insights gained from this research has proposed that freshwater picocyanobacteria may have access to a greater diversity of bioavailable N sources than marine picocyanobacteria, while AA assimilation uptake systems differ within cyanobacteria of freshwater environments though the extent of the influence of this on the microbial community remains to be elucidated. The results of this thesis propose that a contribution to the dominance of picocyanobacteria in the freshwater cyanobacterial community may be due to their assimilation capabilities of a wide range of AAs, though comparisons to their eukaryotic algal competitors are required to infer this advantage over the wider microbial community. This is linked to the morphology of picocyanobacteria, enabling rapid uptake of nutrients - an essential capability for AA uptake due to the short turnover time of these molecules in aquatic environments. Thus, the ability to respond very quickly to nutrient inflow, especially in oligotrophic environments where competition is high, in combination with a large pool of potential bioavailable N sources may contribute to the dominance of picocyanobacteria in freshwater environments. Future research to compare freshwater picocyanobacteria nutrient uptake capabilities to those of eukaryote algae would reveal the nutrient dynamics of the wider phytoplankton ecosystem and aid in elucidating the behavioural patterns of these essential groups.

The focus on N in this thesis neglects multiple other processes and nutrients which are essential for picocyanobacteria proliferation. Though we have been considering AAs as sources of N, they also constitute a carbon (C) source. An essential macronutrient, C can be a limiting factor for freshwater phytoplankton growth, especially when CO_2 becomes depleted by rapid uptake [521]. Isotope flow analysis of dual-labelled ^{13}C,^{15}N-glycine and glutamate in soils demonstrates that N is incorporated at a greater rate than C, promoting the designation of AAs as primarily N sources (though the biogeochemistry of soil is distinct from aquatic environments) [522]. However, supersaturation of CO_2 in unproductive lakes can enhance biomass formation, suggesting the utilisation of AA-C alongside AA-N in oligotrophic environments is highly valuable [521]. In addition to C and N, phosphorus (P) is typically a (co)limiting nutrient in freshwater environments. The picocyanobacterial capabilities for P uptake has, like N, focused on marine environments [385, 523, 524], though research focusing on *phnD*, a component of the phosphonate transporter, has also been conducted in fresh waters [460, 525]. This has identified the bioavailability of organic phosphonates among picocyanobacteria, demonstrating the importance of dissolved organic matter to the bioavailable nutrient pool. Thus, it is insufficient to solely discuss N assimilation capabilities when contemplating picocyanobacterial dominance, and instead the role of P and C, must also be incorporated.

Recent work has postulated that the carboxysome, a C concentrating compartment which surrounds the CO_2-fixing enzyme RUBisCO, also contributes to picocyanobacterial dominance. The inability of RuBisCO to differentiate between CO_2 and O_2 is problematic for cyanobacterial photosynthesis productivity; this is mitigated via the carboxysome, responsible for converting bicarbonate into CO_2 through the action of carbonic anhydrase to increase the local concentration of CO_2 around RuBisCO, preventing photorespiration [526]. There are two forms of carboxysome found in cyanobacteria, an α and a β form. These two forms confer very similar catalytic rates to RuBisCO and function identically, though α-carboxysomes are generally smaller than β-carboxysomes [527]. However, their distribution is interesting - the majority of cyanobacteria possess a β-carboxysome, including typical model strains of *Synechocystis*, *Anabaena*, and *Synechococcus elongatus*. In contrast, picocyanobacteria of the *Syn/Pro* clade possess the α-carboxysome form. Recent research hypothesises that this is a response to carbonate chemistry, with picocyanobacteria dominating chemically stable, often slightly alkaline, environments such as the open ocean and large freshwater lakes where carbonate is the primary form of inorganic C. Meanwhile, β-cyanobacteria are found in environments which feature greater carbonate and oxygen flux, with inorganic C fluctuating between CO_2, bicarbonate, and carbonate [528]. This suggests that C specialisation may contribute to picocyanobacterial dominance, supported by greater encoded diversity of inorganic C uptake mechanisms in β-cyanobacteria [528].

5.4.2 The Complete Role of Dissolved Organic Nitrogen

The bioavailability of AAs has been elucidated in this thesis yet these make up only a small proportion of total DON. With urea, AAs, and nucleotides the most easily characterisable components of DON, unsurprisingly investigation into the bioavailability of these molecules have been the focus of DON research [131, 145, 184, 291, 515, 529, 530]. Humic substances also constitute a major component of DON, yet their molecular diversity is more complex [531]. However, ultrahigh resolution mass spectronomy enables the separation of these compounds and the determination of the individual humic acids, fulvic acids, and humins [532]. The bioavailability of humic substances has been demonstrated [533, 534], yet assimilation mechanisms have not been determined. The high molecular weight of humic substances necessitates extracellular degradation, requiring appropriate molecular machinery to liberate N. Due to this, the composition of the microbial community may be key, wherein taxa encoding the requisite enzymes for humic catabolism, along with the ability to secrete these enzymes, may be vital for humic bioavailability. This would allow the community of plankton, including picocyanobacteria, the ability to utilise a greater proportion of the total N pool than would a single species. On the other hand, while known DON components may have significant roles in contributing to the available N pool, the majority of DON compounds are still unidentified [177]. Further work to clarify the contribution of humic substances and unidentified compounds to the bioavailable DON pool, and subsequent associations with the microbial community, will aid our understanding of the role of DON and the relationship between nutrient source and community composition. This applies especially to oligotrophic environments where DON utilisation is greatest [154].

5.4.3 Limitations and Knowledge Gaps

The five freshwater picocyanobacteria genomes sequenced during this study were not assembled into single, closed chromosomes. As seen in Chapter 2, this can create artefacts during analyses, yielding false negatives. Meanwhile, the majority of freshwater picocyanobacteria utilised in the analyses of Chapter 3 are draft, non-closed, genomes. While the draft genomes of Chapters 2 and 3 are often composed of high coverage reads, hence assumed to cover most (if not all) of the bacterial genome, the uncertainty of missing sequences limits the confidence of gene presence/absence results. This can be resolved through the generation of closed genomes, a technique that is becoming more accessible due to the development of long-read sequencing technologies such as single-molecule real-time sequencing and nanopore sequencing [535]. Closed genomes increase our accuracy of presence/absence investigations, in addition to clarifying genomic loci and rearrangements. With the accessibility of long-read sequencing increasing, closed genomes should become the new standard.

A major limitation of comparative genomic analyses is that targeted proteins are determined based on homology, with molecular characterisation especially lacking in freshwater picocyanobacteria. We have already seen in this thesis that homologous proteins can have different functions in different genera, for example ApcD and its role in photosynthesis in *Synechocystis* sp. PCC 6803 and *Synechococcus elongatus* PCC 7942 (Chapter 2). Another example features the proteins of the ATP-binding cassette (ABC)-type transporter for NO_3^- and NO_2^- in freshwater cyanobacteria; these are very similar to the components of the bicarbonate ABC-type transporter - seemingly homologous proteins transporting different substrates. Due to this, homologous protein function may not be exactly as expected. Additionally, the ability to express identified genes cannot be taken for granted, with the *nif* operon found in

Vulcanococcus limneticus LL, yet not able to be expressed [33]. These issues can be addressed by effective replication of true environmental conditions [470], but also by transcriptomics and proteomics, though these approaches have limitations themselves (as seen in Chapter 4). The most accurate method to determine protein function and their physiological importance is the development of knock-out mutant strains, however this requires the establishment of a model organism. An ideal model organism is amenable to genetic modification; it is as yet unclear how responsive freshwater picocyanobacteria are to genetic manipulation due to the current lack of knock-out experiments, though the potential generation of a freshwater picocyanobacterium model organism will enable a greater understanding of the genomic adaptations of this keystone group.

Significant knowledge gaps exist within our understanding of the cyanobacterial capability to metabolise AAs. For a phylum which is (both historically and presently) so essential for life on Earth, the catabolic mechanisms for AA degradation remain under-studied, though the findings of Chapter 4 suggest successful assimilation of these molecules. To enhance our understanding of the differential role individual AAs have on growth and community composition, our ability to trace the N source to N incorporation is vital; elucidation of these catabolic pathways and identification of the molecular machinery involved will enable us to predict, with greater accuracy, AA assimilation preferences, with implications on picocyanobacterial dominance, community composition and water management of cyanobacterial blooms.

5.5 Final Remarks

The importance of freshwater picocyanobacteria is immense, and their role in global primary production and ecological stability cannot be overstated. Conversely, lack of genomic representation has hindered effective research into the molecular machinery of this essential group, leaving much unexplored.

This thesis represents an attempt to increase our understanding of freshwater picocyanobacteria, focusing on N assimilation capabilities of organic sources. Although much more is required to untangle the complexities of DON uptake, this thesis provides evidence for a broad range of AA uptake in freshwater picocyanobacteria, in addition to prompting the hypothesis of different DON compositions influencing the make up of the microbial community. However, further questions are raised by the work carried out. For example, what are the implications for the variety of AATs we find among freshwater cyanobacteria? Does this have a role in the freshwater dominance of picocyanobacteria? Although not identified in this study, does chitin and glyphosate have a role as N sources in natural environments? Are AAs a more important source of N for freshwater picocyanobacteria compared to larger freshwater cyanobacteria? Does the broad range of AA uptake found in freshwater picocyanobacteria contribute to their dominance? Finally, we've uncovered diversity within freshwater cyanobacteria - what is the diversity found in freshwater picocyanobacteria? This thesis has accomplished its objectives and prompted research questions to further enhance our understanding of freshwater picocyanobacteria.

Overall, this thesis has:

- Increased the representation of freshwater picocyanobacteria at a genomic level
- Highlighted the need for a freshwater picocyanobacterium model organism due to genomic differences with *Synechococcus elongatus*
- Revealed the N assimilation capabilities of freshwater picocyanobacteria, with major differences in organic N transport systems
- Determined widespread AA bioavailability among freshwater picocyanobacteria
- Explored the proteomic response of freshwater picocyanobacteria on AA N-substrate

Appendix A

Supplementary Figures

List of Figures in Appendix A

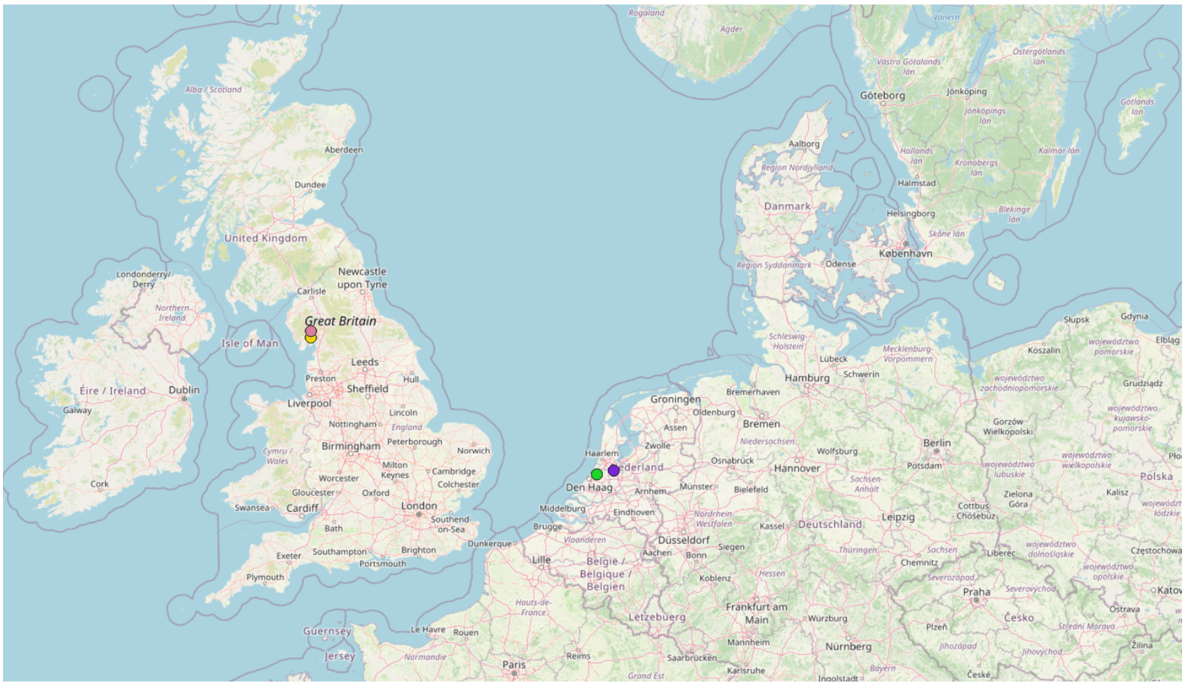

Figure A.1: **Location of picocyanobacteria isolation**. *Synechococcus* sp. CCAP1479/9 and *Synechococcus* sp. CCAP1479/10 were both collected from the south basin of Windermere, the UK (yellow marker). *Synechococcus* sp. CCAP1479/13 was collected from the north basin of Windermere, the UK (red marker). *Synechococcus* sp. CCY0621 was isolated in Leiden, the Netherlands (green marker) and *Synechococcus* sp. CCY9618 was isolated in Vinkeveen, the Netherlands (purple marker).

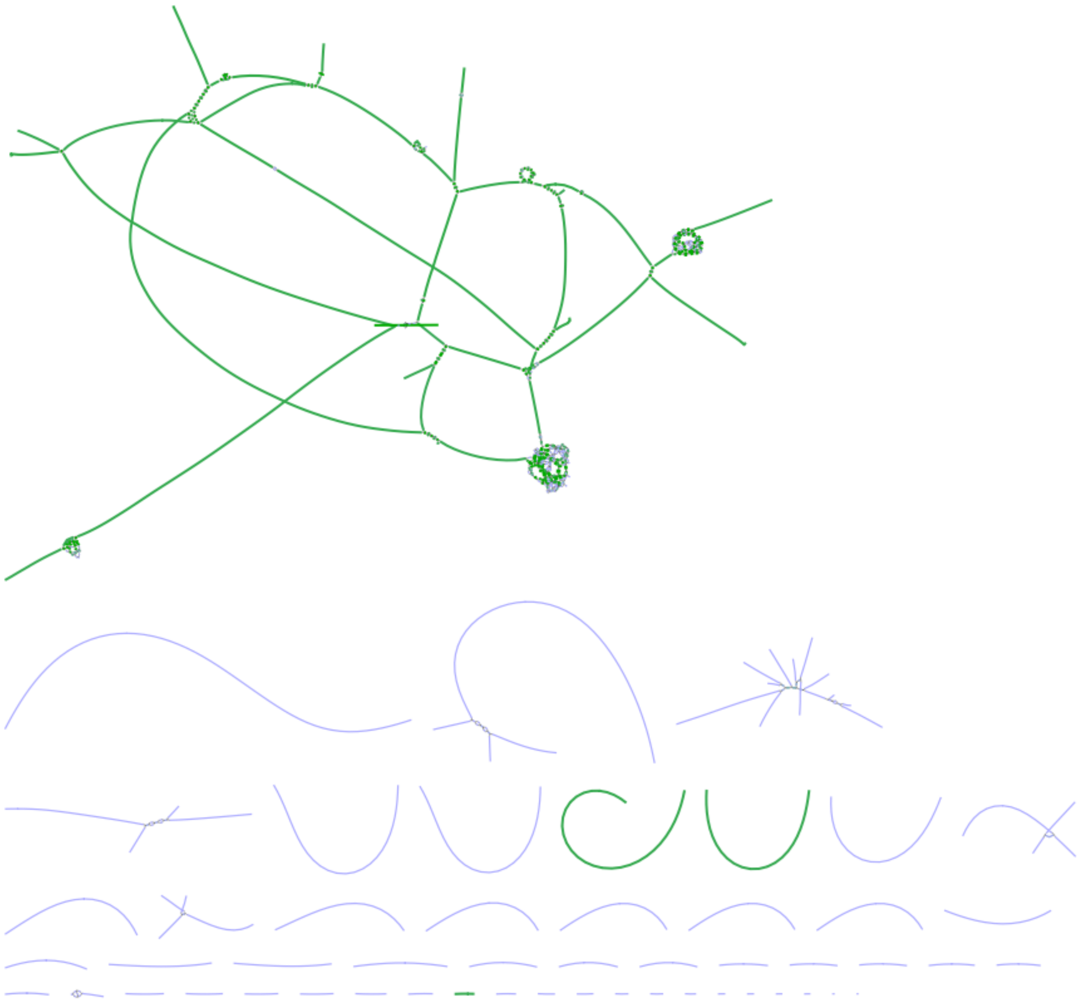

Figure A.2: **Assembled reads of *Synechococcus* sp. CCY0621 using De Bruijn Graph visualisation**. Output from Bandage v0.8.1. Nodes with a depth >243 are in green, nodes with a depth <15 are in lilac.

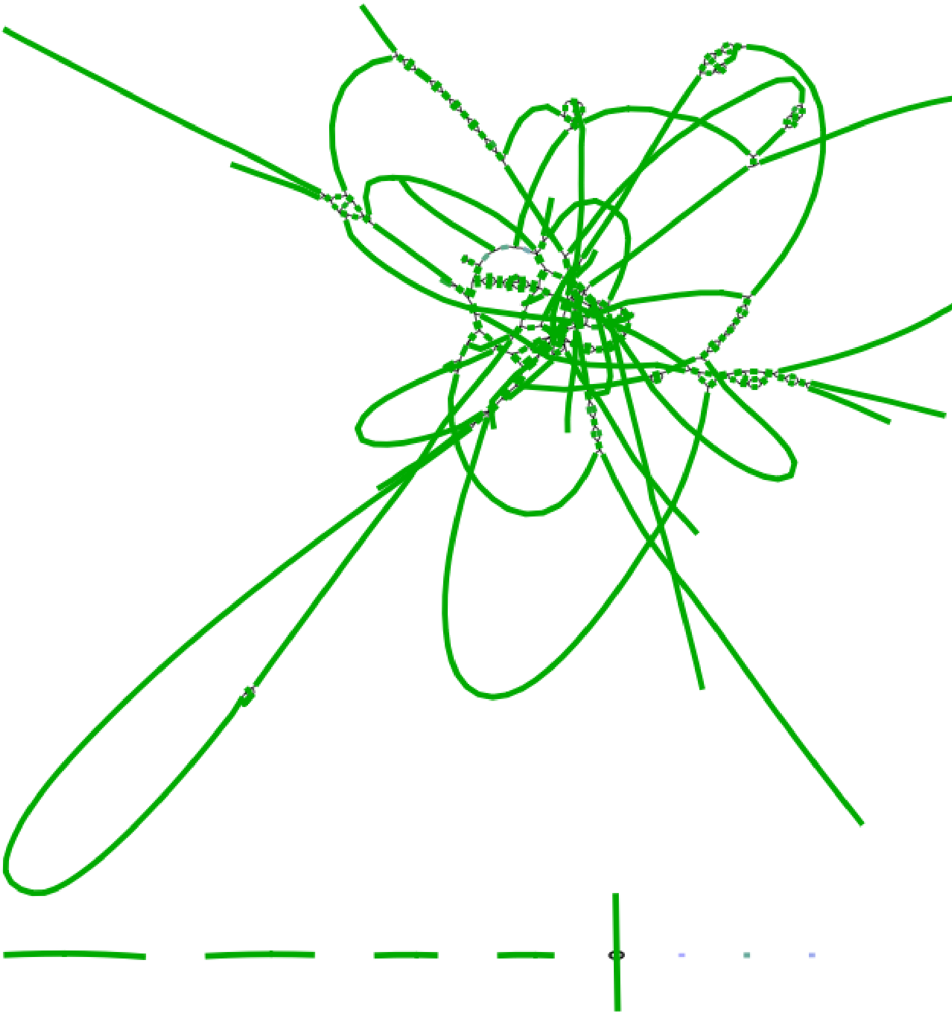

Figure A.3: **Assembled reads of *Synechococcus* sp. CCY9618 using De Bruijn Graph visualisation**. Output from Bandage v0.8.1. Nodes with a depth >175 are in green, nodes with a depth <30 are in lilac.

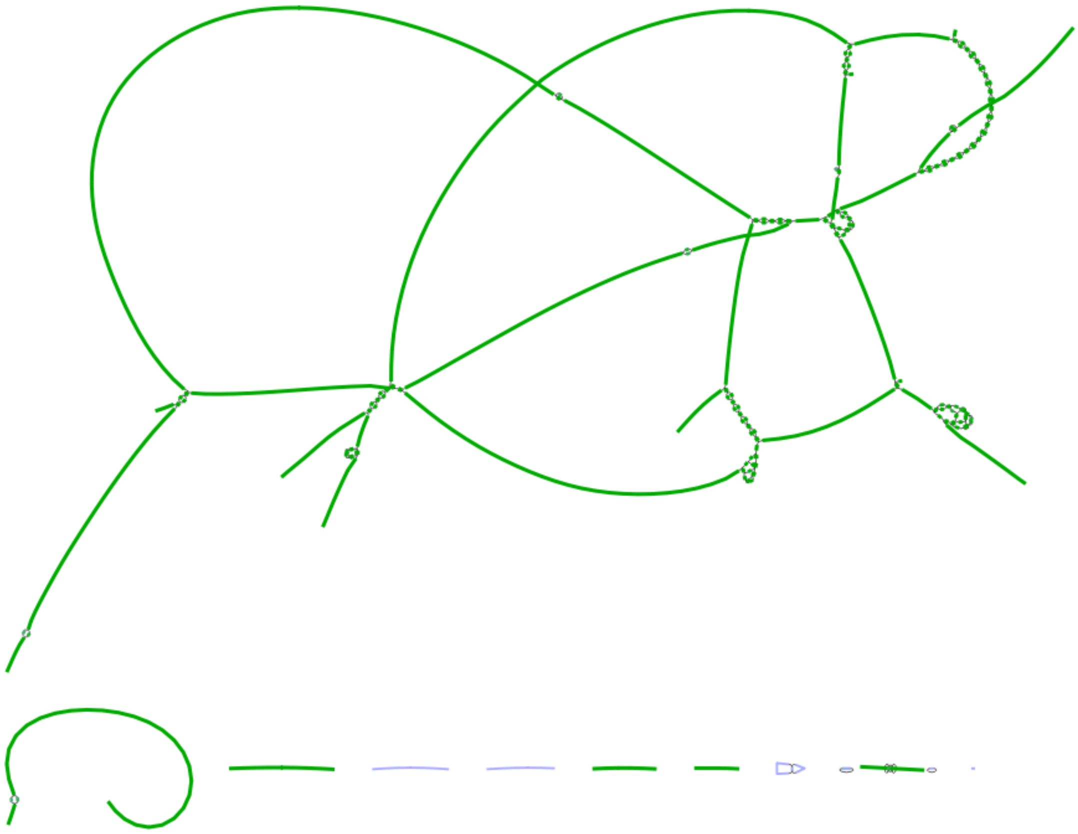

Figure A.4: **Assembled reads of *Synechococcus* sp. CCAP1479/9 using De Bruijn Graph visualisation**. Output from Bandage v0.8.1. Nodes with a depth >181 are in green, nodes with a depth <21 are in lilac.

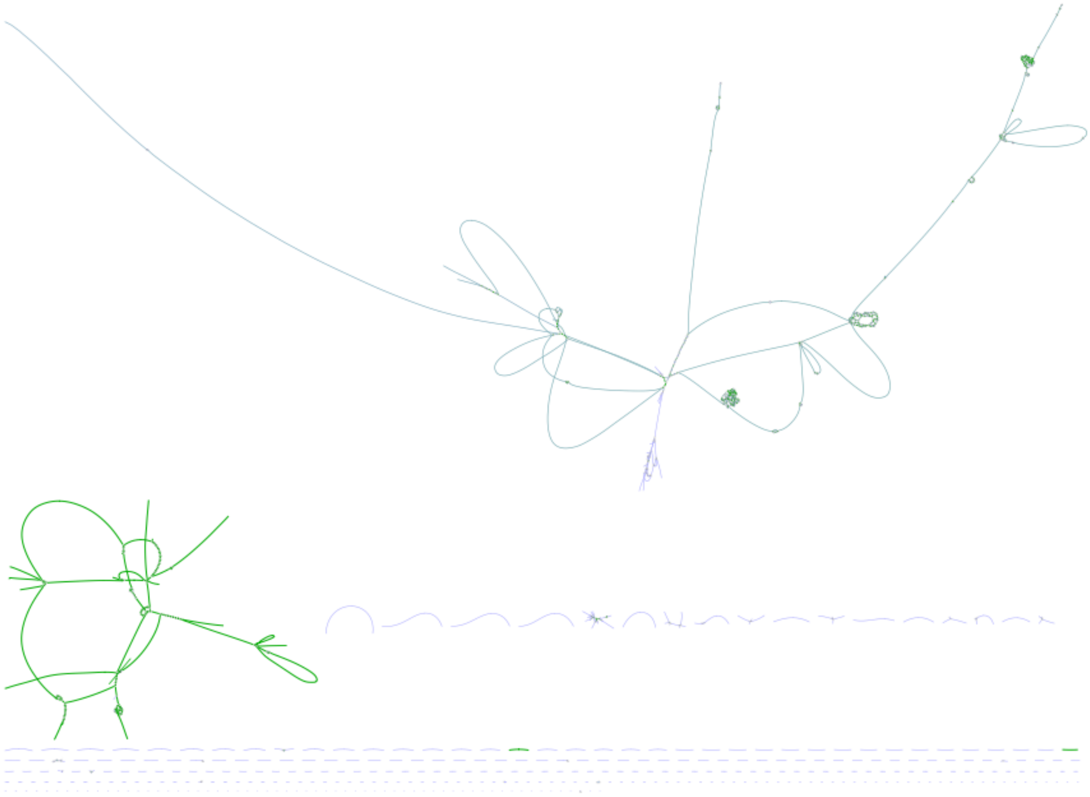

Figure A.5: **Assembled reads of *Synechococcus* sp. CCAP1479/10 using De Bruijn Graph visualisation**. Output from Bandage v0.8.1. Nodes with a depth >119 are in green, nodes with a depth <17 are in lilac.

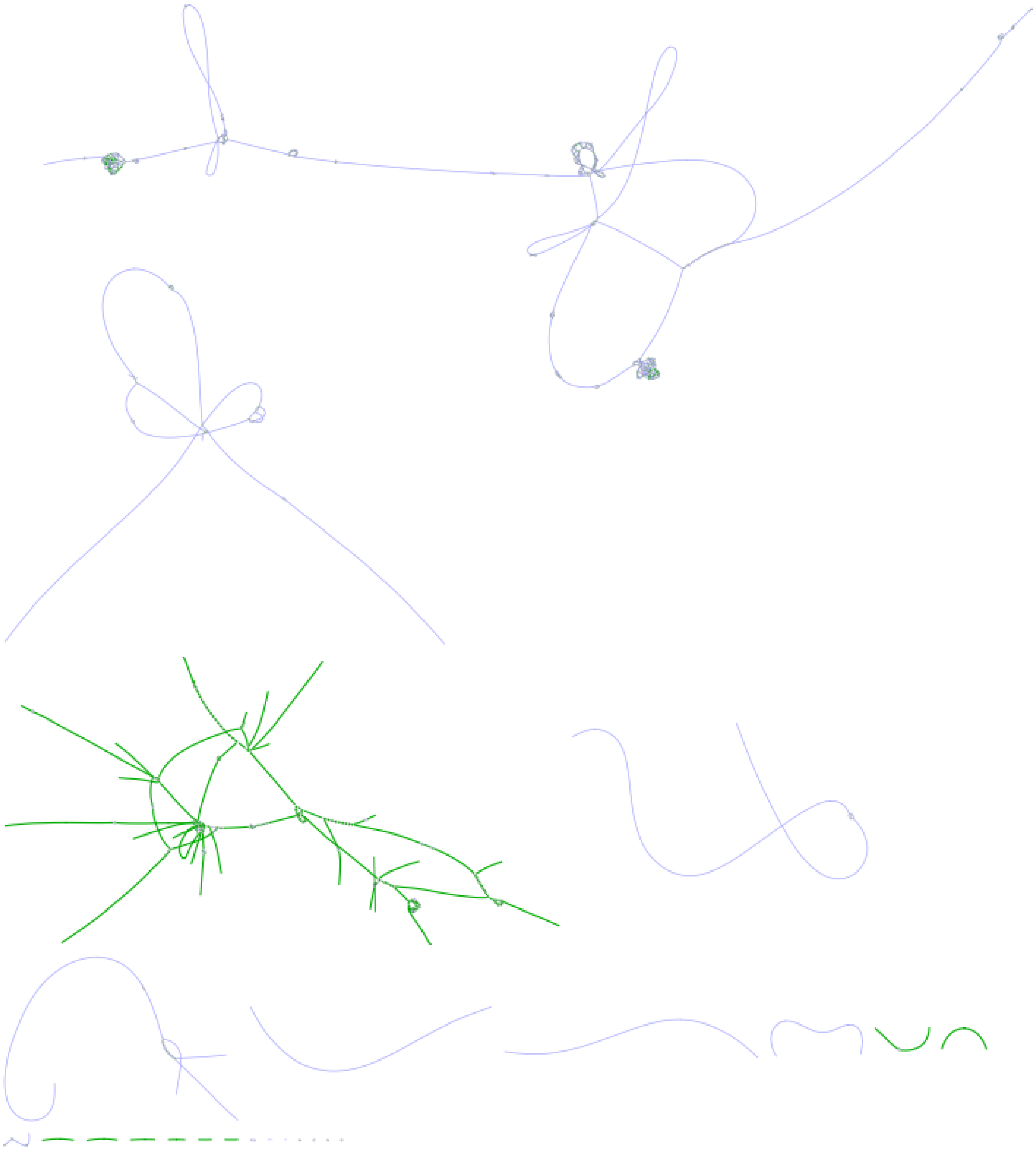

Figure A.6: **Assembled reads of *Synechococcus* sp. CCAP1479/13 using De Bruijn Graph visualisation**. Output from Bandage v0.8.1. Nodes with a depth >146 are in green, nodes with a depth <27 are in lilac.

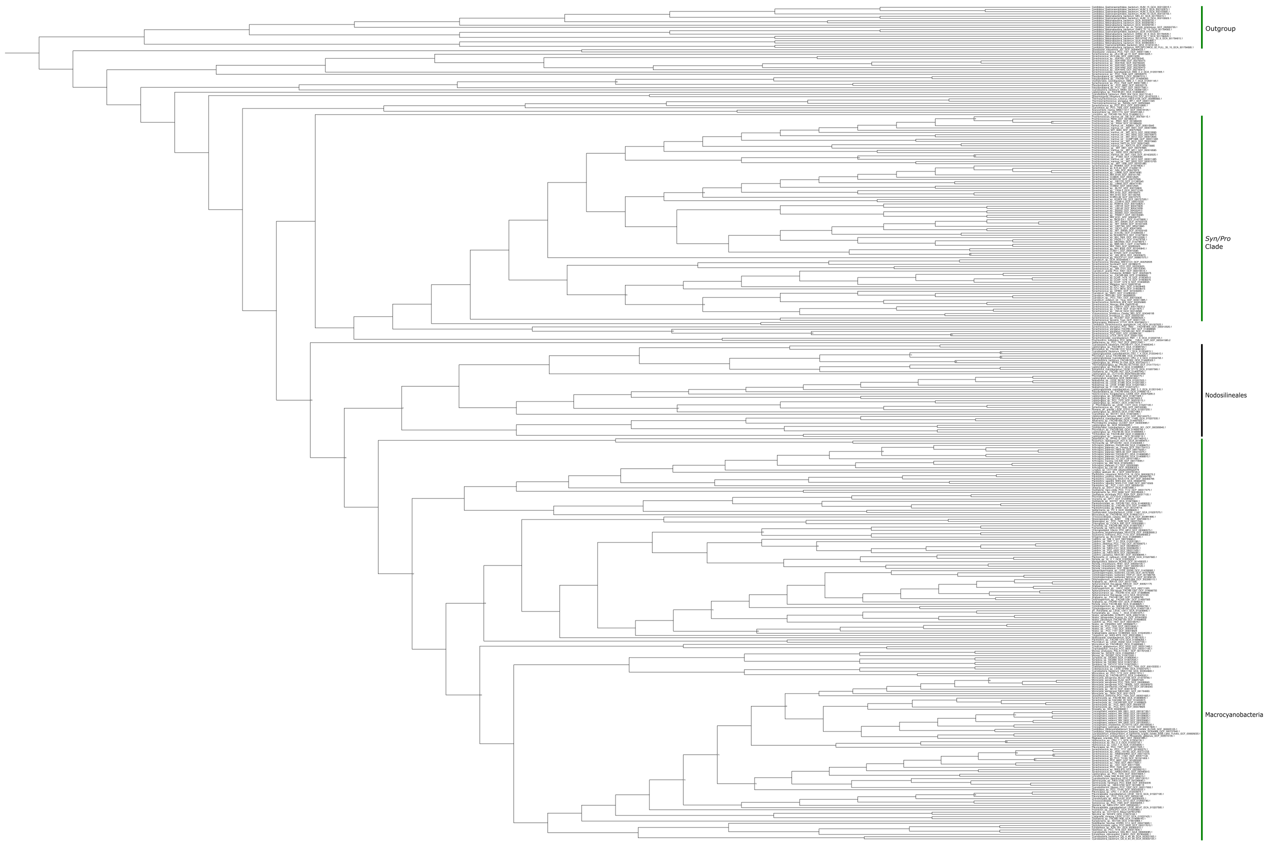

Figure A.7: **Expanded view of the phylogeny featuring all cyanobacteria utilised in a maximum likelihood analysis based off 143 orthologous proteins**. Bootstrap values less than 100 are displayed at nodes. For a full list of taxa included see Supplementary Table B.2.

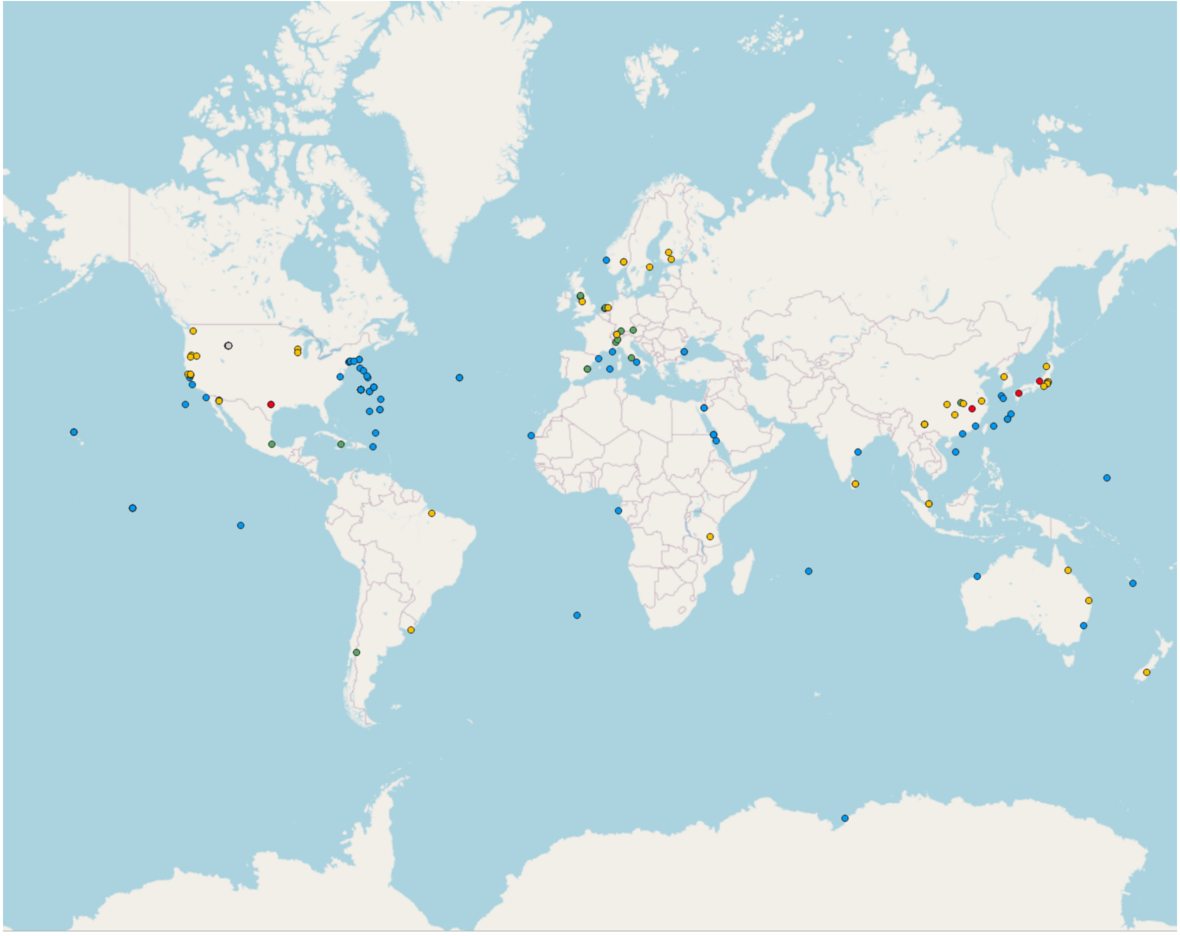

Figure A.8: **Isolation locations of genomes analysed in this study**. Green: Freshwater picocyanobacteria; Red: *Synechococcus elongatus* strains; Grey: Thermal picocyanobacteria; Blue: Marine picocyanobacteria; Orange: Larger freshwater cyanobacteria

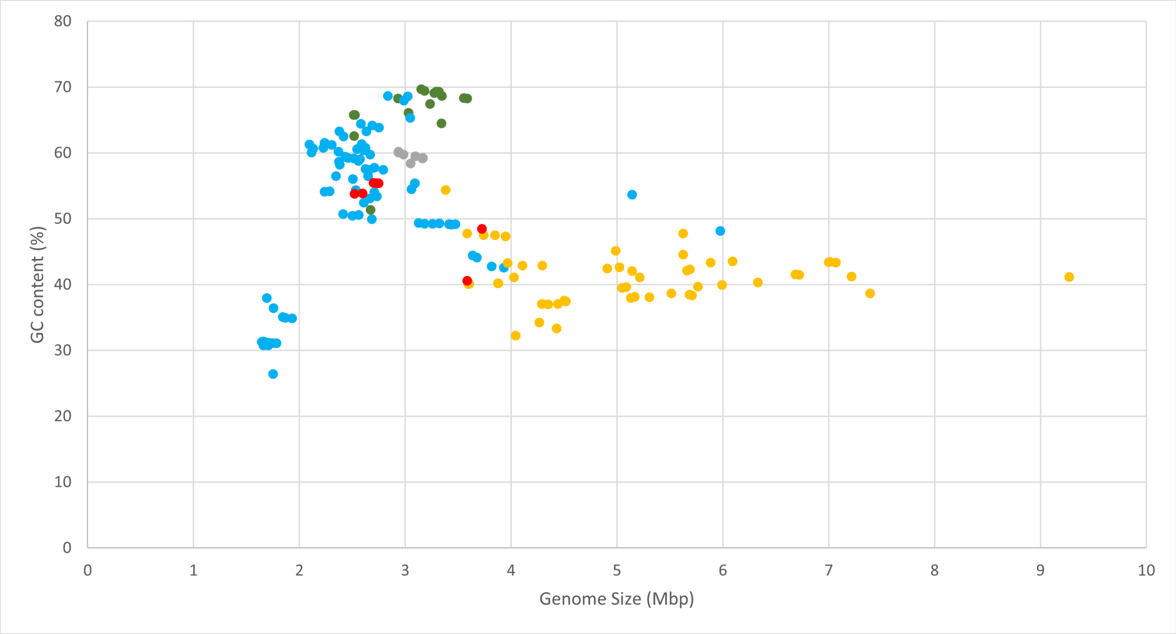

Figure A.9: **Genome size and GC content of genomes analysed**. Green: Freshwater picocyanobacteria; Red: *Synechococcus elongatus* strains; Grey: Thermal picocyanobacteria; Blue: Marine picocyanobacteria; Orange: Larger freshwater cyanobacteria

Figure A.10: **Growth curves of *Synechococcus* sp. CCY9618 and *Synechococcus* sp. CCAP1479/10 cultivated under different N sources at different concentrations.** A high concentration (250 mg N L^{-1} (25 mg N L^{-1} for tyrosine due to poor solubility))(CCY9618: green; CCAP1479/10: blue) and low concentration (1 mg N L^{-1})(CCY9618: orange; CCAP1479/10: purple) of substrate was investigated. The growth curve displays averaged measurements of OD_{750} over 14 days. Error bars represent the standard deviation from three biological replicates.

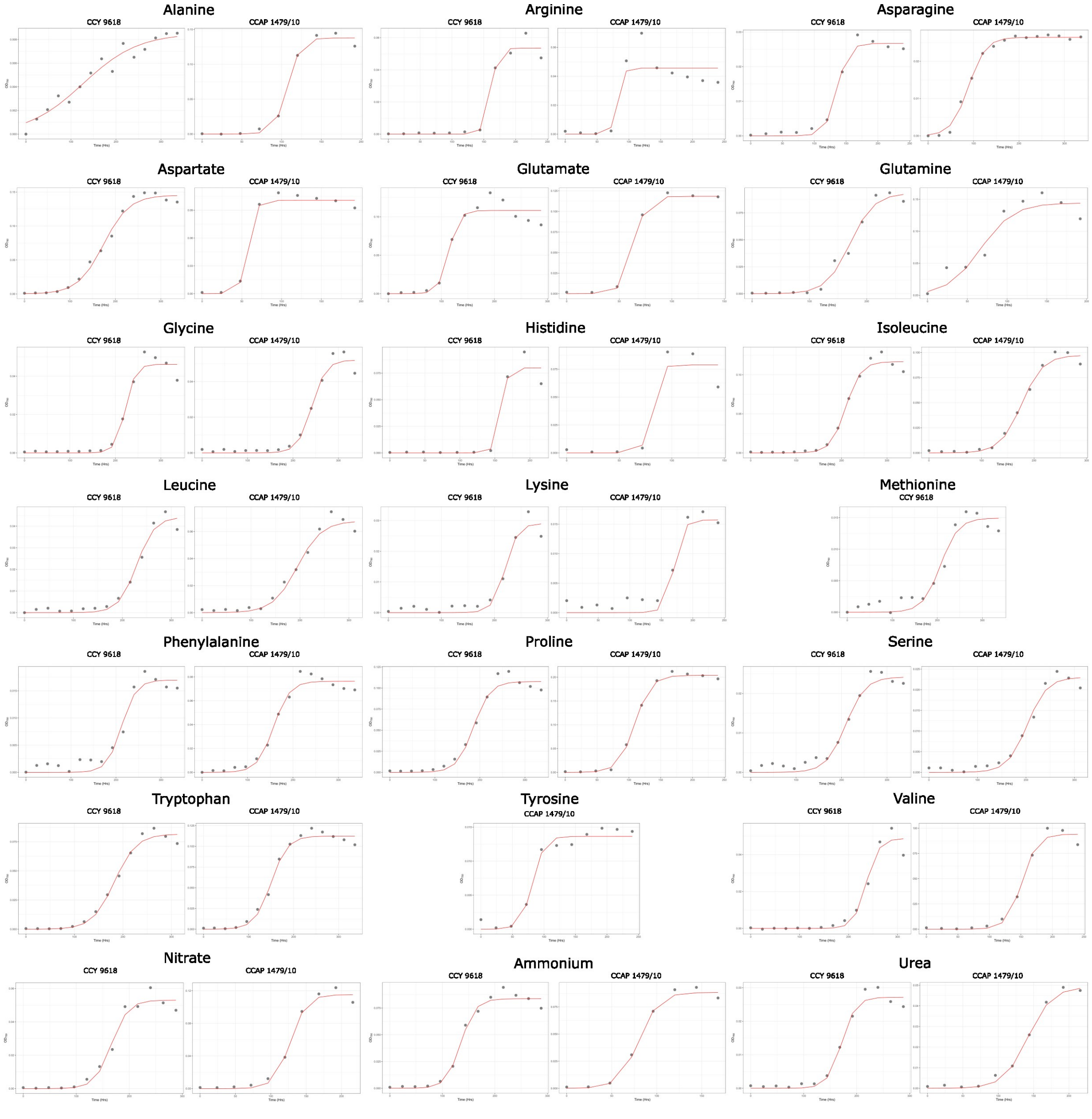

Figure A.11: **Growth rates of *Synechococcus* sp. CCY9618 and *Synechococcus* sp. CCAP1479/10 cultivated under different nitrogen sources at high concentration (250 mg N L^{-1}).** Note that the high concentration of tyrosine was reduced to 25 mg N L^{-1} due to poor solubility. Missing substrates are a result of no growth. Growth rates were calculated using Growthcurver v3.0.1 [480].

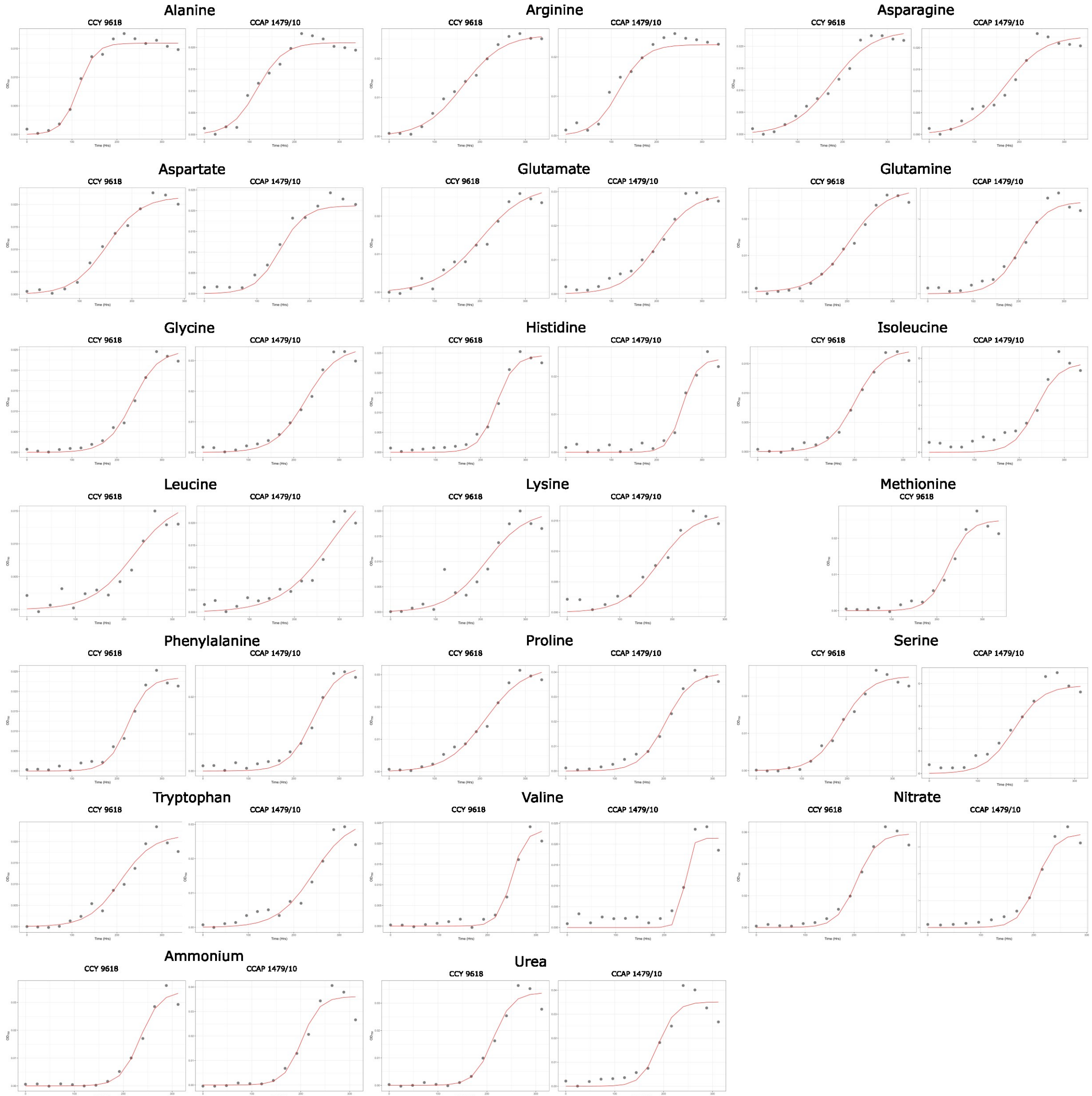

Figure A.12: **Growth rates of *Synechococcus* sp. CCY9618 and *Synechococcus* sp. CCAP1479/10 cultivated under different nitrogen sources at low concentration (1 mg N L^{-1}).** Missing substrates are a result of no growth. Growth rates were calculated using Growthcurver v3.0.1 [480].

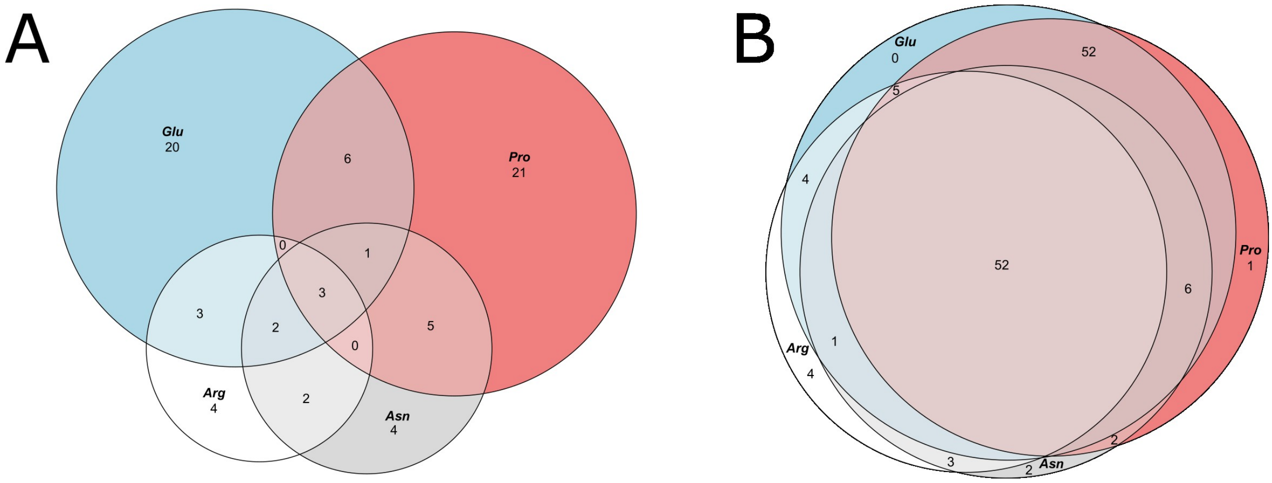

Figure A.13: **Area-proportional Euler diagrams for amino acid nitrogen-substrate differentially expressed protein overlaps compared to nitrogen-starvation.** A) Overlap of proteins up-regulated in *Synechococcus* sp. CCAP149/10 when grown under four AA conditions; B) Overlap of proteins down-regulated in *Synechococcus* sp. CCAP149/10 when grown under four AA conditions.

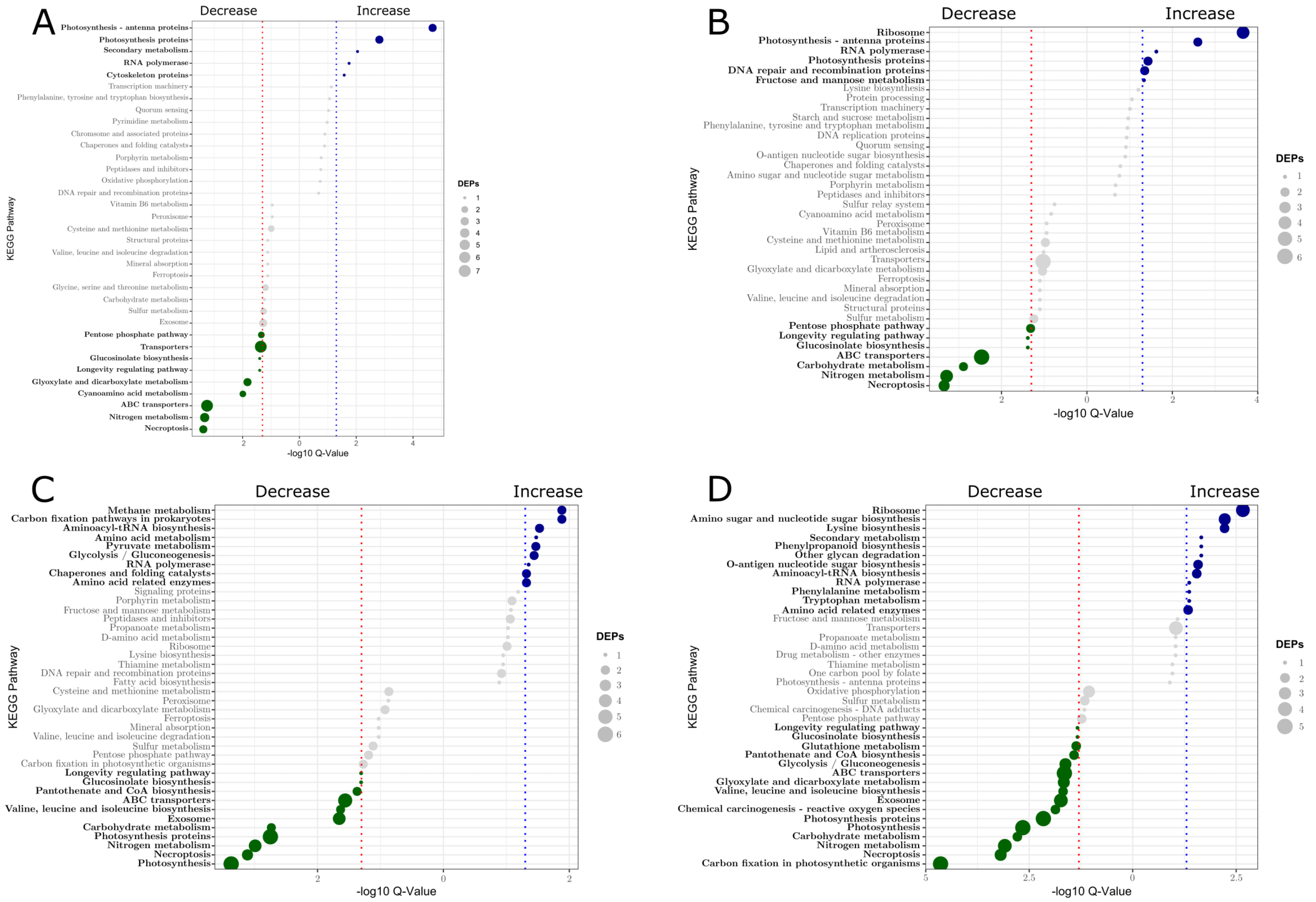

Figure A.14: **KEGG pathway enrichment analysis of growth on amino acid nitrogen-substrate v. nitrogen-starvation**. A) Arginine v. N-starvation; B) Asparagine v. N-starvation; C) Glutamate v. N-starvation; D) Proline v. N-starvation. Top 20 enriched pathways are shown with non-significant pathways in grey, significant pathways are in bold. Blue indicates up-regulated pathways while green indicates down-regulated pathways compared to N-starvation.

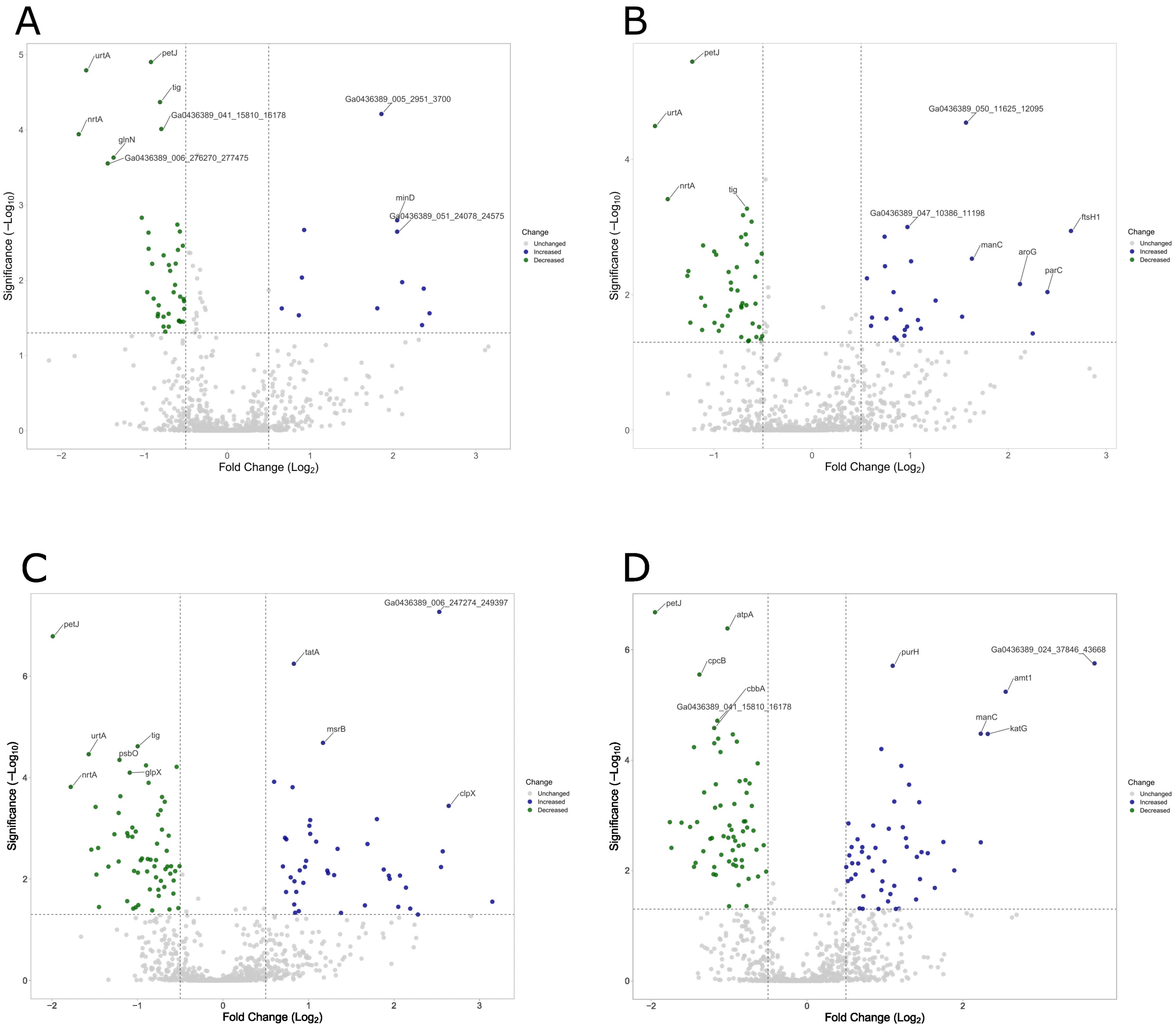

Figure A.15: **Volcano plots of the *Synechococcus* sp. CCAP1479/10 proteome when grown on amino acid nitrogen-substrate compared to NO_3^-.** A) Arginine v. NO_3^-; B) Asparagine v. NO_3^-; C) Glutamate v. NO_3^-; D) Proline v. NO_3^-. Blue values indicate up-regulation, green values indicate down-regulation, grey values indicate no significant difference. DEPs were deemed statistically significant with a q-value less than 0.05 and a $\log_2$ fold change greater than 0.5/less than -0.5. Top hits are listed in Supplementary Table B.22.

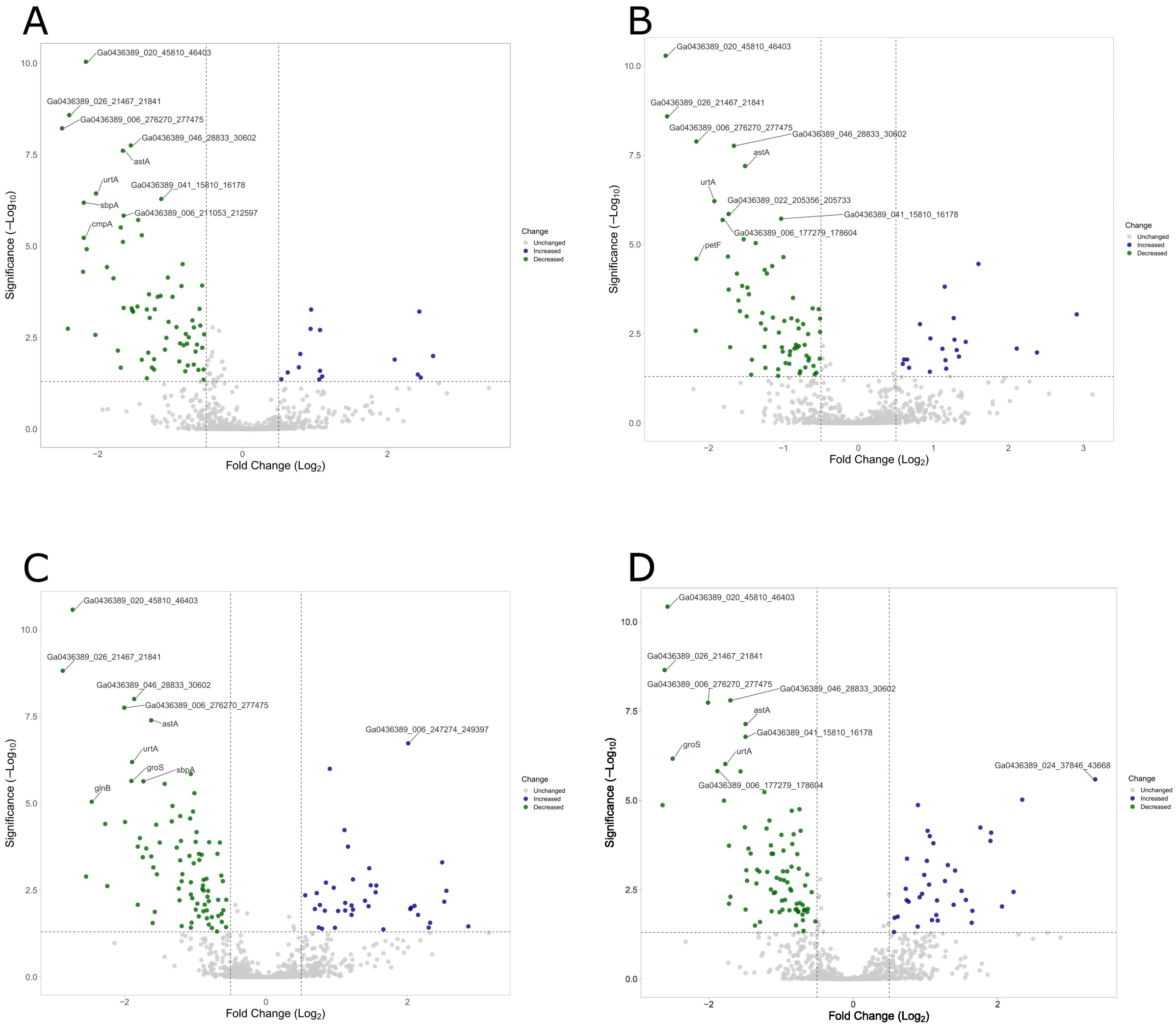

Figure A.16: **Volcano plots of the *Synechococcus* sp. CCAP1479/10 proteome when grown on AA N-substrate compared to N-starvation.** A) Arginine v. N-starvation; B) Asparagine v. N-starvation; C) Glutamate v. N-starvation; D) Proline v. N-starvation. Blue values indicate up-regulation, green values indicate down-regulation, grey values indicate no significant difference. DEPs were deemed statistically significant with a q-value less than 0.05 and a $\log_2$ fold change greater than 0.5/less than -0.5. Top hits are listed in Supplementary Table B.23.

Appendix B

Supplementary Tables

List of Tables in Appendix B

Table B.1: **Average nucleotide identity between newly sequenced strains**

	Synechococcus sp. CCAP1479/9	*Synechococcus* sp. CCAP1479/10	*Synechococcus* sp. CCAP1479/13	*Synechococcus* sp. CCY0621	*Synechococcus* sp. CCY9618
Synechococcus sp. CCAP1479/9		96.2739	96.2836	89.1877	85.8337
Synechococcus sp. CCAP1479/10	96.2737		99.9998	89.3104	85.9498
Synechococcus sp. CCAP1479/13	96.2834	99.9998		89.3368	85.9597
Synechococcus sp. CCY0621	89.1905	89.3125	89.3386		85.4348
Synechococcus sp. CCY9618	85.8327	85.9492	85.9593	85.4336	

Table B.2: **Cyanobacterial strains included in Figure 2.2 and Appendix Figure A.7**

Cyanobacterial Strains
*Acaryochloris_marina*_MBIC11017_GCF_000018105.1
*Acaryochloris*_sp._RCC1774_GCF_003231495.1
aff._Roholtiella_sp._LEGE_12411_GCA_015206995.1
*Alkalinema*_sp._FACHB-956_GCA_014697025.1
*Anabaena*_sp._FACHB-1237_GCA_014696225.1
*Anabaena*_sp__90_GCF_000312705
*Anabaena*_sp__FACHB-1391_GCF_014696755
*Anabaena*_sp__WA102_GCF_001277295
*Anabaenopsis_elenkinii*_CCIBt3563_GCA_015245355.1
*Aphanizomenon_flos-aquae*_FACHB-1287_GCF_014698755
*Aphanizomenon_flos-aquae*_LD13_GCA_001672165
*Aphanizomenon_flos-aquae*_NIES-81_GCF_000521175
*Aphanizomenon*_sp__FACHB-1416_GCF_014698695
*Aphanocapsa_feldmannii*_277cV_GCA_004768415.1
*Arthrospira_maxima*_CS-328_GCF_000173555.1
*Arthrospira_platensis*_C1_GCF_025200965
*Arthrospira_platensis*_FACHB-439_GCA_014698675.1
*Arthrospira_platensis*_FACHB-835_GCA_014698815.1
*Arthrospira_platensis*_FACHB-971_GCA_014698385.1
*Arthrospira_platensis*_NIES-39_GCF_000210375.1
*Arthrospira_platensis*_NIES-46_GCF_009176225.1
*Arthrospira_platensis*_str._Paraca_GCF_000175415.3
*Arthrospira_platensis*_YZ_GCF_001611905.1
*Arthrospira*_sp._O9.13F_GCF_003268325.1
*Calothrix_desertica*_PCC_7102_GCF_007830875.1
*Calothrix_parasitica*_NIES-267_GCF_002368095.1
*Calothrix*_sp._336_3_GCF_000734895.2
*Calothrix*_sp._NIES-3974_GCF_002368395.1
*Calothrix*_sp._NIES-4071_GCF_002368455.1
*Calothrix*_sp._NIES-4101_GCA_004296455.1
*Calothrix*_sp._PCC_6303_GCF_000317435.1
*Calothrix*_sp._PCC_7507_GCF_000316575.1
*Calothrix*_sp._SM1_7_51_GCA_012031285.1
*Candidatus*_Atelocyanobacterium_thalassa_isolate_ALOHA_GCF_000025125.1
*Candidatus*_Atelocyanobacterium_thalassa_isolate_SIO64986_GCF_000737945.1
*Candidatus*_Caenarcaniphilales_bacterium_GCA_013216135.1
*Candidatus*_Gastranaerophilales_bacterium_GCA_012515355.1
*Candidatus*_Gastranaerophilales_bacterium_HUM_10_GCA_002102745.1
*Candidatus*_Gastranaerophilales_bacterium_HUM_12_GCA_002102825.1
*Candidatus*_Gastranaerophilales_bacterium_HUM_15_GCA_002102815.1
*Candidatus*_Gastranaerophilales_bacterium_HUM_3_GCA_002102975.1
*Candidatus*_Gastranaerophilales_bacterium_HUM_5_GCA_002103085.1
*Candidatus*_Gastranaerophilus_sp._ex_Termes_propinquus_GCF_002933735.1
*Candidatus*_Melainabacteria_bacterium_GCA_003242885.1
*Candidatus*_Melainabacteria_bacterium_GCA_003258725.1

Table B.2 continued from previous page

Cyanobacterial Strains
*Candidatus*_Melainabacteria_bacterium_GCA_003258785.1
*Candidatus*_Melainabacteria_bacterium_GCA_003258795.1
*Candidatus*_Melainabacteria_bacterium_GCA_003963305.1
*Candidatus*_Melainabacteria_bacterium_GWA2_34_9_GCA_001784535.1
*Candidatus*_Melainabacteria_bacterium_GWF2_32_7_GCA_001786525.1
*Candidatus*_Melainabacteria_bacterium_GWF2_37_15_GCA_001784565.1
*Candidatus*_Melainabacteria_bacterium_MEL.A1_GCA_001765415.1
*Candidatus*_Melainabacteria_bacterium_RIFCSPLOWO2_02_FULL_35_15_GCA_001784585.1
*Candidatus*_Melainabacteria_bacterium_RIFOXYA2_FULL_32_9_GCA_001784615.1
*Candidatus*_Synechococcus_spongiarum_142_GCA_001007625.1
cf._*Phormidesmis*_sp._LEGE_11477_GCA_015207105.1
*Chamaesiphon_minutus*_PCC_6605_GCF_000317145.1
*Chlorogloea*_sp._CCALA_695_GCF_003003885.1
*Chlorogloeopsis_fritschii*_PCC_6912_GCF_003990575.1
*Chondrocystis*_sp._NIES-4102_GCF_002368355.1
*Chroococcidiopsis_cubana*_SAG_39.79_GCF_003991895.1
*Chroococcidiopsis*_sp._PCC_6712_GCF_012932785.1
*Coleofasciculus_chthonoplastes*_PCC_7420_GCF_000155555.1
*Coleofasciculus*_sp._LEGE_07092_GCA_015207375.1
*Crinalium_epipsammum*_PCC_9333_GCF_000317495.1
*Crocosphaera_chwakensis*_CCY0110_GCF_000169335.1
*Crocosphaera_subtropica*_ATCC_51142_GCF_000017845.1
*Crocosphaera_watsonii*_WH_0003_GCF_000235665.1
*Crocosphaera_watsonii*_WH_0005_GCF_001050835.1
*Crocosphaera_watsonii*_WH_0401_GCF_001039615.1
*Crocosphaera_watsonii*_WH_0402_GCF_001039635.1
*Crocosphaera_watsonii*_WH_8501_GCF_000167195.1
*Crocosphaera_watsonii*_WH_8502_GCF_001039555.1
Cyanobacteria_bacterium_CRU_2_1_GCA_012034815.1
Cyanobacteria_bacterium_FACHB-471_GCA_014695345.1
Cyanobacteria_bacterium_FACHB-502_GCA_014695325.1
Cyanobacteria_bacterium_GSL.Bin1_GCA_009909085.1
Cyanobacteria_bacterium_M5B4_GCA_002861535.1
Cyanobacteria_bacterium_PMG_004_GCA_004173145.1
Cyanobacteria_bacterium_QS_4_48_99_GCA_003021505.1
Cyanobacteria_bacterium_QS_8_64_29_GCA_003022125.1
Cyanobacteria_bacterium_UBA11162_GCA_003504865.1
Cyanobacterium_aponinum_0216_GCF_009710015.1
cyanobacterium_endosymbiont_of_Epithemia_turgida_isolate_EtSB_Lake_Yunoko_GCF_000829235.1
cyanobacterium_endosymbiont_of_Rhopalodia_gibberula_GCF_003574135.1
Cyanobacterium_stanieri_PCC_7202_GCF_000317655.1
*Cyanobium_gracile*_PCC_6307_GCF_000316515.1
*Cyanobium*_NIES-981_GCF_900088535
*Cyanobium*_sp._GCA_003249045.1
*Cyanobium*_sp._NS01_GCF_014280235.1
*Cyanobium*_sp__PCC_7001_GCF_000155635

Table B.2 continued from previous page

Cyanobacterial Strains
*Cyanobium_usitatum*_str._Tous_GCF_003011885.1
*Cyanothece*_sp._PCC_7425_GCF_000022045.1
*Cyanothece*_sp._SIO1E1_GCA_010672835.1
*Cylindrospermopsis_raciborskii*_CS-505_GCF_001676585
*Cylindrospermopsis_raciborskii*_ITEP-A1_GCF_001586755
*Cylindrospermopsis_raciborskii*_MVCC14_GCF_001858125
*Cylindrospermum*_sp._FACHB-282_GCA_014697195.1
*Cylindrospermum*_sp._NIES-4074_GCA_003994795.1
*Dactylococcopsis_salina*_PCC_8305_GCF_000317615.1
*Desertifilum*_sp._IPPAS_B-1220_GCF_001746915.1
*Dolichospermum_compactum*_NIES-806_GCF_002368115.1
*Dolichospermum*_sp__FACHB-1091_GCF_014697585
*Dolichospermum*_sp__UHCC_0352_GCF_009711925
*Euhalothece_natronophila*_Z-M001_GCF_007904085.1
*Euhalothece*_sp._KZN_001_GCA_002900215.1
filamentous_cyanobacterium_LEGE_07170_GCA_015207365.1
filamentous_cyanobacterium_LEGE_11480_GCA_015207035.1
*Fischerella*_sp._FACHB-380_GCA_014697535.1
*Fischerella*_sp._NIES-4106_GCF_002368315.1
*Geitlerinema*_sp._FC_II_GCA_002286845.1
*Geitlerinema*_sp._PCC_7407_GCF_000317045.1
*Geminocystis_herdmanii*_PCC_6308_GCF_000332235
*Geminocystis*_sp._NIES-3708_GCF_001548095.1
*Geminocystis*_sp__NIES-3709_GCF_001548115
*Gloeobacter_kilaueensis*_JS1_GCF_000484535.1
*Gloeobacter_violaceus*_PCC_7421_GCF_000011385.1
*Gloeocapsa*_sp._PCC_73106_GCF_000332035.1
*Gloeocapsa*_sp__PCC_7428_GCF_000317555
*Gloeocapsopsis*_sp._AAB1___1H9_GCF_009728615.1
*Gloeomargarita_lithophora*_Alchichica-D10_GCF_001870225.1
*Gloeothece_citriformis*_PCC_7424_GCF_000021825.1
*Halomicronema_hongdechloris*_C2206_GCF_002075285.3
*Halothece*_sp._PCC_7418_GCF_000317635.1
*Hormoscilla*_sp._SP12CHS1_GCA_014324065.1
*Hydrococcus*_sp._CRU_1_1_GCA_012034135.1
*Hydrococcus*_sp._CSU_1_8_GCA_012034605.1
*Hydrococcus*_sp._RU_2_2_GCA_012032735.1
*Iningainema*_sp._BLCCT55_GCA_014698965.1
*Kamptonema*_sp._PCC_6506_GCF_000180455.1
*Kamptonema*_sp._SIO1D9_GCA_010672865.1
*Komarekiella*_sp.__clone_1__GCA_014831675.1
*Leptolyngbya_antarctica*_GCA_003241845.1
*Leptolyngbya_boryana*_IAM_M-101_GCF_002142475.1
*Leptolyngbya*_sp.__hensonii__GCF_001939115.1
*Leptolyngbya*_sp._CCY15150_953e7925cd814f54
*Leptolyngbya*_sp._FACHB-16_GCA_014696095.1

Table B.2 continued from previous page

Cyanobacterial Strains
*Leptolyngbya*_sp._FACHB-261_GCA_014696065.1
*Leptolyngbya*_sp._FACHB-36_GCA_014695955.1
*Leptolyngbya*_sp._FACHB-671_GCA_014695745.1
*Leptolyngbya*_sp._GCA_003249025.1
*Leptolyngbya*_sp._IPPAS_B-1204_GCA_003724315.1
*Leptolyngbya*_sp._PCC_7375_GCF_000316115.1
*Leptolyngbya*_sp._PCC_7376_GCF_000316605.1
*Leptolyngbya*_sp._SIO1E4_GCA_010672825.2
*Leptolyngbya*_sp._SIO4C1_GCA_010671975.1
*Leptolyngbya*_sp._SIO4C5_GCA_010671965.1
*Leptolyngbya*_sp._SIOISBB_GCA_010671905.1
Leptolyngbyaceae_cyanobacterium_CRU_2_3_GCA_012034795.1
Leptolyngbyaceae_cyanobacterium_CSU_1_4_GCA_012034615.1
Leptolyngbyaceae_cyanobacterium_SM2_5_2_GCA_012031045.1
*Limnospira*_sp._BM_GCA_014250495.1
*Limnothrix_rosea*_IAM_M-220_GCF_001904615.1
*Limnothrix*_sp._FACHB-708_GCA_014698215.1
*Lusitaniella_coriacea*_LEGE_07157_GCA_015207425.1
*Lyngbya_aestuarii*_BL_J_GCF_000478195.2
*Lyngbya*_sp._CCY1209_a9a2ee4593ab2f78
*Mastigocoleus_testarum*_BC008_GCF_001456025.1
*Microcoleus*_sp._FACHB-1515_GCA_014696105.1
*Microcoleus*_sp._FACHB-68_GCA_014695715.1
*Microcoleus*_sp._FACHB-831_GCA_014695585.1
*Microcoleus*_sp._FACHB-SPT15_GCA_014849525.1
*Microcoleus*_sp._PCC_7113_GCF_000317515.1
*Microcystis_aeruginosa*_BLCCF158_GCF_014218765.1
*Microcystis_aeruginosa*_KLA2_GCF_008257445
*Microcystis_aeruginosa*_NIES-2481_GCF_001704955
*Microcystis_aeruginosa*_PCC_7005_GCF_000599945
*Microcystis_aeruginosa*_PCC_7806SL_GCF_002095975
*Microcystis_panniformis*_FACHB-1757_GCF_001264245
*Microcystis*_sp__0824_GCF_003112475
*Microcystis*_sp__MC19_GCF_003019735
*Moorea_producens*_PAL-8-15-08-1_GCF_001767235.1
*Moorea*_sp._SIO2B7_GCA_010672555.1
*Moorea*_sp._SIO3C2_GCA_010692465.1
*Nodosilinea*_sp._LEGE_06152_GCA_015207525.1
*Nodosilinea*_sp._LEGE_07088_GCA_015207395.1
*Nodosilinea*_sp._LEGE_07298_GCA_015207265.1
*Nodosilinea*_sp._P-1105_GCF_012911975.1
*Nostoc_paludosum*_FACHB-159_GCF_014698835
*Nostoc*_sp._CENA543_GCF_002896875.1
*Nostoc*_sp._PCC_7524_GCF_000316645.1
*Nostoc*_sp__PCC_7107_GCF_000316625
*Nostoc*_sp__PCC_7120_GCF_000009705

Table B.2 continued from previous page

Cyanobacterial Strains
*Nostoc_sphaeroides*_CCNUC1_GCA_009372195.1
*Nostoc_sphaeroides*_Kutzing_En_GCF_003443655
Nostocaceae_cyanobacterium_GCA_013821625.1
*Okeania*_sp._SIO1I7_GCA_010672585.1
*Oscillatoria_acuminata*_PCC_6304_GCF_000317105.1
*Oscillatoria_nigro-viridis*_PCC_7112_GCF_000317475.1
*Oscillatoria*_sp._FACHB-1406_GCA_014698145.1
*Oscillatoria*_sp._FACHB-1407_GCA_014697545.1
*Oscillatoria*_sp._SIO1A7_GCA_010672965.1
Oscillatoriales_cyanobacterium_C42_A2020_001_GCF_000309945.1
Oscillatoriales_cyanobacterium_LEGE_11467_GCA_015207075.1
Oscillatoriales_cyanobacterium_SM2_2_1_GCA_012031145.1
*Oxynema*_sp._AP17_GCF_012295525.1
*Phormidesmis_priestleyi*_ULC007_GCF_003003695.1
*Phormidium*_sp._CCY1219_a10f3efcff4a5351
*Phormidium*_sp._FACHB-592_GCA_014695795.1
*Phormidium*_sp._LEGE_05292_GCA_015207735.1
*Phormidium_tenue*_FACHB-886_GCA_014764505.1
*Phormidium_tenue*_NIES-30_GCF_001904775.1
*Planktothricoides*_sp._FACHB-1261_GCA_014698535.1
*Planktothricoides*_sp__FACHB-1370_GCF_014698175
*Planktothricoides*_sp_SR001_GCF_001276715
*Planktothrix_agardhii*_NIES-204_GCA_003609755
*Planktothrix_agardhii*_NIVA-CYA_1268_GCF_000710505
*Planktothrix_prolifica*_NIVA-CYA_406_GCF_000464765
*Planktothrix_rubescens*_NIVA-CYA_18_GCA_900009275.2
*Planktothrix_rubescens*_NIVA-CYA_407_GCF_000464785
*Planktothrix*_sp._FACHB-1375_GCA_014696265.1
*Planktothrix*_sp__PCC_11201_GCF_900009135
*Plectonema*_cf._radiosum_LEGE_06105_GCA_015207665.1
*Pleurocapsa*_sp._CRU_1_2_GCA_012033675.1
*Pleurocapsa*_sp._PCC_7327_GCF_000317025.1
*Pleurocapsa*_sp__PCC_7319_GCF_000332195
Pleurocapsales_cyanobacterium_LEGE_06147_GCA_015207585.1
Pleurocapsales_cyanobacterium_LEGE_10410_GCA_015207195.1
*Prochlorococcus_marinus*_NATL2A_GCF_000012465
*Prochlorococcus_marinus*_str._MIT_1323_GCF_001632025.1
*Prochlorococcus_marinus*_str._SB_GCF_000760115.1
*Prochlorococcus_marinus*_str__AS9601_GCF_000015645
*Prochlorococcus_marinus*_str__CCMP1986_GCF_000011465
*Prochlorococcus_marinus*_str__MIT_9211_GCF_000018585
*Prochlorococcus_marinus*_str__MIT_9215_GCF_000018065
*Prochlorococcus_marinus*_str__MIT_9301_GCF_000015965
*Prochlorococcus_marinus*_str__MIT_9302_GCF_000759975
*Prochlorococcus_marinus*_str__MIT_9303_GCF_000015705
*Prochlorococcus_marinus*_str__MIT_9312_GCF_000012645

Table B.2 continued from previous page

Cyanobacterial Strains
*Prochlorococcus_marinus*_str__MIT_9313_GCF_000011485
*Prochlorococcus_marinus*_str__MIT_9515_GCF_000015665
*Prochlorococcus_marinus*_str__NATL1A_GCF_000015685
*Prochlorococcus*_MIT_0604_GCF_000757845
*Prochlorococcus*_RS50_GCF_001989415
*Prochlorococcus*_sp__MIT_0801_GCF_000757865
*Prochlorococcus*_sp__MIT_1306_GCF_001631985
*Prochlorococcus*_sp__P1363_GCF_012933595
*Prochlorococcus*_sp__RS01_GCF_001989435
*Prochlorococcus*_sp__RS04_GCF_001989455
*Prochlorococcus*_sp__SS52_GCF_000760375
*Prochloron*_sp._SP5CPC1_GCA_014323965.1
*Prochlorothrix_hollandica*_PCC_9006___CALU_1027_GCF_000341585.2
*Pseudanabaena*_sp._ABRG5-3_GCF_003967015.1
*Pseudanabaena*_sp._FACHB-2040_GCA_014696715.1
*Pseudanabaena*_sp._PCC_7367_GCF_000317065.1
*Pseudanabaena*_sp__FACHB-723_GCF_014696385
*Pseudanabaena*_sp__PCC_6802_GCF_000332175
*Richelia_intracellularis*_GCF_000613065.1
*Richelia_intracellularis*_HH01_GCF_000350105.1
*Richelia_intracellularis*_HM01_GCF_000350125.1
*Richelia_sinica*_FACHB-800_GCA_014698825.1
*Richelia*_sp._SL_2_1_GCA_012033205.1
*Rippkaea_orientalis*_PCC_8801_GCF_000021805.1
*Romeria*_aff._gracilis_LEGE_07310_GCA_015207255.1
*Roseofilum_reptotaenium*_AO1-A_GCA_001890975.1
*Rubidibacter_lacunae*_KORDI_51-2_GCF_000473895.1
*Scytonema_hofmannii*_PCC_7110_GCF_000346485.2
*Scytonema_tolypothrichoides*_VB-61278_GCF_000828085.3
*Snowella*_sp._GCA_003249095.1
*Sphaerospermopsis*_sp._LEGE_00249_GCF_014299995.1
*Spirulina*_sp._CCY15215_84eac1c87f91d783
*Spirulina*_sp._SIO3F2_GCA_010672165.1
*Stanieria*_sp._NIES-3757_GCF_002355455.1
*Symploca*_sp._SIO1C2_GCA_010672925.1
*Symploca*_sp._SIO2B6_GCA_010672535.1
*Symploca*_sp._SIO2E9_GCA_010692645.1
*Symploca*_sp._SIO3E6_GCA_010672185.1
Synechococcaceae_cyanobacterium_SM2_3_2_GCA_012031565.1
Synechococcales_cyanobacterium_RM1_1_8_GCA_012032745.1
*Synechococcus*_Alchichica_8F6_GCF_002252665
*Synechococcus*_Atexcac_6H9_Ga0078145
*Synechococcus*_CC9311_GCF_000014585
*Synechococcus*_CC9605_GCF_000012625
*Synechococcus*_CC9902_GCF_000012505
*Synechococcus*_Constance_BO8801_GCF_002252675

Table B.2 continued from previous page

Cyanobacterial Strains
*Synechococcus_elongatus*_FACHB-1061_GCF_014698895
*Synechococcus_elongatus*_FACHB-242_GCF_014698415
*Synechococcus_elongatus*_PCC_7942___FACHB-805_GCF_000012525.1
*Synechococcus*_Huapia_1G10_GCF_002252625
*Synechococcus*_JA-3-3Ab_GCF_000013205
*Synechococcus*_KORDI-49_GCF_000737575
*Synechococcus*_KORDI-52_GCF_000737595
*Synechococcus_lacustris*_Tous_GCF_003011125
*Synechococcus*_Maggiore_4A10_Ga0078148
*Synechococcus*_Mondsee_MW101C3_GCF_002252635
*Synechococcus*_PCC_6301_GCF_022984195
*Synechococcus*_PCC_7003_GCF_001693255
*Synechococcus*_PCC_8807_GCF_001693295
*Synechococcus*_sp._7002_GCF_900177825.1
*Synechococcus*_sp._A18-25c_GCF_014280035.1
*Synechococcus*_sp._BIOS-E4-1_GCF_014279995.1
*Synechococcus*_sp._BMK-MC-1_GCF_014279955.1
*Synechococcus*_sp._CB0101_GCF_000179235.2
*Synechococcus*_sp._GFB01_GCF_001039265.1
*Synechococcus*_sp._JA-2-3B_a2-13_GCF_000013225.1
*Synechococcus*_sp._KORDI-100_GCF_000737535.1
*Synechococcus*_sp._LTW-R_GCF_014217875.1
*Synechococcus*_sp._MEDNS5_GCF_014279875.1
*Synechococcus*_sp._NIES-970_GCF_002356215.1
*Synechococcus*_sp._PCC_6312_GCF_000316685.1
*Synechococcus*_sp._PCC_7117_GCF_001693275.1
*Synechococcus*_sp._PCC_73109_GCF_001521855.1
*Synechococcus*_sp._PCC_7502_GCF_000317085.1
*Synechococcus*_sp._PROS-7-1_GCF_014279795.1
*Synechococcus*_sp._RCC307_GCF_000063525.1
*Synechococcus*_sp._RS9902_GCF_014279635.1
*Synechococcus*_sp._RS9916_GCF_000153825.1
*Synechococcus*_sp._RSCCF101_GCF_008807075.1
*Synechococcus*_sp._WH_7805_GCF_000153285.1
*Synechococcus*_sp._WH_8020_GCF_001040845.1
*Synechococcus*_sp__60AY4M2_GCF_002760375
*Synechococcus*_sp__63AY4M1_GCF_002760395
*Synechococcus*_sp__63AY4M2_GCF_002760475
*Synechococcus*_sp__65AY640_GCF_002760445
*Synechococcus*_sp__65AY6A5_GCF_002760415
*Synechococcus*_sp__65AY6Li_GCF_002760345
*Synechococcus*_sp__BDU_130192_GCF_002721235
*Synechococcus*_sp__BL107_GCF_000153805
*Synechococcus*_sp__BS55D_GCF_004332415
*Synechococcus*_sp__BS56D_GCF_004332405
*Synechococcus*_sp__CC9616_GCF_000515235

Table B.2 continued from previous page

Cyanobacterial Strains
*Synechococcus*_sp__FACHB-909_GCF_014698945
*Synechococcus*_sp__GEYO_GCF_900473955
*Synechococcus*_sp__HB1133_GCF_011365345
*Synechococcus*_sp__MIT_S9504_GCF_001632105
*Synechococcus*_sp__MIT_S9508_GCF_001632165
*Synechococcus*_sp__MIT_S9509_GCF_001631935
*Synechococcus*_sp__N26_GCF_900473975
*Synechococcus*_sp__NKBG042902_GCF_000715475
*Synechococcus*_sp__NKBG15041c_GCF_000485815
*Synechococcus*_sp__OG1_GCF_900177365
*Synechococcus*_sp__PCC_11901_GCF_005577135
*Synechococcus*_sp__PCC_7335_GCF_000155595
*Synechococcus*_sp__PCC_7336_GCF_000332275
*Synechococcus*_sp__RS9917_GCF_000153065
*Synechococcus*_sp__UW105_GCF_900473935
*Synechococcus*_sp__UW140_GCF_900474295
*Synechococcus*_sp__UW179A_GCF_900473965
*Synechococcus*_sp__UW69_GCF_900474185
*Synechococcus*_sp__UW86_GCF_900474085
*Synechococcus*_sp__WH_5701_GCF_000153045
*Synechococcus*_sp__WH_8016_GCF_000230675
*Synechococcus*_sp__XM-24_GCA_003149585
*Synechococcus*_sp__YX04-3_GCA_003712165
*Synechococcus*_sp_A15-44_GCF_014280115
*Synechococcus*_sp_CCAP_1479_10_GCF_019038515
*Synechococcus*_sp_CCAP_1479_13_GCF_019038475
*Synechococcus*_sp_CCAP_1479_9_GCF_019038535
*Synechococcus*_sp_CCY_0621_GCF_019038465
*Synechococcus*_sp_CCY_9618_GCF_019038415
*Synechococcus*_sp_NOUM97013_GCF_014279815
*Synechococcus*_sp_SYN20_GCF_014279555
*Synechococcus*_SynAce01_GCF_001885215
*Synechococcus*_UTEX_2973_GCF_000817325
*Synechococcus*_WH_7803_GCF_000063505
*Synechococcus*_WH_8101_GCF_004209775
*Synechococcus*_WH_8102_GCF_000195975
*Synechococcus*_WH_8103_GCF_001182765
*Synechococcus*_WH_8109_GCF_000161795
*Synechocystis*_sp._FACHB-383_GCA_014698605.1
*Synechocystis*_sp__FACHB-929_GCF_014698625
*Synechocystis*_sp__PCC_6714_GCF_000478825
*Synechocystis*_sp__PCC_6803_GCF_000009725
*Synechocystis*_sp_CACIAM_05_GCF_010222815
*Thermoleptolyngbya*_sp._PKUAC-SCTA183_GCF_013177315.1
*Thermosynechococcus_elongatus*_BP-1_GCF_000011345
*Thermosynechococcus*_sp__NK55a_GCF_000505665

Table B.2 continued from previous page

Cyanobacterial Strains
*Thermosynechococcus_vulcanus*_NIES-2134_GCF_003990665.1
*Tolypothrix*_sp._NIES-4075_GCF_002218085.1
*Trichocoleus*_sp._FACHB-262_GCA_014696035.1
*Vulcanococcus_limneticus*_Candia_3B3_GCF_024346195
*Vulcanococcus_limneticus*_LL_GCF_002252705
*Xenococcus*_sp._PCC_7305_GCF_000332055.1

Table B.3: **Statistical analysis of COG functional annotation of the newly sequenced strains as total gene numbers.** Student's T-test compares the number of genes in the newly sequenced strains to the genes found in *Synechococcus elongatus* strains.

COG	***Synechococcus* sp. CCAP1479/9**	***Synechococcus* sp. CCAP1479/10**	***Synechococcus* sp. CCAP1479/13**	***Synechococcus* sp. CCY0621**	***Synechococcus* sp. CCY9618**	***Synechococcus elongatus* PCC 7942**	***Synechococcus elongatus* UTEX 2973**	***Synechococcus elongatus* PCC 6301**	***Synechococcus elongatus* FACHB-242**	***Synechococcus elongatus* FACHB-1061**	**T-test p-value**
J	163	161	161	161	160	163	166	165	167	167	0.001165
K	161	155	154	138	116	114	115	110	114	114	0.004995
L	116	121	122	127	146	108	116	112	114	114	0.035265
B	2	2	2	2	2	2	2	2	2	2	-
D	35	42	42	37	28	27	27	25	27	27	0.004655
V	41	43	43	51	47	31	31	31	31	31	0.000051
T	92	100	98	89	62	85	117	109	117	117	0.054321
M	199	201	203	204	186	134	148	145	148	148	0.000001
N	19	21	20	17	18	26	26	26	26	25	0.000015
U	71	73	75	71	62	34	55	53	55	54	0.002424
O	111	114	115	114	108	94	99	101	100	100	0.000063
C	220	229	230	220	224	199	199	202	198	199	0.000004
G	112	111	113	104	103	71	82	80	82	82	0.000010
E	181	184	181	174	163	138	137	132	138	138	0.000008
F	86	89	89	88	83	98	97	96	98	98	0.000026
H	186	183	185	185	181	167	171	171	173	173	0.000016
I	79	80	79	88	70	54	54	51	54	54	0.000021
P	142	151	151	162	122	152	163	157	163	164	0.079997
Q	46	45	45	46	39	27	44	42	44	44	0.295648
S	669	711	710	704	629	572	737	715	745	742	0.643511

Table B.4: **COG functional annotation of the newly sequenced strains as a percentage of the genome.** Student's T-test compares the number of genes in the newly sequenced strains to the genes found in *Synechococcus elongatus* strains.

COG	***Synechococcus* sp. CCAP1479/9**	***Synechococcus* sp. CCA1479/10**	***Synechococcus* sp. CCAP1479/13**	***Synechococcus* sp. CCY0621**	***Synechococcus* sp. CCY9618**	***Synechococcus elongatus* PCC 7942**	***Synechococcus elongatus* UTEX 2973**	***Synechococcus elongatus* PCC 6301**	***Synechococcus elongatus* FACHB-242**	***Synechococcus elongatus* FACHB-1061**	**T-test p-value**
J	0.048454	0.046789	0.046721	0.047256	0.051463	0.061232	0.060962	0.065295	0.060354	0.060332	5.93E-06
K	0.04786	0.045045	0.044689	0.040505	0.037311	0.042825	0.042233	0.04353	0.0412	0.041185	0.655504
L	0.034483	0.035164	0.035403	0.037276	0.04696	0.040571	0.0426	0.044321	0.0412	0.041185	0.126993
B	0.000595	0.000581	0.00058	0.000587	0.000643	0.000751	0.000734	0.000791	0.000723	0.000723	2.95E-05
D	0.010404	0.012206	0.012188	0.01086	0.009006	0.010143	0.009916	0.009893	0.009758	0.009754	0.123295
V	0.012188	0.012496	0.012478	0.014969	0.015117	0.011645	0.011385	0.012268	0.011203	0.011199	0.023331
T	0.027348	0.029061	0.028439	0.026123	0.019942	0.031931	0.042967	0.043134	0.042284	0.042269	0.000731
M	0.059156	0.058413	0.058909	0.059877	0.059826	0.050338	0.054352	0.05738	0.053488	0.053468	0.00157
N	0.005648	0.006103	0.005804	0.00499	0.00579	0.009767	0.009548	0.010289	0.009396	0.009032	6.08E-07
U	0.021106	0.021215	0.021764	0.020839	0.019942	0.012772	0.020198	0.020973	0.019877	0.019509	0.168149
O	0.032996	0.03313	0.033372	0.033461	0.034738	0.035312	0.036357	0.039968	0.03614	0.036127	0.005957
C	0.065398	0.06655	0.066744	0.064573	0.072049	0.074756	0.073081	0.079937	0.071558	0.071893	0.007298
G	0.033294	0.032258	0.032792	0.030525	0.03313	0.026672	0.030114	0.031658	0.029635	0.029624	0.016883
E	0.053805	0.053473	0.052525	0.051071	0.052428	0.051841	0.050312	0.052236	0.049874	0.049855	0.02981
F	0.025565	0.025865	0.025827	0.025829	0.026697	0.036814	0.035622	0.03799	0.035417	0.035405	6.17E-08
H	0.055291	0.053182	0.053685	0.0543	0.058218	0.062735	0.062798	0.067669	0.062523	0.0625	0.000194
I	0.023484	0.023249	0.022925	0.025829	0.022515	0.020285	0.019831	0.020182	0.019516	0.019509	0.00026
P	0.042212	0.043883	0.043819	0.047549	0.039241	0.0571	0.05986	0.062129	0.058909	0.059249	7.12E-06
Q	0.013674	0.013078	0.013059	0.013502	0.012544	0.010143	0.016159	0.01662	0.015902	0.015896	0.185412
S	0.19887	0.206626	0.206036	0.206633	0.202316	0.214876	0.270657	0.282944	0.269245	0.268064	0.001416

Table B.5: **List of differentially identified KEGG functional annotations.**

KEGG Function ID	**Function Name**	***Synechococcus* or *elongatus***
KO:K00002	alcohol dehydrogenase (NADP+) [EC:1.1.1.2]	*Synechococcus*
KO:K00016	L-lactate dehydrogenase [EC:1.1.1.27]	*Synechococcus*
KO:K00116	malate dehydrogenase (quinone) [EC:1.1.5.4]	*Synechococcus*
KO:K00227	Delta7-sterol 5-desaturase [EC:1.14.19.20]	*Synechococcus*
KO:K00274	monoamine oxidase [EC:1.4.3.4]	*elongatus*
KO:K00384	thioredoxin reductase (NADPH) [EC:1.8.1.9]	*elongatus*
KO:K00404	cytochrome c oxidase cbb3-type subunit I [EC:7.1.1.9]	*elongatus*
KO:K00425	cytochrome bd ubiquinol oxidase subunit I [EC:7.1.1.7]	*elongatus*
KO:K00426	cytochrome bd ubiquinol oxidase subunit II [EC:7.1.1.7]	*elongatus*
KO:K00477	phytanoyl-CoA hydroxylase [EC:1.14.11.18]	*Synechococcus*
KO:K00508	linoleoyl-CoA desaturase [EC:1.14.19.3]	*Synechococcus*
KO:K00537	arsenate reductase (glutaredoxin) [EC:1.20.4.1]	*Synechococcus*
KO:K00556	tRNA (guanosine-2'-O-)-methyltransferase [EC:2.1.1.34]	*Synechococcus*
KO:K00568	2-polyprenyl-6-hydroxyphenyl methylase / 3-demethylubiquinone-9 3-methyltransferase [EC:2.1.1.222 2.1.1.64]	*Synechococcus*
KO:K00588	caffeoyl-CoA O-methyltransferase [EC:2.1.1.104]	*elongatus*
KO:K00598	trans-aconitate 2-methyltransferase [EC:2.1.1.144]	*Synechococcus*
KO:K00612	carbamoyltransferase [EC:2.1.3.-]	*Synechococcus*
KO:K00651	homoserine O-succinyltransferase/O-acetyltransferase [EC:2.3.1.46 2.3.1.31]	*elongatus*
KO:K00657	diamine N-acetyltransferase [EC:2.3.1.57]	*Synechococcus*
KO:K00684	leucyl/phenylalanyl-tRNA—protein transferase [EC:2.3.2.6]	*elongatus*
KO:K00754	L-malate glycosyltransferase [EC:2.4.1.-]	*Synechococcus*
KO:K00835	valine–pyruvate aminotransferase [EC:2.6.1.66]	*elongatus*
KO:K00852	ribokinase [EC:2.7.1.15]	*Synechococcus*
KO:K00941	hydroxymethylpyrimidine/phosphomethylpyrimidine kinase [EC:2.7.1.49 2.7.4.7]	*elongatus*
KO:K00970	poly(A) polymerase [EC:2.7.7.19]	*Synechococcus*
KO:K00979	3-deoxy-manno-octulosonate cytidylyltransferase (CMP-KDO synthetase) [EC:2.7.7.38]	*elongatus*
KO:K01008	selenide, water dikinase [EC:2.7.9.3]	*Synechococcus*
KO:K01046	triacylglycerol lipase [EC:3.1.1.3]	*Synechococcus*
KO:K01090	protein phosphatase [EC:3.1.3.16]	*elongatus*
KO:K01126	glycerophosphoryl diester phosphodiesterase [EC:3.1.4.46]	*Synechococcus*
KO:K01144	exodeoxyribonuclease V [EC:3.1.11.5]	*Synechococcus*
KO:K01153	type I restriction enzyme, R subunit [EC:3.1.21.3]	*elongatus*
KO:K01154	type I restriction enzyme, S subunit [EC:3.1.21.3]	*elongatus*
KO:K01174	micrococcal nuclease [EC:3.1.31.1]	*elongatus*
KO:K01176	alpha-amylase [EC:3.2.1.1]	*Synechococcus*
KO:K01179	endoglucanase [EC:3.2.1.4]	*Synechococcus*
KO:K01185	lysozyme [EC:3.2.1.17]	*elongatus*
KO:K01299	carboxypeptidase Taq [EC:3.4.17.19]	*Synechococcus*
KO:K01356	repressor LexA [EC:3.4.21.88]	*Synechococcus*
KO:K01406	serralysin [EC:3.4.24.40]	*Synechococcus*
KO:K01423	beta-barrel assembly-enhancing protease [EC:3.4.-.-]	*elongatus*
KO:K01425	glutaminase [EC:3.5.1.2]	*Synechococcus*
KO:K01426	amidase [EC:3.5.1.4]	*Synechococcus*
KO:K01428	urease subunit alpha [EC:3.5.1.5]	*Synechococcus*

Table B.5 continued from previous page

KEGG Function ID	Function Name	*Synechococcus* or *elongatus*
KO:K01429	urease subunit beta [EC:3.5.1.5]	*Synechococcus*
KO:K01430	urease subunit gamma [EC:3.5.1.5]	*Synechococcus*
KO:K01437	aspartoacylase [EC:3.5.1.15]	*Synechococcus*
KO:K01455	formamidase [EC:3.5.1.49]	*Synechococcus*
KO:K01480	agmatinase [EC:3.5.3.11]	*Synechococcus*
KO:K01488	adenosine deaminase [EC:3.5.4.4]	*elongatus*
KO:K01489	cytidine deaminase [EC:3.5.4.5]	*elongatus*
KO:K01506	5-aminopentanamidase [EC:3.5.1.30]	*Synechococcus*
KO:K01512	acylphosphatase [EC:3.6.1.7]	*Synechococcus*
KO:K01534	Zn2+/Cd2+-exporting ATPase [EC:7.2.2.12 7.2.2.21]	*Synechococcus*
KO:K01551	arsenite/tail-anchored protein-transporting ATPase [EC:7.3.2.7 7.3.-.-]	*elongatus*
KO:K01569	oxalate decarboxylase [EC:4.1.1.2]	*elongatus*
KO:K01580	glutamate decarboxylase [EC:4.1.1.15]	*Synechococcus*
KO:K01582	lysine decarboxylase [EC:4.1.1.18]	*Synechococcus*
KO:K01623	fructose-bisphosphate aldolase, class I [EC:4.1.2.13]	*Synechococcus*
KO:K01627	2-dehydro-3-deoxyphosphooctonate aldolase (KDO 8-P synthase) [EC:2.5.1.55]	*elongatus*
KO:K01658	anthranilate synthase component II [EC:4.1.3.27]	*elongatus*
KO:K01664	para-aminobenzoate synthetase component II [EC:2.6.1.85]	*Synechococcus*
KO:K01674	carbonic anhydrase [EC:4.2.1.1]	*elongatus*
KO:K01754	threonine dehydratase [EC:4.3.1.19]	*Synechococcus*
KO:K01792	glucose-6-phosphate 1-epimerase [EC:5.1.3.15]	*elongatus*
KO:K01823	isopentenyl-diphosphate Delta-isomerase [EC:5.3.3.2]	*elongatus*
KO:K01854	UDP-galactopyranose mutase [EC:5.4.99.9]	*Synechococcus*
KO:K01893	asparaginyl-tRNA synthetase [EC:6.1.1.22]	*elongatus*
KO:K01894	glutamyl-Q tRNA(Asp) synthetase [EC:6.1.1.-]	*elongatus*
KO:K01949	glutamate—methylamine ligase [EC:6.3.4.12]	*elongatus*
KO:K01953	asparagine synthase (glutamine-hydrolysing) [EC:6.3.5.4]	*Synechococcus*
KO:K01989	putative tryptophan/tyrosine transport system substrate-binding protein	*Synechococcus*
KO:K01991	polysaccharide biosynthesis/export protein	*Synechococcus*
KO:K02013	iron complex transport system ATP-binding protein [EC:7.2.2.-]	*Synechococcus*
KO:K02018	molybdate transport system permease protein	*elongatus*
KO:K02022	HlyD family secretion protein	*Synechococcus*
KO:K02032	peptide/nickel transport system ATP-binding protein	*elongatus*
KO:K02039	phosphate transport system protein	*elongatus*
KO:K02041	phosphonate transport system ATP-binding protein [EC:7.3.2.2]	*Synechococcus*
KO:K02042	phosphonate transport system permease protein	*Synechococcus*
KO:K02044	phosphonate transport system substrate-binding protein	*Synechococcus*
KO:K02048	sulfate/thiosulfate transport system substrate-binding protein	*elongatus*
KO:K02068	putative ABC transport system ATP-binding protein	*Synechococcus*
KO:K02069	putative ABC transport system permease protein	*Synechococcus*
KO:K02189	cobalt-precorrin 5A hydrolase [EC:3.7.1.12]	*elongatus*
KO:K02217	ferritin [EC:1.16.3.2]	*Synechococcus*
KO:K02286	phycocyanin-associated rod linker protein	*elongatus*
KO:K02426	cysteine desulfuration protein SufE	*Synechococcus*
KO:K02479	two-component system, NarL family, response regulator	*Synechococcus*
KO:K02480	two-component system, NarL family, sensor kinase [EC:2.7.13.3]	*Synechococcus*

Table B.5 continued from previous page

KEGG Function ID	Function Name	*Synechococcus* or *elongatus*
KO:K02484	two-component system, OmpR family, sensor kinase [EC:2.7.13.3]	*Synechococcus*
KO:K02485	two-component system, response regulator	*elongatus*
KO:K02495	oxygen-independent coproporphyrinogen III oxidase [EC:1.3.98.3]	*Synechococcus*
KO:K02638	plastocyanin	*elongatus*
KO:K02642	cytochrome b6-f complex subunit 6	*elongatus*
KO:K02658	twitching motility two-component system response regulator PilH	*elongatus*
KO:K02660	twitching motility protein PilJ	*elongatus*
KO:K02662	type IV pilus assembly protein PilM	*elongatus*
KO:K02706	photosystem II P680 reaction center D2 protein [EC:1.10.3.9]	*elongatus*
KO:K03087	RNA polymerase nonessential primary-like sigma factor	*elongatus*
KO:K03187	urease accessory protein	*Synechococcus*
KO:K03188	urease accessory protein	*Synechococcus*
KO:K03189	urease accessory protein	*Synechococcus*
KO:K03190	urease accessory protein	*Synechococcus*
KO:K03270	3-deoxy-D-manno-octulosonate 8-phosphate phosphatase (KDO 8-P phosphatase) [EC:3.1.3.45]	*elongatus*
KO:K03281	chloride channel protein, CIC family	*elongatus*
KO:K03287	outer membrane factor, OMF family	*Synechococcus*
KO:K03294	basic amino acid/polyamine antiporter, APA family	*Synechococcus*
KO:K03298	drug/metabolite transporter, DME family	*Synechococcus*
KO:K03306	inorganic phosphate transporter, PiT family	*elongatus*
KO:K03307	solute:Na+ symporter, SSS family	*elongatus*
KO:K03312	glutamate:Na+ symporter, ESS family	*Synechococcus*
KO:K03322	manganese transport protein	*Synechococcus*
KO:K03327	multidrug resistance protein, MATE family	*elongatus*
KO:K03395	aminoglycoside 3-N-acetyltransferase I [EC:2.3.1.60]	*Synechococcus*
KO:K03408	purine-binding chemotaxis protein CheW	*elongatus*
KO:K03427	type I restriction enzyme M protein [EC:2.1.1.72]	*elongatus*
KO:K03442	small conductance mechanosensitive channel	*Synechococcus*
KO:K03559	biopolymer transport protein ExbD	*elongatus*
KO:K03572	DNA mismatch repair protein MutL	*elongatus*
KO:K03574	8-oxo-dGTP diphosphatase [EC:3.6.1.55]	*elongatus*
KO:K03579	ATP-dependent helicase HrpB [EC:3.6.4.13]	*Synechococcus*
KO:K03581	exodeoxyribonuclease V alpha subunit [EC:3.1.11.5]	*Synechococcus*
KO:K03582	exodeoxyribonuclease V beta subunit [EC:3.1.11.5]	*Synechococcus*
KO:K03583	exodeoxyribonuclease V gamma subunit [EC:3.1.11.5]	*Synechococcus*
KO:K03585	membrane fusion protein, multidrug efflux system	*Synechococcus*
KO:K03594	bacterioferritin [EC:1.16.3.1]	*Synechococcus*
KO:K03596	GTP-binding protein LepA	*elongatus*
KO:K03605	hydrogenase maturation protease [EC:3.4.23.-]	*elongatus*
KO:K03606	putative colanic acid biosysnthesis UDP-glucose lipid carrier transferase	*elongatus*
KO:K03630	DNA repair protein RadC	*elongatus*
KO:K03638	molybdopterin adenylyltransferase [EC:2.7.7.75]	*Synechococcus*
KO:K03665	GTPase	*elongatus*
KO:K03699	putative hemolysin	*Synechococcus*
KO:K03707	thiaminase (transcriptional activator TenA) [EC:3.5.99.2]	*elongatus*

Table B.5 continued from previous page

KEGG Function ID	**Function Name**	***Synechococcus* or *elongatus***
KO:K03724	ATP-dependent helicase Lhr and Lhr-like helicase [EC:3.6.4.-]	*Synechococcus*
KO:K03733	integrase/recombinase XerC	*elongatus*
KO:K03741	arsenate reductase (thioredoxin) [EC:1.20.4.4]	*Synechococcus*
KO:K03744	LemA protein	*Synechococcus*
KO:K03778	D-lactate dehydrogenase [EC:1.1.1.28]	*elongatus*
KO:K03781	catalase [EC:1.11.1.6]	*elongatus*
KO:K03828	putative acetyltransferase [EC:2.3.1.-]	*elongatus*
KO:K03830	putative acetyltransferase [EC:2.3.1.-]	*Synechococcus*
KO:K03839	flavodoxin I	*elongatus*
KO:K03841	fructose-1,6-bisphosphatase I [EC:3.1.3.11]	*elongatus*
KO:K03867	UDP-glucose:tetrahydrobiopterin glucosyltransferase [EC:2.4.1.-]	*elongatus*
KO:K03972	phage shock protein E	*elongatus*
KO:K04027	ethanolamine utilization protein EutM	*Synechococcus*
KO:K04028	ethanolamine utilization protein EutN	*Synechococcus*
KO:K04033	AraC family transcriptional regulator, ethanolamine operon transcriptional activator	*Synechococcus*
KO:K04488	nitrogen fixation protein NifU and related proteins	*elongatus*
KO:K04749	anti-sigma B factor antagonist	*Synechococcus*
KO:K04758	ferrous iron transport protein A	*Synechococcus*
KO:K04759	ferrous iron transport protein B	*Synechococcus*
KO:K04771	serine protease Do [EC:3.4.21.107]	*elongatus*
KO:K05275	pyridoxine 4-dehydrogenase [EC:1.1.1.65]	*Synechococcus*
KO:K05300	L-amino acid dehydrogenase [EC:1.4.5.-]	*Synechococcus*
KO:K05341	amylosucrase [EC:2.4.1.4]	*Synechococcus*
KO:K05378	phycoerythrin-associated linker protein	*Synechococcus*
KO:K05382	phycoerythrin-associated linker protein	*elongatus*
KO:K05521	ADP-ribosyl-[dinitrogen reductase] hydrolase [EC:3.2.2.24]	*elongatus*
KO:K05586	bidirectional [NiFe] hydrogenase diaphorase subunit [EC:7.1.1.2]	*elongatus*
KO:K05592	ATP-dependent RNA helicase DeaD [EC:3.6.4.13]	*Synechococcus*
KO:K05770	translocator protein	*Synechococcus*
KO:K05841	sterol 3beta-glucosyltransferase [EC:2.4.1.173]	*elongatus*
KO:K05844	ribosomal protein S6–L-glutamate ligase [EC:6.3.2.-]	*Synechococcus*
KO:K05846	osmoprotectant transport system permease protein	*Synechococcus*
KO:K06006	periplasmic protein CpxP/Spy	*elongatus*
KO:K06041	arabinose-5-phosphate isomerase [EC:5.3.1.13]	*elongatus*
KO:K06131	cardiolipin synthase A/B [EC:2.7.8.-]	*Synechococcus*
KO:K06177	tRNA pseudouridine32 synthase / 23S rRNA pseudouridine746 synthase [EC:5.4.99.28 5.4.99.29]	*elongatus*
KO:K06178	23S rRNA pseudouridine2605 synthase [EC:5.4.99.22]	*elongatus*
KO:K06199	fluoride exporter	*Synechococcus*
KO:K06223	DNA adenine methylase [EC:2.1.1.72]	*elongatus*
KO:K06346	spoIIIJ-associated protein	*elongatus*
KO:K06878	tRNA-binding protein	*elongatus*
KO:K06917	tRNA 2-selenouridine synthase [EC:2.9.1.3]	*Synechococcus*
KO:K06919	putative DNA primase/helicase	*elongatus*
KO:K07024	sucrose-6-phosphatase [EC:3.1.3.24]	*Synechococcus*
KO:K07053	3’,5’-nucleoside bisphosphate phosphatase [EC:3.1.3.97]	*elongatus*

Table B.5 continued from previous page

KEGG Function ID	Function Name	*Synechococcus* or *elongatus*
KO:K07058	membrane protein	*Synechococcus*
KO:K07100	putative phosphoribosyl transferase	*Synechococcus*
KO:K07114	Ca-activated chloride channel homolog	*Synechococcus*
KO:K07172	antitoxin MazE	*Synechococcus*
KO:K07192	flotillin	*elongatus*
KO:K07213	copper chaperone	*elongatus*
KO:K07232	glutathione-specific gamma-glutamylcyclotransferase [EC:4.3.2.7]	*elongatus*
KO:K07243	high-affinity iron transporter	*elongatus*
KO:K07253	phenylpyruvate tautomerase [EC:5.3.2.1]	*elongatus*
KO:K07267	porin	*Synechococcus*
KO:K07340	inner membrane protein	*Synechococcus*
KO:K07393	glutathionyl-hydroquinone reductase [EC:1.8.5.7]	*Synechococcus*
KO:K07397	putative redox protein	*Synechococcus*
KO:K07448	restriction system protein	*elongatus*
KO:K07465	putative RecB family exonuclease	*elongatus*
KO:K07478	putative ATPase	*Synechococcus*
KO:K07496	putative transposase	*elongatus*
KO:K07497	putative transposase	*Synechococcus*
KO:K07577	putative mRNA 3-end processing factor	*Synechococcus*
KO:K07638	two-component system, OmpR family, osmolarity sensor histidine kinase EnvZ [EC:2.7.13.3]	*Synechococcus*
KO:K07659	two-component system, OmpR family, phosphate regulon response regulator OmpR	*Synechococcus*
KO:K07727	putative transcriptional regulator	*elongatus*
KO:K07755	arsenite methyltransferase [EC:2.1.1.137]	*Synechococcus*
KO:K08218	MFS transporter, PAT family, beta-lactamase induction signal transducer AmpG	*Synechococcus*
KO:K08276	ecotin	*Synechococcus*
KO:K08303	U32 family peptidase [EC:3.4.-.-]	*elongatus*
KO:K08304	membrane-bound lytic murein transglycosylase A [EC:4.2.2.-]	*elongatus*
KO:K08368	MFS transporter, putative metabolite transport protein	*elongatus*
KO:K08679	UDP-glucuronate 4-epimerase [EC:5.1.3.6]	*Synechococcus*
KO:K08693	2',3'-cyclic-nucleotide 2'-phosphodiesterase / 3'-nucleotidase / 5'-nucleotidase [EC:3.1.4.16 3.1.3.6 3.1.3.5]	*elongatus*
KO:K08696	carbon dioxide concentrating mechanism protein CcmK	*elongatus*
KO:K08697	carbon dioxide concentrating mechanism protein CcmL	*elongatus*
KO:K08698	carbon dioxide concentrating mechanism protein CcmM	*elongatus*
KO:K08699	carbon dioxide concentrating mechanism protein CcmN	*elongatus*
KO:K08700	carbon dioxide concentrating mechanism protein CcmO	*elongatus*
KO:K08715	inward rectifier potassium channel	*Synechococcus*
KO:K08904	photosystem II Psb28-2 protein	*elongatus*
KO:K08964	methylthioribulose-1-phosphate dehydratase [EC:4.2.1.109]	*elongatus*
KO:K08966	2-hydroxy-3-keto-5-methylthiopentenyl-1-phosphate phosphatase [EC:3.1.3.87]	*elongatus*
KO:K08994	ion channel-forming bestrophin family protein	*elongatus*
KO:K09015	Fe-S cluster assembly protein SufD	*elongatus*
KO:K09125	queuosine precursor transporter	*Synechococcus*
KO:K09181	acetyltransferase	*Synechococcus*
KO:K09474	acid phosphatase (class A) [EC:3.1.3.2]	*Synechococcus*
KO:K09777	extracellular matrix regulatory protein A	*elongatus*

Table B.5 continued from previous page

KEGG Function ID	**Function Name**	***Synechococcus* or *elongatus***
KO:K09815	zinc transport system substrate-binding protein	*Synechococcus*
KO:K09816	zinc transport system permease protein	*Synechococcus*
KO:K09817	zinc transport system ATP-binding protein [EC:7.2.2.20]	*Synechococcus*
KO:K09818	manganese/iron transport system substrate-binding protein	*elongatus*
KO:K09819	manganese/iron transport system permease protein	*elongatus*
KO:K09820	manganese/iron transport system ATP-binding protein	*elongatus*
KO:K09825	Fur family transcriptional regulator, peroxide stress response regulator	*elongatus*
KO:K09836	beta-carotene/zeaxanthin 4-ketolase [EC:1.14.99.63 1.14.99.64]	*Synechococcus*
KO:K09880	enolase-phosphatase E1 [EC:3.1.3.77]	*elongatus*
KO:K10211	4,4'-diaponeurosporenoate glycosyltransferase [EC:2.4.1.-]	*Synechococcus*
KO:K10212	glycosyl-4,4'-diaponeurosporenoate acyltransferase [EC:2.3.1.-]	*Synechococcus*
KO:K10255	acyl-lipid omega-6 desaturase (Delta-12 desaturase) [EC:1.14.19.23 1.14.19.45]	*Synechococcus*
KO:K10536	agmatine deiminase [EC:3.5.3.12]	*elongatus*
KO:K10747	DNA ligase 1 [EC:6.5.1.1 6.5.1.6 6.5.1.7]	*Synechococcus*
KO:K10778	AraC family transcriptional regulator, regulatory protein of adaptative response / methylated-DNA-[protein]-cysteine methyltransferase [EC:2.1.1.63]	*elongatus*
KO:K11085	ATP-binding cassette, subfamily B, bacterial MsbA [EC:7.5.2.6]	*Synechococcus*
KO:K11105	potassium/hydrogen antiporter	*Synechococcus*
KO:K11327	membrane fusion protein, cation efflux system	*Synechococcus*
KO:K11328	two-component system, OmpR family, Ni(II)-sensor and/or redox sensor kinase NrsS [EC:2.7.13.3]	*Synechococcus*
KO:K11332	two-component system, OmpR family, response regulator NblR	*elongatus*
KO:K11356	two-component system, sensor histidine kinase and response regulator [EC:2.7.13.3]	*elongatus*
KO:K11443	two-component system, cell cycle response regulator DivK	*elongatus*
KO:K11520	two-component system, OmpR family, manganese sensing sensor histidine kinase [EC:2.7.13.3]	*elongatus*
KO:K11521	two-component system, OmpR family, manganese sensing response regulator	*elongatus*
KO:K11522	two-component system, chemotaxis family, response regulator PixG	*elongatus*
KO:K11523	two-component system, chemotaxis family, response regulator PixH	*elongatus*
KO:K11524	positive phototaxis protein PixI	*elongatus*
KO:K11525	methyl-accepting chemotaxis protein PixJ	*elongatus*
KO:K11526	two-component system, chemotaxis family, sensor histidine kinase and response regulator PixL	*elongatus*
KO:K11601	manganese transport system substrate-binding protein	*Synechococcus*
KO:K11602	manganese transport system permease protein	*Synechococcus*
KO:K11603	manganese transport system ATP-binding protein [EC:7.2.2.5]	*Synechococcus*
KO:K11751	5'-nucleotidase / UDP-sugar diphosphatase [EC:3.1.3.5 3.6.1.45]	*elongatus*
KO:K11780	7,8-didemethyl-8-hydroxy-5-deazariboflavin synthase [EC:4.3.1.32]	*elongatus*
KO:K11781	5-amino-6-(D-ribitylamino)uracil—L-tyrosine 4-hydroxyphenyl transferase [EC:2.5.1.147]	*elongatus*
KO:K11927	ATP-dependent RNA helicase RhlE [EC:3.6.4.13]	*Synechococcus*
KO:K11952	bicarbonate transport system ATP-binding protein [EC:7.3.2.-]	*elongatus*
KO:K11954	neutral amino acid transport system substrate-binding protein	*elongatus*
KO:K11955	neutral amino acid transport system permease protein	*elongatus*
KO:K11956	neutral amino acid transport system permease protein	*elongatus*
KO:K11957	neutral amino acid transport system ATP-binding protein	*elongatus*
KO:K11958	neutral amino acid transport system ATP-binding protein	*elongatus*

Table B.5 continued from previous page

KEGG Function ID	**Function Name**	***Synechococcus* or *elongatus***
KO:K11959	urea transport system substrate-binding protein	*Synechococcus*
KO:K11960	urea transport system permease protein	*Synechococcus*
KO:K11961	urea transport system permease protein	*Synechococcus*
KO:K11962	urea transport system ATP-binding protein	*Synechococcus*
KO:K11963	urea transport system ATP-binding protein	*Synechococcus*
KO:K12251	N-carbamoylputrescine amidase [EC:3.5.1.53]	*elongatus*
KO:K12260	sulfiredoxin [EC:1.8.98.2]	*elongatus*
KO:K12421	fatty acid CoA ligase FadD9	*Synechococcus*
KO:K12452	CDP-4-dehydro-6-deoxyglucose reductase, E1 [EC:1.17.1.1]	*Synechococcus*
KO:K12454	CDP-paratose 2-epimerase [EC:5.1.3.10]	*Synechococcus*
KO:K12957	alcohol/geraniol dehydrogenase (NADP+) [EC:1.1.1.2 1.1.1.183]	*Synechococcus*
KO:K13007	Fuc2NAc and GlcNAc transferase [EC:2.4.1.-]	*Synechococcus*
KO:K13421	uridine monophosphate synthetase [EC:2.4.2.10 4.1.1.23]	*Synechococcus*
KO:K13444	formylglycine-generating enzyme [EC:1.8.3.7]	*Synechococcus*
KO:K13628	iron-sulfur cluster assembly protein	*elongatus*
KO:K13643	Rrf2 family transcriptional regulator, iron-sulfur cluster assembly transcription factor	*elongatus*
KO:K13745	L-2,4-diaminobutyrate decarboxylase [EC:4.1.1.86]	*Synechococcus*
KO:K14337	alpha-1,6-mannosyltransferase [EC:2.4.1.-]	*Synechococcus*
KO:K14340	mannosyltransferase [EC:2.4.1.-]	*Synechococcus*
KO:K14606	lycopene cyclase CruP [EC:5.5.1.19]	*elongatus*
KO:K14623	DNA-damage-inducible protein D	*elongatus*
KO:K14742	tRNA threonylcarbamoyladenosine biosynthesis protein TsaB	*Synechococcus*
KO:K15034	ribosome-associated protein	*Synechococcus*
KO:K15256	tRNA (cmo5U34)-methyltransferase [EC:2.1.1.-]	*Synechococcus*
KO:K15383	MtN3 and saliva related transmembrane protein	*Synechococcus*
KO:K15579	nitrate/nitrite transport system ATP-binding protein	*elongatus*
KO:K15977	putative oxidoreductase	*Synechococcus*
KO:K16052	MscS family membrane protein	*Synechococcus*
KO:K16264	cobalt-zinc-cadmium efflux system protein	*Synechococcus*
KO:K16291	L,D-transpeptidase ErfK/SrfK	*Synechococcus*
KO:K16554	polysaccharide biosynthesis transport protein	*Synechococcus*
KO:K16694	teichuronic acid exporter	*Synechococcus*
KO:K16703	colanic acid/amylovoran biosynthesis glycosyltransferase [EC:2.4.-.-]	*Synechococcus*
KO:K16870	N-acetylglucosaminyl-diphospho-decaprenol L-rhamnosyltransferase [EC:2.4.1.289]	*Synechococcus*
KO:K16881	mannose-1-phosphate guanylyltransferase / phosphomannomutase [EC:2.7.7.13 5.4.2.8]	*elongatus*
KO:K16898	ATP-dependent helicase/nuclease subunit A [EC:3.1.-.- 3.6.4.12]	*elongatus*
KO:K16915	nickel transport protein	*Synechococcus*
KO:K17081	prohibitin 2	*Synechococcus*
KO:K17218	sulfide:quinone oxidoreductase [EC:1.8.5.4]	*Synechococcus*
KO:K17472	Rrf2 family transcriptional regulator, cysteine metabolism repressor	*elongatus*
KO:K17830	digeranylgeranylglycerophospholipid reductase [EC:1.3.1.101 1.3.7.11]	*elongatus*
KO:K18199	cyclohexyl-isocyanide hydratase [EC:4.2.1.103]	*elongatus*
KO:K18208	renalase [EC:1.6.3.5]	*elongatus*
KO:K18234	virginiamycin A acetyltransferase [EC:2.3.1.-]	*elongatus*
KO:K18676	glucosamine kinase [EC:2.7.1.8]	*Synechococcus*
KO:K18829	antitoxin VapB	*Synechococcus*

Table B.5 continued from previous page

KEGG Function ID	**Function Name**	***Synechococcus* or *elongatus***
KO:K18911	L-histidine Nalpha-methyltransferase [EC:2.1.1.44]	*Synechococcus*
KO:K18918	RHH-type transcriptional regulator, rel operon repressor / antitoxin RelB	*Synechococcus*
KO:K18997	chaperone modulatory protein CbpM	*Synechococcus*
KO:K19156	antitoxin PrlF	*Synechococcus*
KO:K19789	DNA repair protein RadD	*elongatus*
KO:K20276	large repetitive protein	*Synechococcus*
KO:K20444	O-antigen biosynthesis protein [EC:2.4.1.-]	*Synechococcus*
KO:K20866	glucose-1-phosphatase [EC:3.1.3.10]	*elongatus*
KO:K21020	diguanylate cyclase [EC:2.7.7.65]	*Synechococcus*
KO:K21023	diguanylate cyclase [EC:2.7.7.65]	*Synechococcus*
KO:K21071	ATP-dependent phosphofructokinase / diphosphate-dependent phosphofructokinase [EC:2.7.1.11 2.7.1.90]	*elongatus*
KO:K21140	[CysO sulfur-carrier protein]-S-L-cysteine hydrolase [EC:3.13.1.6]	*Synechococcus*
KO:K21344	D-glycero-beta-D-manno-heptose-7-phosphate kinase [EC:2.7.1.167]	*elongatus*
KO:K21345	D-glycero-beta-D-manno-heptose 1-phosphate adenylyltransferase [EC:2.7.7.70]	*elongatus*
KO:K21498	antitoxin HigA-1	*elongatus*
KO:K21639	L-arginine oxidase [EC:1.4.3.25]	*elongatus*
KO:K21703	LysR family transcriptional regulator, low CO2-responsive transcriptional regulator	*elongatus*
KO:K21903	ArsR family transcriptional regulator, lead/cadmium/zinc/bismuth-responsive transcriptional repressor	*elongatus*
KO:K21908	membrane protein HdeD	*Synechococcus*
KO:K21919	BTB/POZ domain-containing protein KCTD9	*Synechococcus*
KO:K22044	moderate conductance mechanosensitive channel	*elongatus*
KO:K22205	S-adenosyl-L-methionine hydrolase (adenosine-forming) [EC:3.13.1.8]	*elongatus*
KO:K22250	poly(3-hydroxyoctanoate) depolymerase [EC:3.1.1.76]	*Synechococcus*
KO:K22278	peptidoglycan-N-acetylglucosamine deacetylase [EC:3.5.1.104]	*Synechococcus*
KO:K22450	aralkylamine N-acetyltransferase [EC:2.3.1.87]	*elongatus*
KO:K22468	polyphosphate kinase [EC:2.7.4.1]	*Synechococcus*
KO:K23139	cytochrome P450 family 110 [EC:1.14.-.-]	*elongatus*
KO:K23242	manganese efflux pump family protein	*Synechococcus*
KO:K23753	AMP-polyphosphate phosphotransferase [EC:2.7.4.33]	*elongatus*
KO:K23994	solute carrier family 9B (sodium/hydrogen exchanger), member 1/2	*elongatus*
KO:K24157	thioredoxin-dependent peroxiredoxin [EC:1.11.1.24]	*Synechococcus*
KO:K24158	thioredoxin-dependent peroxiredoxin [EC:1.11.1.24]	*elongatus*

Table B.6: **Genomes selected for comparative genomic analyses**. BUSCO scores above 90% were the threshold for high quality genome assembly. F-Pcy: freshwater picocyanobacteria; SE-Cy: *Synechococcus elongatus* strains; T-Pcy: thermal picocyanobacteria; M-Pcy: marine picocyanobacteria; LF-Cy: larger freshwater cyanobacteria.

Group	**Strain**	**Isolation Location**	**Accession**	**Genome Size (Mbp)**	**BUSCO Complete (%)**
F-Pcy	Cyanobium Gracile PCC 6307	Lake water, Wisconsin, USA	GCF_000316515.1	3.34	97%
F-Pcy	Cyanobium usitatum Tous	Tous Reservoir, Spain	GCF_003011885.1	2.52	96%
F-Pcy	Synechococcus lacustris Tous	Tous Reservoir, Spain	GCF_003011125.1	2.67	92%
F-Pcy	Synechococcus sp CCAP 1479_10	South Basin, Windermere, UK	GCF_019038515.1	3.31	95%
F-Pcy	Synechococcus sp CCAP 1479_13	North Basin, Windermere, UK	GCF_019038475.1	3.30	95%
F-Pcy	Synechococcus sp CCAP 1479_9	South Basin, Windermere, UK	GCF_019038535.1	3.29	96%
F-Pcy	Synechococcus sp CCY 0621	Leiden, The Netherlands	GCF_019038465.1	3.23	96%
F-Pcy	Synechococcus sp CCY 9618	Vinkeveen, The Netherlands	GCF_019038415.1	2.93	96%
F-Pcy	Synechococcus sp. MW101C3	Lake Mondsee, Austria	GCF_002252635.1	3.03	96%
F-Pcy	Synechococcus sp. 1G10	Lake Nahuel Huapia, Argentina	GCF_002252625.1	3.34	95%
F-Pcy	Synechococcus sp. Alchichicha_8F6	Lake Alchichicha, Mexico	GCF_002252665.1	2.51	96%
F-Pcy	Synechococcus sp. Atexcac_6H9	Lake Atexcac, Mexico	Ga0078145 (JGI)	2.52	96%
F-Pcy	Synechococcus sp. BO8801	Lake Constance, Germany	GCF_002252675.1	3.27	95%
F-Pcy	Synechococcus sp. FACHB-909	Baohu Lake, China	GCF_014698945.1	3.18	97%
F-Pcy	Synechococcus sp. Maggiore 4A10	Lake Maggiore, Italy	Ga0078148 (JGI)	3.15	96%
F-Pcy	Vulcanococcus limneticus LL	Lake Albano, Italy	GCF_002252705.1	3.55	96%
F-Pcy	Vulcanococcus limneticus sp. 3B3	Lake Candia, Italy	GCF_024346195.1	3.58	96%
SE-Cy	Synechococcus elongatus FACHB-1061	Changjiang River, China	GCF_014698895.1	2.73	98%
SE-Cy	Synechococcus elongatus FACHB-242	Freshwater, USA	GCF_014698415.1	2.74	98%
SE-Cy	Synechococcus elongatus PCC 6301	Waller Creek, Austin, USA	GCF_022984195.1	2.70	97%
SE-Cy	Synechococcus elongatus PCC 7942	California, USA	GCF_014698905.1	2.74	98%
SE-Cy	Synechococcus sp. PCC 6312	California, USA	GCF_000316685.1	3.72	96%
SE-Cy	Synechococcus sp. PCC 7502	Sphagnum bog near Kastanienbaum, Switzerland	GCF_000317085.1	3.58	95%
SE-Cy	Synechococcus elongatus sp. UTEX 2973	Waller Creek, Austin, USA	GCF_000817325.1	2.74	98%
SE-Cy	Thermosynechococcus elongatus BP-1	Beppu Hot Spring, Japan	GCF_000011345.1	2.59	94%
SE-Cy	Thermosynechococcus sp. NK55a	Nakabusa Hot Spring, Japan	GCF_000505665.1	2.52	91%
T-Pcy	Synechococcus JA-2-3B'a(2-13)	Octopus Spring, Wyoming, USA	GCF_000013225.1	3.047	91%

Table B.6 continued from previous page

Group	Strain	Isolation Location	Accession	Genome Size (Mbp)	BUSCO Complete (%)
T-Pcy	Synechococcus JA-3-3Ab	Octopus Spring, Wyoming, USA	GCF_000013205.1	2.933	90%
T-Pcy	Synechococcus sp. 60AY4M2	Mushroom Spring, USA	GCF_002760375.1	3.163	90%
T-Pcy	Synechococcus sp. 63AY4M1	Mushroom Spring, USA	GCF_002760395.1	3.159	90%
T-Pcy	Synechococcus sp. 63AY4M2	Mushroom Spring, USA	GCF_002760475.1	3.094	90%
T-Pcy	Synechococcus sp. 65AY640	Mushroom Spring, USA	GCF_002760445.1	3.161	90%
T-Pcy	Synechococcus sp. 65AY6A5	Mushroom Spring, USA	GCF_002760415.1	2.982	90%
T-Pcy	Synechococcus sp. 65AY6Li	Mushroom Spring, USA	GCF_002760345.1	2.933	90%
M-Pcy	Cyanobium sp. NIES-981	Iriomoteejima Island, Okinawa, Japan	GCF_900088535.1	3.022	95%
M-Pcy	Cyanobium sp. PCC 7001	Long Island, USA	GCF_000155635.1	2.834	96%
M-Pcy	Prochlorococcus marinus str. AS9601	Arabian Sea	GCF_000015645.1	1.670	94%
M-Pcy	Prochlorococcus marinus str. MIT 9211	Equatorial Pacific	GCF_000018585.1	1.689	95%
M-Pcy	Prochlorococcus marinus str. MIT 9215	Equatorial Pacific	GCF_000018065.1	1.739	94%
M-Pcy	Prochlorococcus marinus str. MIT 9301	Sargasso Sea	GCF_000015965.1	1.642	95%
M-Pcy	Prochlorococcus marinus str. MIT 9302	Sargasso Sea	GCF_000759975.1	1.745	94%
M-Pcy	Prochlorococcus marinus str. MIT 9303	Sargasso Sea	GCF_000015705.1	2.683	96%
M-Pcy	Prochlorococcus marinus str. MIT 9312	North Atlantic Ocean	GCF_000012645.1	1.709	95%
M-Pcy	Prochlorococcus marinus str. MIT 9313	North Atlantic Ocean	GCF_000011485.1	2.411	97%
M-Pcy	Prochlorococcus marinus str. MIT 9515	Equatorial Pacific	GCF_000015665.1	1.704	93%
M-Pcy	Prochlorococcus marinus str. NATL1A	North Atlantic Ocean	GCF_000015685.1	1.865	94%
M-Pcy	Prochlorococcus marinus str. NATL2A	North Atlantic Ocean	GCF_000012465.1	1.843	94%
M-Pcy	Prochlorococcus marinus subsp. marinus str. CCMP1375	Sargasso Sea	GCF_000007925.1	1.751	94%
M-Pcy	Prochlorococcus marinus subsp. pastoris str. CCMP1986	Mediterranean Sea	GCF_000011465.1	1.658	94%
M-Pcy	Prochlorococcus sp. MIT 0604	Pacific Ocean	GCF_000757845.1	1.780	94%
M-Pcy	Prochlorococcus sp. MIT 0801	Sargasso Sea	GCF_000757865.1	1.929	94%
M-Pcy	Prochlorococcus sp. MIT 1306	Pacific Ocean	GCF_001631985.1	2.499	97%
M-Pcy	Prochlorococcus sp. P1363	Station ALOHA, North Pacific subtropical gyre	GCF_012933595.1	2.558	97%
M-Pcy	Prochlorococcus sp. RS01	Red Sea	GCF_001989435.1	1.658	95%
M-Pcy	Prochlorococcus sp. RS04	Red Sea	GCF_001989455.1	1.656	95%
M-Pcy	Prochlorococcus sp. RS50	Central Red Sea	GCF_001989415.1	1.656	95%
M-Pcy	Prochlorococcus sp. SS52	Sargasso Sea	GCF_000760375.1	1.754	94%
M-Pcy	Synechococcus sp. A15-127	Atlantic Ocean	GCF_014280215.1	2.543	98%

Table B.6 continued from previous page

Group	Strain	Isolation Location	Accession	Genome Size (Mbp)	BUSCO Complete (%)
M-Pcy	Synechococcus sp. A15-44	West Africa Upwelling	GCF_014280115.1	2.622	98%
M-Pcy	Synechococcus sp. BDU 130192	Kakinada, Andhra Pradesh, India	GCF_002721235.1	3.257	97%
M-Pcy	Synechococcus sp. BL107	Blanes Bay, Spain	GCF_000153805.1	2.283	98%
M-Pcy	Synechococcus sp. BMK-MC-1	Tyrrhenian Sea, Mediterranean	GCF_014279955.1	2.601	98%
M-Pcy	Synechococcus sp. BS55D	Black Sea	GCF_004332415.1	2.235	97%
M-Pcy	Synechococcus sp. BS56D	Black Sea	GCF_004332405.1	2.304	97%
M-Pcy	Synechococcus sp. CB0101	Chesapeake Bay Estuary, Maryland, USA	GCF_000179235.2	2.686	97%
M-Pcy	Synechococcus sp. CC9311	California Current	GCF_000014585.1	2.607	98%
M-Pcy	Synechococcus sp. CC9605	California Current	GCF_000012625.1	2.511	98%
M-Pcy	Synechococcus sp. CC9616	California Current	GCF_000515235.1	2.646	97%
M-Pcy	Synechococcus sp. CC9902	California Current	GCF_000012505.1	2.235	98%
M-Pcy	Synechococcus sp. GEYO	Sargasso Sea	GCF_900473955.1	2.344	97%
M-Pcy	Synechococcus sp. HB1133	Sydney Beach, Australia	GCF_011365345.1	2.379	97%
M-Pcy	Synechococcus sp. KORDI-100	Tropical Pacific Ocean	GCF_000737535.1	2.789	97%
M-Pcy	Synechococcus sp. KORDI-49	East China Sea	GCF_000737575.1	2.586	98%
M-Pcy	Synechococcus sp. KORDI-52	East China Sea	GCF_000737595.1	2.572	98%
M-Pcy	Synechococcus sp. LTW-R	Hong Kong	GCF_014217875.1	2.416	98%
M-Pcy	Synechococcus sp. MIT S9504	Equatorial Pacific	GCF_001632105.1	3.087	97%
M-Pcy	Synechococcus sp. MIT S9508	Equatorial Pacific	GCF_001632165.1	2.502	98%
M-Pcy	Synechococcus sp. MIT S9509	Equatorial Pacific	GCF_001631935.1	3.088	97%
M-Pcy	Synechococcus sp. N26	Sargasso Sea	GCF_900473975.1	2.127	98%
M-Pcy	Synechococcus sp. NIES-970	Rikuhama Beach, Japan	GCF_002356215.1	3.124	97%
M-Pcy	Synechococcus sp. NKBG042902	Okinawa, Japan	GCF_000715475.1	3.319	94%
M-Pcy	Synechococcus sp. NKBG15041c	Okinawa, Japan	GCF_000485815.1	3.180	93%
M-Pcy	Synechococcus sp. NOUM97013	Equatorial Pacific	GCF_014279815.1	2.553	97%
M-Pcy	Synechococcus sp. OG1	Marine	GCF_900177365.1	3.434	95%
M-Pcy	Synechococcus sp. PCC 11901	Pulau Ubin, Singapore	GCF_005577135.1	3.472	97%
M-Pcy	Synechococcus sp. PCC 7002	Magueyes Island, Puerto Rico, USA	GCF_000019485.1	3.410	97%
M-Pcy	Synechococcus sp. PCC 7003	Greenwich, Connecticut, USA	GCF_001693255.1	3.813	97%
M-Pcy	Synechococcus sp. PCC 7117	Port Hedland, Australia	GCF_001693275.1	3.930	97%
M-Pcy	Synechococcus sp. PCC 73109	City Island, New York, USA	GCF_001521855.1	3.674	97%
M-Pcy	Synechococcus sp. PCC 7335	Puerto Penasco, Mexico	GCF_000155595.1	5.974	97%
M-Pcy	Synechococcus sp. PCC 7336	East Pacific Ocean	GCF_000332275.1	5.141	92%

Table B.6 continued from previous page

Group	Strain	Isolation Location	Accession	Genome Size (Mbp)	BUSCO Complete (%)
M-Pcy	Synechococcus sp. PCC 8807	Port Gentil, Gabon	GCF_001693295.1	3.633	97%
M-Pcy	Synechococcus sp. RCC307	Mediterranean Sea	GCF_000063525.1	2.225	97%
M-Pcy	Synechococcus sp. RS9916	Red Sea	GCF_000153825.1	2.664	98%
M-Pcy	Synechococcus sp. RS9917	Red Sea	GCF_000153065.1	2.580	96%
M-Pcy	Synechococcus sp. RSCCF101	Red Sea	GCF_008807075.1	2.985	91%
M-Pcy	Synechococcus sp. SYN20	Raunefjorden, Atlantic Ocean	GCF_014279555.1	2.729	98%
M-Pcy	Synechococcus sp. SynAce01	Ace Lake, Antarctica	GCF_001885215.1	2.751	95%
M-Pcy	Synechococcus sp. UW105	Sargasso Sea	GCF_900473935.1	2.659	98%
M-Pcy	Synechococcus sp. UW140	Sargasso Sea	GCF_900474295.1	2.704	98%
M-Pcy	Synechococcus sp. UW179A	Sargasso Sea	GCF_900473965.1	3.058	97%
M-Pcy	Synechococcus sp. UW179B	Sargasso Sea	GCF_900474245.1	2.532	98%
M-Pcy	Synechococcus sp. UW69	Sargasso Sea	GCF_900474185.1	2.372	98%
M-Pcy	Synechococcus sp. UW86	Sargasso Sea	GCF_900474085.1	2.091	98%
M-Pcy	Synechococcus sp. WH 5701	Long Island Sound	GCF_000153045.1	3.044	97%
M-Pcy	Synechococcus sp. WH 7803	Sargasso Sea	GCF_000063505.1	2.367	98%
M-Pcy	Synechococcus sp. WH 7805	Sargasso Sea	GCF_000153285.1	2.620	97%
M-Pcy	Synechococcus sp. WH 8016	Sargasso Sea	GCF_000230675.1	2.707	98%
M-Pcy	Synechococcus sp. WH 8020	Sargasso Sea	GCF_001040845.1	2.661	98%
M-Pcy	Synechococcus sp. WH 8101	Woods Hole, USA	GCF_004209775.1	2.630	97%
M-Pcy	Synechococcus sp. WH 8102	Caribbean Sea	GCF_000195975.1	2.434	98%
M-Pcy	Synechococcus sp. WH 8103	Sargasso Sea	GCF_001182765.1	2.430	98%
M-Pcy	Synechococcus sp. WH 8109	Sargasso Sea	GCF_000161795.2	2.112	98%
M-Pcy	Synechococcus sp. XM-24	Xiamen Estuary, China	GCA_003149585.1	2.377	97%
M-Pcy	Synechococcus sp. YX04-3	South China Sea	GCA_003712165.1	2.460	99%
LF-Cy	Anabaena sp. 90	Lake Vesijdrvi, Finland	GCF_000312705.1	5.306	99%
LF-Cy	Anabaena sp. FACHB-1391	Dianchi Lake, China	GCF_014696755.1	4.497	98%
LF-Cy	Anabaena sp. WA102	Anderson Lake, USA	GCF_001277295.1	5.705	98%
LF-Cy	Aphanizomenon flos-aquae FACHB-1287	Lake, China	GCF_014698755.1	4.347	100%
LF-Cy	Aphanizomenon flos-aquae LD13	Upper Klamath Lake, USA	GCA_001672165.1	4.439	99%
LF-Cy	Aphanizomenon flos-aquae NIES-81	Lake Kasumigaura, Japan	GCF_000521175.1	5.683	99%
LF-Cy	Aphanizomenon sp. FACHB-1416	Dianchi Lake, China	GCF_014698695.1	4.287	99%
LF-Cy	Arthrospira platensis C1	Freshwater	GCF_025200965.1	6.089	97%
LF-Cy	Arthrospira platensis FACHB-971	Lake, China	GCF_014698385.1	5.624	99%
LF-Cy	Cylindrospermopsis raciborskii CS-505	Solmon Dam, Australia	GCF_001676585.1	3.879	97%

Table B.6 continued from previous page

Group	Strain	Isolation Location	Accession	Genome Size (Mbp)	BUSCO Complete (%)
LF-Cy	Cylindrospermopsis raciborskii ITEP-A1	Arcoverde Reservoir, Brazil	GCF_001586755.1	3.604	98%
LF-Cy	Cylindrospermopsis raciborskii MVCC14	Laguna Blanca, Uruguay	GCF_001858125.1	3.595	96%
LF-Cy	Cylindrospermopsis raciborskii S07	Lake Wivenhoe, Australia	GCF_002893205.1	3.872	98%
LF-Cy	Dolichospermum compactum NIES-806	Rostherne Mere, UK	GCF_002368115.1	5.166	100%
LF-Cy	Dolichospermum sp. FACHB-1091	Freshwater, China	GCF_014697585.1	4.516	99%
LF-Cy	Dolichospermum sp. UHCC 0352	Lake Tuusulanjarvi, Finland	GCF_009711925.1	5.129	93%
LF-Cy	Geminocystis herdmanii PCC-6308	Lake, Wisconsin, USA	GCF_000332235.1	4.263	97%
LF-Cy	Geminocystis sp. NIES-3708	Chiba, Narashino, Japan	GCF_001548095.1	4.042	97%
LF-Cy	Geminocystis sp. NIES-3709	Chiba, Narashino, Japan	GCF_001548115.1	4.426	97%
LF-Cy	Gloeocapsa sp. PCC 73106	Sphagnum bog, Switzerland	GCF_000332035.1	4.025	97%
LF-Cy	Gloeocapsa sp. PCC 7428	Maha Oya, Sri Lanka	GCF_000317555.1	5.883	99%
LF-Cy	Microcystis aeruginosa KLA2	Copco Reservoir, USA	GCF_008257445.1	5.213	99%
LF-Cy	Microcystis aeruginosa NIES-2481	Lake Kasumigaura, Japan	GCF_001704955.2	4.293	99%
LF-Cy	Microcystis aeruginosa PCC 7005	Madison, USA	GCF_000599945.1	4.906	100%
LF-Cy	Microcystis aeruginosa PCC 7806SL	Freshwater, Netherlands	GCF_002095975.1	5.139	100%
LF-Cy	Microcystis panniformis FACHB-1757	Lake Taihu, China	GCF_001264245.1	5.687	92%
LF-Cy	Microcystis sp. 0824	Lake Hachiro, Japan	GCF_003112475.1	4.103	100%
LF-Cy	Microcystis sp. MC19	Lake Seo, South Korea	GCF_003019735.1	5.020	100%
LF-Cy	Nostoc paludosum FACHB-159	Hunan Province, China	GCF_014698835.1	9.270	100%
LF-Cy	Nostoc sp. PCC 7107	Point Reyes Penisula, USA	GCF_000316625.1	6.330	100%
LF-Cy	Nostoc sp. PCC 7120	Freshwater Lake , USA	GCF_000009705.1	7.212	100%
LF-Cy	Nostoc sp. PCC 7524	Hot Spring, Amparai District, Sri Lanka	GCF_000316645.1	6.719	100%
LF-Cy	Nostoc sphaeroides Kutzing En	Enshi, China	GCF_003443655.1	6.684	98%
LF-Cy	Planktothricoides sp. FACHB-1261	China	GCF_014698535.1	7.011	98%
LF-Cy	Planktothricoides sp. FACHB-1370	Beihu Lake, China	GCF_014698175.1	6.999	98%
LF-Cy	Planktothricoides sp. SR001	Freshwater Reservoir, Singapore	GCF_001276715.1	7.067	98%
LF-Cy	Planktothrix agardhii NIES-204	Lake Kasumigaura, Japan	GCA_003609755.1	5.084	98%
LF-Cy	Planktothrix agardhii NIVA-CYA-1268	Lake Langsjön, Sweden	GCF_000710505.1	5.046	96%
LF-Cy	Planktothrix prolifica NIVA-CYA 406	Steinsfjorden, Norway	GCF_000464765.1	5.763	97%
LF-Cy	Planktothrix rubescens sp. NIVA-CYA407	Steinsfjorden, Norway	GCF_000464785.1	5.510	98%
LF-Cy	Planktothrix sp. PCC 11201	Waitaki River, New Zealand	GCF_900009135.1	5.988	98%

Table B.6 continued from previous page

Group	Strain	Isolation Location	Accession	Genome Size (Mbp)	BUSCO Complete (%)
LF-Cy	Pleurocapsa sp. PCC 7319	Puerto Penasco, USA	GCF_000332195.1	7.387	98%
LF-Cy	Pleurocapsa sp. PCC-7327	Hunter's Hot Spring, USA	GCF_000317025.1	4.987	99%
LF-Cy	Prochlorothrix hollandica PCC 9006	Lake Loosdrecht, The Netherlands	GCF_000332315.1	3.378	96%
LF-Cy	Pseudanabaena sp. ABRG5-3	Shizuoka, Japan	GCF_003967015.1	5.658	95%
LF-Cy	Pseudanabaena sp. FACHB-723	Dianchi Lake, China	GCF_014696385.1	3.964	94%
LF-Cy	Pseudanabaena sp. PCC 6802	Berkeley, USA	GCF_000332175.1	5.622	95%
LF-Cy	Synechocystis sp. CACIAM 05	Acude Bolonia, Brazil	GCF_010222815.1	3.581	97%
LF-Cy	Synechocystis sp. FACHB-929	Japan	GCF_014698625.1	3.847	97%
LF-Cy	Synechocystis sp. PCC 6714	Oakland, California	GCF_000478825.2	3.739	97%
LF-Cy	Synechocystis sp. PCC 6803	Oakland, California	GCF_000009725.1	3.947	97%

Table B.7: **List of genes screened among cyanobacteria.** Protein ID refers to the KO code or experimentally determined gene locus. Amino acid biosynthesis genes were found via KEGG pathway searches.

Gene	Protein ID	Reference	Gene	Protein ID	Reference
ntcA	K21561	[536]	*natH*	K09971	[254]
ntcB	K09681	[216]	*bgtA*	K17063	[254]
glnB	K04751	[537]	*bgtB*	K17062	[254]
icd	K00031	[538]	*natI*	K01999	[255]
pipX	K14518	[208]	*natJ*	K01997	[255]
spkE	K08884	[539]	*natK*	K01998	[255]
amt1	K03320	[220]	*natL*	K01995	[255]
amt2	sll1017	[220]	*natM*	K01996	[255]
amt3	sll0537	[220]	*gtrA*	sll1102	[259]
amt4	alr0990	[222]	*gtrB*	sll1103	[259]
amtB	K03320	[221]	*gtrC*	K02030	[259]
nrtA	K15576	[229]	*gltS*	K03312	[259]
nrtB	K15577	[229]	*dmeA*	Synpcc7942_1024	[258]
nrtC	K15578	[229]	*agcS*	K03310	[260]
nrtD	K15579	[229]	*oppC*	K02034	[261]
nrtP	K02575	[540]	*dppA*	K02035	[262]
focA	K21993	[232]	*dppB*	K02033	[262]
nrtS	syc1385_c	[233]	*dppC*	K02034	[262]
cynA	K15576	[541]	*dppD*	K02032	[262]
cynB	K15577	[541]	*dppF*	K02032	[262]
cynD	K15578	[541]	*glnA*	K01915	[227]
cynS	K01725	[542]	*glnN*	K01915	[227]
cynH	Ga0125238_111683	[235]	*glsF*	K00284	[263]
urtA	K11959	[247]	*gltD*	K00266	[543]

Table B.7 continued from previous page

Gene	Protein ID	Reference	Gene	Protein ID	Reference
urtB	K11960	[247]	*gifA*	ssl1911	[544]
urtC	K11961	[247]	*gifB*	sll1515	[544]
urtD	K11962	[247]	*gdhA*	K00262	[264]
urtE	K11963	[247]	*narB*	K00367	[545]
natA	K11957	[253]	*nirA*	K00366	[546]
natB	K11954	[253]	*nirB*	Synpcc7942_1241	[215]
natC	K11955	[253]	*ureA*	K01430	[249]
natD	K11956	[253]	*ureB*	K01429	[249]
natE	K11958	[253]	*ureC*	K01428	[249]
natF	K09969	[254]	*ureD*	K03190	[249]
natG	K09970	[254]	*ureE*	K03187	[249]
ureF	K03188	[249]	*chiA*	K01183	[389]
ureG	K03189	[249]	*choA*	K01233	[389]
ureJ	K03192	[547]	*phnD*	K02044	[525]
carA	K01956	[548]	*amiA*	K01426	[549]
carB	K01955	[548]	*sufS*	K11717	[550]
narM	Synpcc7942_0933	[551]	*ilvB*	K01652	[344]
cphA	K03802	[552]	*ilvN*	K01653	[344]
cphB	K13282	[552]	*ilvC*	K00053	[344]
iadA	K01424	[191]	*ilvD*	K01687	[344]
nblA	pfam04485	[553]	*livE*	K00826	[344]
speA	K01585	[266]	*leuA*	K01649	[344]
speB	K01480	[266]	*leuD*	K01704	[344]
aguA	K10536	[266]	*leuC*	K01703	[344]
aguB	K12251	[266]	*leuB*	K00052	[344]
puo	K03343	[266]	*metH*	K00548	[344]
putA	K00128	[266]	*pheA*	K04518	[344]

Table B.7 continued from previous page

Gene	Protein ID	Reference	Gene	Protein ID	Reference
arcA	sll1336	[266]	*trpE*	K01657	[344]
arcC	K00926	[266]	*trpG*	K01664	[344]
aoxA	K05300	[266]	*trpD*	K00766	[344]
syNOS	S7335_573	[267]	*trpF*	K01817	[344]
nifD	K02586	[554]	*trpC*	K01609	[344]
nifK	K02591	[555]	*trpA*	K01695	[344]
fdxN	SMa0811	[550]	*trpB*	K01696	[344]
nifU	K13819	[550]	*tyrA*	K15226	[344]
nifB	K02585	[550]	*asnB*	K01953	[344]
nifE	K02587	[556]	*asnS*	K01893	[344]
nifH	K02588	[555]	*aspS*	K01876	[344]
nifN	K02592	[556]	*gatA*	K02433	[344]
nifX	K02596	[557]	*gatB*	K02434	[344]
nifW	K02595	[557]	*gatC*	K02435	[344]
nifV	K02594	[558]	*cysE*	K00640	[344]
nifZ	K02597	[558]	*cysM1*	K01738	[344]
nifT	K02593	[558]	*cysM2*	K01738	[344]
chdA	K22278	[389]	*serA*	K00058	[344]
aspC	K00812	[344]	*glyA*	K00600	[344]
lysC	K00928	[344]	*hisG*	K00765	[344]
asd	K00133	[344]	*hisIE*	K11755	[344]
thrA	K00003	[344]	*hisA*	K01814	[344]
thrB	K00872	[344]	*hisF*	K02500	[344]
thrC1	K01733	[344]	*hisH*	K02501	[344]
thrC2	K01733	[344]	*hisB*	K01693	[344]
argJ	K00620	[344]	*hisC*	K00817	[344]
argB	K00930	[344]	*hisD*	K00013	[344]

Table B.7 continued from previous page

Gene	Protein ID	Reference	Gene	Protein ID	Reference
argC	K00145	[344]	*dapA*	K01714	[344]
argD	K00821	[344]	*dapB*	K00215	[344]
argF	K00611	[344]	*aspB*	K10206	[344]
argG	K01940	[344]	*dapF*	K01778	[344]
argH	K01755	[344]	*lysA*	K01586	[344]
kprS	K00948	[344]	*proB*	K00931	[344]
hisZ	K02502	[344]	*proA*	K00147	[344]
ald	K00259	[344]	*proC*	K00286	[344]

Table B.8: **Query sequences used for BLASTP of nitrogen assimilatory genes**

N Source	Gene	Accession Number	Species
Amino Acid	*agcS*	WP_015226123.1	*Halothece* sp. PCC 7418
Metabolism	*aguA*	WP_011378474.1	*Synechococcus_elongatus*_PCC_7942
Metabolism	*aguB*	WP_011378338.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*ald*	Ga0078150_104155	*Synechococcus* sp. 1G10
		WP_011429984.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011127372.1	*Synechococcus*_sp_WH8102
		WP_066910947.1	*Synechococcus*_sp__MIT9509
		WP_011825163.1	*Prochlorococcus_marinus*_MIT9303
		WP_011378176.1	*Synechococcus_elongatus*_PCC_7942
Amides	*amiA*	WP_011378077.1	*Synechococcus_elongatus*_PCC_7942
Ammonium	*amt1*	WP_011377582.1	*Synechococcus_elongatus*_PCC_7942
Ammonium	*amt2*	WP_162327372.1	*Synechocystis* sp. CACIAM 05
Ammonium	*amt3*	WP_162329244.1	*Synechocystis* sp. CACIAM 05
Ammonium	*amt4*	WP_044520794.1	*Nostoc*_sp__PCC_7120
Ammonium	*amtB*	WP_011378385.1	*Synechococcus_elongatus*_PCC_7942
Metabolism	*aoxA*	WP_162328701.1	*Synechocystis* sp. CACIAM 05
Metabolism	*arcA*	WP_162329280.1	*Synechocystis* sp. CACIAM 05
Metabolism	*arcC*	WP_162329465.1	*Synechocystis* sp. CACIAM 05
Amino Acid Biosynthesis	*argB*	WP_094554544.1	*Synechococcus* sp. 1G10
		WP_011429220.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011128640.1	*Synechococcus*_sp_WH8102
		WP_066905848.1	*Synechococcus*_sp__MIT9509
		WP_011825399.1	*Prochlorococcus_marinus*_MIT9303
		WP_011823296.1	*Prochlorococcus_marinus*_NATL1A
		WP_011244792.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*argC*	Ga0078150_101229	*Synechococcus* sp. 1G10
		WP_011429549.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011128126.1	*Synechococcus*_sp_WH8102
		WP_066907555.1	*Synechococcus*_sp__MIT9509
		WP_063394957.1	*Prochlorococcus_marinus*_MIT9303
		WP_011823675.1	*Prochlorococcus_marinus*_NATL1A
		WP_011242437.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*argD*	Ga0078150_102151	*Synechococcus* sp. 1G10
		WP_011430379.1	*Synechococcus*_sp_JA-3-3Ab
		WP_042503704.1	*Synechococcus*_sp_WH8102
		WP_066906857.1	*Synechococcus*_sp__MIT9509
		WP_041374860.1	*Prochlorococcus_marinus*_MIT9303

Table B.8 continued from previous page

N Source	Gene	Accession Number	Species
Amino Acid Biosynthesis	*argD*	WP_071813355.1	*Prochlorococcus_marinus*_NATL1A
		WP_011377800.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*argF*	Ga0078150_10453	*Synechococcus* sp. 1G10
		WP_011431589.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011128450.1	*Synechococcus*_sp_WH8102
		WP_066907073.1	*Synechococcus*_sp__MIT9509
		WP_011826540.1	*Prochlorococcus_marinus*_MIT9303
		WP_011824242.1	*Prochlorococcus_marinus*_NATL1A
		WP_011243902.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*argG*	WP_011429604.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011129366.1	*Synechococcus*_sp_WH8102
Amino Acid Biosynthesis	*argH*	WP_011429025.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011126894.1	*Synechococcus*_sp_WH8102
Amino Acid Biosynthesis	*argJ*	WP_011430182.1	*Synechococcus*_sp_JA-3-3Ab
		WP_042504501.1	*Synechococcus*_sp_WH8102
Amino Acid Biosynthesis	*asd*	Ga0078150_10527	*Synechococcus* sp. 1G10
		WP_011430691.1	*Synechococcus*_sp_JA-3-3Ab
		WP_042504046.1	*Synechococcus*_sp_WH8102
		WP_066908499.1	*Synechococcus*_sp__MIT9509
		WP_011824760.1	*Prochlorococcus_marinus*_MIT9303
Amino Acid Biosynthesis	*asnB*	WP_011129306.1	*Synechococcus*_sp_WH8102
		WP_074159255.1	*Synechococcus*_sp__MIT9509
		WP_011823568.1	*Prochlorococcus_marinus*_NATL1A
Amino Acid Biosynthesis	*asnS*	WP_162329229.1	*Synechocystis* sp. CACIAM 05
Amino Acid Biosynthesis	*aspB*	Ga0078150_11274	*Synechococcus* sp. 1G10
		WP_011431255.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011129003.1	*Synechococcus*_sp_WH8102
		WP_066905005.1	*Synechococcus*_sp__MIT9509
		WP_041375215.1	*Prochlorococcus_marinus*_MIT9303
		WP_011824469.1	*Prochlorococcus_marinus*_NATL1A
		WP_011242999.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*aspC*	Ga0078150_12147	*Synechococcus* sp. 1G10
		WP_011429705.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011128036.1	*Synechococcus*_sp_WH8102
		WP_066905320.1	*Synechococcus*_sp__MIT9509
		WP_011826075.1	*Prochlorococcus_marinus*_MIT9303
		WP_011823445.1	*Prochlorococcus_marinus*_NATL1A
		WP_011243875.1	*Synechococcus_elongatus*_PCC_7942

Table B.8 continued from previous page

N Source	Gene	Accession Number	Species
Amino Acid Biosynthesis	*aspS*	WP_011432789.1	*Synechococcus*_sp_JA-3-3Ab
Amino Acid	*bgtA*	WP_162328001.1	*Synechocystis* sp. CACIAM 05
Amino Acid	*bgtB*	WP_162327858.1	*Synechocystis* sp. CACIAM 05
Metabolism	*carA*	Ga0078150_109139	*Synechococcus* sp. 1G10
		WP_011429111.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011127891.1	*Synechococcus*_sp_WH8102
		WP_066906513.1	*Synechococcus*_sp__MIT9509
		WP_011826212.1	*Prochlorococcus_marinus*_MIT9303
		WP_011823622.1	*Prochlorococcus_marinus*_NATL1A
		WP_011378327.1	*Synechococcus_elongatus*_PCC_7942
Metabolism	*carB*	Ga0078150_10136	*Synechococcus* sp. 1G10
		WP_011430580.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011127695.1	*Synechococcus*_sp_WH8102
		WP_066905373.1	*Synechococcus*_sp__MIT9509
		WP_011826029.1	*Prochlorococcus_marinus*_MIT9303
		WP_011823594.1	*Prochlorococcus_marinus*_NATL1A
		WP_011243131.1	*Synechococcus_elongatus*_PCC_7942
Chitin	*chdA*	Ga0078150_11933	*Synechococcus* sp. 1G10
		WP_011429655.1	*Synechococcus*_sp_JA-3-3Ab
Chitin	chiA	WP_066911065.1	*Synechococcus*_sp__MIT9509
		WP_011377708.1	*Synechococcus_elongatus*_PCC_7942
Chitin	*choA*	WP_010995999.1	*Nostoc*_sp__PCC_7120
Metabolism	*cphA*	Ga0078150_11858	*Synechococcus* sp. 1G10
Metabolism	*cphB*	WP_010872518.1	*Synechocystis*_sp_PCC_6803
Nitrate/Nitrite	*cynA*	WP_011129340.1	*Synechococcus*_sp_WH8102
		WP_011378317.1	*Synechococcus_elongatus*_PCC_7942
Nitrate/Nitrite	*cynB*	WP_011129339.1	*Synechococcus*_sp_WH8102
		WP_011244297.1	*Synechococcus_elongatus*_PCC_7942
Nitrate/Nitrite	*cynD*	WP_011129338.1	*Synechococcus*_sp_WH8102
		WP_0110551680.1	*Synechococcus_elongatus*_PCC_7942
Nitrate/Nitrite	*cynH*	WP_042504028.1	*Synechococcus*_sp_WH8102
Nitrate/Nitrite	*cynS*	Ga0078150_105173	*Synechococcus* sp. 1G10
		WP_071820881.1	*Synechococcus*_sp_WH8102
		WP_074159326.1	*Synechococcus*_sp__MIT9509
		WP_011822872.1	*Prochlorococcus_marinus*_NATL1A
		WP_011244299.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*cysE*	Ga0078150_13712	*Synechococcus* sp. 1G10
		WP_011429460.1	*Synechococcus*_sp_JA-3-3Ab

Table B.8 continued from previous page

N Source	Gene	Accession Number	Species
Amino Acid Biosynthesis	*cysE*	WP_011126968.1	*Synechococcus*_sp_WH8102
		WP_066911339.1	*Synechococcus*_sp__MIT9509
		WP_011824838.1	*Prochlorococcus_marinus*_MIT9303
		WP_011824592.1	*Prochlorococcus_marinus*_NATL1A
		WP_011243996.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*cysM1*	Ga0078150_11720	*Synechococcus* sp. 1G10
		WP_011430534.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011127540.1	*Synechococcus*_sp_WH8102
		WP_066911739.1	*Synechococcus*_sp__MIT9509
		WP_011826729.1	*Prochlorococcus_marinus*_MIT9303
		WP_011822984.1	*Prochlorococcus_marinus*_NATL1A
		WP_011055164.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*cysM2*	Ga0078150_10381	*Synechococcus* sp. 1G10
		WP_011129078.1	*Synechococcus*_sp_WH8102
		WP_041375247.1	*Prochlorococcus_marinus*_MIT9303
		WP_011055156.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*dapA*	Ga0078150_10526	*Synechococcus* sp. 1G10
		WP_049749712.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011126945.1	*Synechococcus*_sp_WH8102
		WP_066908496.1	*Synechococcus*_sp__MIT9509
		WP_011824761.1	*Prochlorococcus_marinus*_MIT9303
		WP_041700641.1	*Prochlorococcus_marinus*_NATL1A
		WP_011244557.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*dapB*	Ga0078150_102107	*Synechococcus* sp. 1G10
		WP_011430162.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011127684.1	*Synechococcus*_sp_WH8102
		WP_066905391.1	*Synechococcus*_sp__MIT9509
		WP_011826039.1	*Prochlorococcus_marinus*_MIT9303
		WP_011823738.1	*Prochlorococcus_marinus*_NATL1A
		WP_011378334.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*dapF*	Ga0078150_101218	*Synechococcus* sp. 1G10
		WP_041439019.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011128118.1	*Synechococcus*_sp_WH8102
		WP_066907533.1	*Synechococcus*_sp__MIT9509
		WP_011823679.1	*Prochlorococcus_marinus*_NATL1A
		WP_011244477.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid	*dmeA*	WP_011242834.1	*Synechococcus_elongatus*_PCC_7942
Peptides	*dppA*	Ga0078146_104134	*Synechococcus* sp. BO8801

Table B.8 continued from previous page

N Source	Gene	Accession Number	Species
Peptides	*dppB*	Ga0078146_104135	*Synechococcus* sp. BO8801
Peptides	*dppC*	Ga0078146_11623	*Synechococcus* sp. BO8801
Peptides	*dppD*	Ga0078146_11132	*Synechococcus* sp. BO8801
Peptides	*dppF*	WP_185831897.1	*Pseudomonas aeruginosa*
Nitrogen Fixation	*fdxN*	WP_015081224.1	*Anabaena*_sp__90
Nitrate/Nitrite	*focA*	WP_041374994.1	*Prochlorococcus_marinus*_MIT9303
		WP_038655547.1	*Prochlorococcus_marinus*_NATL1A
Amino Acid Biosynthesis	*gatA*	WP_011430266.1	*Synechococcus*_sp_JA-3-3Ab
Amino Acid Biosynthesis	*gatB*	WP_162328642.1	*Synechocystis* sp. CACIAM 05
Amino Acid Biosynthesis	*gatC*	WP_162329184.1	*Synechocystis* sp. CACIAM 05
Metabolism	*gdhA*	WP_011826248.1	*Prochlorococcus_marinus*_MIT9303
Metabolism	*gifA*	WP_011242952.1	*Synechococcus_elongatus*_PCC_7942
		WP_028948617.1	*Synechocystis*_sp__PCC_6714
Metabolism	*gifB*	WP_010871622.1	*Synechocystis*_sp__PCC_6803
Metabolism	*glnA*	Ga0078150_101329	*Synechococcus* sp. 1G10
		WP_011430212.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011127938.1	*Synechococcus*_sp_WH8102
		WP_066906366.1	*Synechococcus*_sp__MIT9509
		WP_011826275.1	*Prochlorococcus_marinus*_MIT9303
		WP_011823483.1	*Prochlorococcus_marinus*_NATL1A
		WP_011378345.1	*Synechococcus_elongatus*_PCC_7942
Regulation	*glnB*	Ga0078150_104180	*Synechococcus* sp. 1G10
		WP_011430352.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011127336.1	*Synechococcus*_sp_WH8102
		WP_066910997.1	*Synechococcus*_sp__MIT9509
		WP_011825143.1	*Prochlorococcus_marinus*_MIT9303
		WP_011295338.1	*Prochlorococcus_marinus*_NATL1A
		WP_011243504.1	*Synechococcus_elongatus*_PCC_7942
Metabolism	*glnN*	Ga0078150_10944	*Synechococcus* sp. 1G10
		WP_011243650.1	*Synechococcus_elongatus*_PCC_7942
Metabolism	*glsF*	Ga0078150_11258	*Synechococcus* sp. 1G10
		WP_011431483.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011128988.1	*Synechococcus*_sp_WH8102
		WP_011377778.1	*Synechococcus_elongatus*_PCC_7942
Metabolism	*gltD*	WP_010871407.1	*Synechocystis*_sp__PCC_6803
Amino Acid	*gltS*	WP_162327631.1	*Synechocystis* sp. CACIAM 05
Amino Acid	*gtrA*	WP_162328652.1	*Synechocystis* sp. CACIAM 05
Amino Acid	*gtrB*	WP_162328651.1	*Synechocystis* sp. CACIAM 05

Table B.8 continued from previous page

N Source	Gene	Accession Number	Species
Amino Acid	*gtrC*	WP_162328650.1	*Synechocystis* sp. CACIAM 05
Amino Acid Biosynthesis	*glyA*	Ga0078150_107174	*Synechococcus* sp. 1G10
		WP_049749769.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011127135.1	*Synechococcus*_sp_WH8102
		WP_066905027.1	*Synechococcus*_sp__MIT9509
		WP_011827054.1	*Prochlorococcus_marinus*_MIT9303
		WP_011823107.1	*Prochlorococcus_marinus*_NATL1A
		WP_011243543.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*hisA*	Ga0078150_12437	*Synechococcus* sp. 1G10
		WP_011431306.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011127638.1	*Synechococcus*_sp_WH8102
		WP_066905638.1	*Synechococcus*_sp__MIT9509
		WP_011826375.1	*Prochlorococcus_marinus*_MIT9303
		WP_011823540.1	*Prochlorococcus_marinus*_NATL1A
		WP_011244575.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*hisB*	WP_011431413.1	*Synechococcus* sp. 1G10
		WP_011127102.1	*Synechococcus*_sp_JA-3-3Ab
		WP_066905059.1	*Synechococcus*_sp_WH8102
		WP_011827089.1	*Synechococcus*_sp__MIT9509
		WP_011294280.1	*Prochlorococcus_marinus*_MIT9303
		WP_011243691.1	*Prochlorococcus_marinus*_NATL1A
Amino Acid Biosynthesis	*hisC*	Ga0078150_106199	*Synechococcus* sp. 1G10
		WP_011430707.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011129171.1	*Synechococcus*_sp_WH8102
		WP_066909177.1	*Synechococcus*_sp__MIT9509
		WP_011827319.1	*Prochlorococcus_marinus*_MIT9303
		WP_011823043.1	*Prochlorococcus_marinus*_NATL1A
		WP_011377834.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*hisD*	Ga0078150_104207	*Synechococcus* sp. 1G10
		WP_011429494.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011127477.1	*Synechococcus*_sp_WH8102
		WP_082823413.1	*Synechococcus*_sp__MIT9509
		WP_011825110.1	*Prochlorococcus_marinus*_MIT9303
		WP_011824422.1	*Prochlorococcus_marinus*_NATL1A
		WP_011244814.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*hisF*	Ga0078150_11011	*Synechococcus* sp. 1G10
		WP_011431091.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011128537.1	*Synechococcus*_sp_WH8102

Table B.8 continued from previous page

N Source	Gene	Accession Number	Species
Amino Acid Biosynthesis	*hisF*	WP_066906807.1	*Synechococcus*_sp__MIT9509
		WP_011826678.1	*Prochlorococcus*_*marinus*_MIT9303
		WP_011823231.1	*Prochlorococcus*_*marinus*_NATL1A
		WP_011378309.1	*Synechococcus*_*elongatus*_PCC_7942
Amino Acid Biosynthesis	*hisG*	Ga0078150_10325	*Synechococcus* sp. 1G10
		WP_041438996.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011128406.1	*Synechococcus*_sp_WH8102
		WP_066905720.1	*Synechococcus*_sp__MIT9509
		WP_011130564.1	*Prochlorococcus*_*marinus*_MIT9303
		WP_011294521.1	*Prochlorococcus*_*marinus*_NATL1A
		WP_011243066.1	*Synechococcus*_*elongatus*_PCC_7942
Amino Acid Biosynthesis	*hisH*	Ga0078150_102270	*Synechococcus* sp. 1G10
		WP_011429853.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011127589.1	*Synechococcus*_sp_WH8102
		WP_011825569.1	*Prochlorococcus*_*marinus*_MIT9303
		WP_011824103.1	*Prochlorococcus*_*marinus*_NATL1A
		WP_011244777.1	*Synechococcus*_*elongatus*_PCC_7942
Amino Acid Biosynthesis	*hisIE*	Ga0078150_103105	*Synechococcus* sp. 1G10
		WP_083759185.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011128369.1	*Synechococcus*_sp_WH8102
		WP_066905658.1	*Synechococcus*_sp__MIT9509
		WP_011826456.1	*Prochlorococcus*_*marinus*_MIT9303
		WP_011823360.1	*Prochlorococcus*_*marinus*_NATL1A
		WP_011244411.1	*Synechococcus*_*elongatus*_PCC_7942
Amino Acid Biosynthesis	*hisZ*	Ga0078150_101243	*Synechococcus* sp. 1G10
		WP_011430366.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011128140.1	*Synechococcus*_sp_WH8102
		WP_066907583.1	*Synechococcus*_sp__MIT9509
		WP_011826160.1	*Prochlorococcus*_*marinus*_MIT9303
		WP_011823664.1	*Prochlorococcus*_*marinus*_NATL1A
		WP_011243038.1	*Synechococcus*_*elongatus*_PCC_7942
Metabolism	*iadA*	WP_010998063.1	*Nostoc*_sp__PCC_7120
		WP_010873103.1	*Synechocystis*_sp_PCC_6803
Regulation	*icd*	Ga0078150_11360	*Synechococcus* sp. 1G10
		WP_011430312.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011127042.1	*Synechococcus*_sp_WH8102
		WP_066905093.1	*Synechococcus*_sp__MIT9509
		WP_011827151.1	*Prochlorococcus*_*marinus*_MIT9303

Table B.8 continued from previous page

N Source	Gene	Accession Number	Species
Regulation	*icd*	WP_011824551.1	*Prochlorococcus_marinus*_NATL1A
		WP_011378158.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*ilvB*	Ga0078150_103157	*Synechococcus* sp. 1G10
		WP_011430887.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011128606.1	*Synechococcus*_sp_WH8102
		WP_074159277.1	*Synechococcus*_sp__MIT9509
		WP_011825433.1	*Prochlorococcus_marinus*_MIT9303
		WP_011823319.1	*Prochlorococcus_marinus*_NATL1A
		WP_011377461.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*ilvC*	Ga0078150_102132	*Synechococcus* sp. 1G10
		WP_011429161.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011128514.1	*Synechococcus*_sp_WH8102
		WP_066906821.1	*Synechococcus*_sp__MIT9509
		WP_011826626.1	*Prochlorococcus_marinus*_MIT9303
		WP_011824291.1	*Prochlorococcus_marinus*_NATL1A
		WP_039755692.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*ilvD*	Ga0078150_101371	*Synechococcus* sp. 1G10
		WP_041438388.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011127988.1	*Synechococcus*_sp_WH8102
		WP_066905445.1	*Synechococcus*_sp__MIT9509
		WP_011826324.1	*Prochlorococcus_marinus*_MIT9303
		WP_011823517.1	*Prochlorococcus_marinus*_NATL1A
		WP_011243210.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*ilvN*	Ga0078150_1284	*Synechococcus* sp. 1G10
		WP_011430944.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011127547.1	*Synechococcus*_sp_WH8102
		WP_066910056.1	*Synechococcus*_sp__MIT9509
		WP_011825509.1	*Prochlorococcus_marinus*_MIT9303
		WP_011294852.1	*Prochlorococcus_marinus*_NATL1A
		WP_011243982.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*kprS*	Ga0078150_1018	*Synechococcus* sp. 1G10
		WP_011431657.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011127831.1	*Synechococcus*_sp_WH8102
		WP_066910070.1	*Synechococcus*_sp__MIT9509
		WP_041375086.1	*Prochlorococcus_marinus*_MIT9303
		WP_011294759.1	*Prochlorococcus_marinus*_NATL1A
		WP_011378321.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*leuA*	Ga0078150_102283	*Synechococcus* sp. 1G10

Table B.8 continued from previous page

N Source	Gene	Accession Number	Species
Amino Acid Biosynthesis	*leuA*	WP_011431034.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011127595.1	*Synechococcus*_sp_WH8102
		WP_066909917.1	*Synechococcus*_sp__MIT9509
		WP_011825575.1	*Prochlorococcus*_*marinus*_MIT9303
		WP_011824098.1	*Prochlorococcus*_*marinus*_NATL1A
		WP_011378025.1	*Synechococcus*_*elongatus*_PCC_7942
Amino Acid Biosynthesis	*leuB*	Ga0078150_12312	*Synechococcus* sp. 1G10
		WP_011430627.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011127649.1	*Synechococcus*_sp_WH8102
		WP_066905485.1	*Synechococcus*_sp__MIT9509
		WP_011826363.1	*Prochlorococcus*_*marinus*_MIT9303
		WP_011823530.1	*Prochlorococcus*_*marinus*_NATL1A
		WP_0113780580.1	*Synechococcus*_*elongatus*_PCC_7942
Amino Acid Biosynthesis	*leuC*	Ga0078150_107171	*Synechococcus* sp. 1G10
		WP_011429196.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011127138.1	*Synechococcus*_sp_WH8102
		WP_066904744.1	*Synechococcus*_sp__MIT9509
		WP_011827051.1	*Prochlorococcus*_*marinus*_MIT9303
		WP_011823105.1	*Prochlorococcus*_*marinus*_NATL1A
		WP_011244507.1	*Synechococcus*_*elongatus*_PCC_7942
Amino Acid Biosynthesis	*leuD*	Ga0078150_107170	*Synechococcus* sp. 1G10
		WP_041438503.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011127139.1	*Synechococcus*_sp_WH8102
		WP_066904742.1	*Synechococcus*_sp__MIT9509
		WP_011827050.1	*Prochlorococcus*_*marinus*_MIT9303
		WP_011823104.1	*Prochlorococcus*_*marinus*_NATL1A
		WP_011243872.1	*Synechococcus*_*elongatus*_PCC_7942
Amino Acid Biosynthesis	*livE*	Ga0078150_101205	*Synechococcus* sp. 1G10
		WP_011429887.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011128103.1	*Synechococcus*_sp_WH8102
		WP_066907496.1	*Synechococcus*_sp__MIT9509
		WP_011826129.1	*Prochlorococcus*_*marinus*_MIT9303
		WP_011823688.1	*Prochlorococcus*_*marinus*_NATL1A
		WP_011242829.1	*Synechococcus*_*elongatus*_PCC_7942
Amino Acid Biosynthesis	*lysA*	Ga0078150_13018	*Synechococcus* sp. 1G10
		WP_011431665.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011127801.1	*Synechococcus*_sp_WH8102
		WP_066912293.1	*Synechococcus*_sp__MIT9509

Table B.8 continued from previous page

N Source	Gene	Accession Number	Species
Amino Acid Biosynthesis	*lysA*	WP_011825648.1	*Prochlorococcus_marinus*_MIT9303
		WP_011824072.1	*Prochlorococcus_marinus*_NATL1A
		WP_011243563.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*lysC*	Ga0078150_10518	*Synechococcus* sp. 1G10
		WP_011430578.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011126948.1	*Synechococcus*_sp_WH8102
		WP_066908481.1	*Synechococcus*_sp__MIT9509
		WP_011824771.1	*Prochlorococcus_marinus*_MIT9303
		WP_011824616.1	*Prochlorococcus_marinus*_NATL1A
		WP_039756002.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*metH*	Ga0078150_101204	*Synechococcus* sp. 1G10
		WP_011431208.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011128102.1	*Synechococcus*_sp_WH8102
		WP_066907770.1	*Synechococcus*_sp__MIT9509
		WP_011826128.1	*Prochlorococcus_marinus*_MIT9303
		WP_011823689.1	*Prochlorococcus_marinus*_NATL1A
		WP_011378007.1	*Synechococcus_elongatus*_PCC_7942
Metabolism	*narB*	Ga0078150_11114	*Synechococcus* sp. 1G10
		WP_011429517.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011129317.1	*Synechococcus*_sp_WH8102
		WP_066908785.1	*Synechococcus*_sp__MIT9509
		WP_011377931.1	*Synechococcus_elongatus*_PCC_7942
Metabolism	*narM*	WP_011242921.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid	*natA*	WP_162328838.1	*Synechocystis* sp. CACIAM 05
Amino Acid	*natB*	WP_162328984.1	*Synechocystis* sp. CACIAM 05
Amino Acid	*natC*	WP_162328527.1	*Synechocystis* sp. CACIAM 05
Amino Acid	*natD*	WP_162328421.1	*Synechocystis* sp. CACIAM 05
Amino Acid	*natE*	WP_162327940.1	*Synechocystis* sp. CACIAM 05
Amino Acid	*natF*	WP_044522113.1	*Nostoc*_sp__PCC_7120
Amino Acid	*natG*	WP_010998304.1	*Nostoc*_sp__PCC_7120
Amino Acid	*natH*	WP_010998305.1	*Nostoc*_sp__PCC_7120
Amino Acid	*natI*	WP_044521412.1	*Nostoc*_sp__PCC_7120
Amino Acid	*natJ*	WP_010996692.1	*Nostoc*_sp__PCC_7120
Amino Acid	*natK*	WP_010996694.1	*Nostoc*_sp__PCC_7120
Amino Acid	*natL*	WP_010996695.1	*Nostoc*_sp__PCC_7120
Amino Acid	*natM*	WP_044522999.1	*Nostoc*_sp__PCC_7120
Metabolism	*nblA*	WP_011378330.1	*Synechococcus_elongatus*_PCC_7942
Nitrogen Fixation	*nifB*	WP_011430655.1	*Synechococcus*_sp_JA-3-3Ab

Table B.8 continued from previous page

N Source	Gene	Accession Number	Species
Nitrogen Fixation	*nifD*	WP_011430650.1	*Synechococcus*_sp_JA-3-3Ab
Nitrogen Fixation	*nifE*	WP_011430661.1	*Synechococcus*_sp_JA-3-3Ab
Nitrogen Fixation	*nifH*	WP_011430651.1	*Synechococcus*_sp_JA-3-3Ab
Nitrogen Fixation	*nifK*	WP_011430649.1	*Synechococcus*_sp_JA-3-3Ab
Nitrogen Fixation	*nifN*	WP_011430662.1	*Synechococcus*_sp_JA-3-3Ab
Amino Acid Biosynthesis	*sufS*	WP_011430653.1	*Synechococcus*_sp_JA-3-3Ab
Nitrogen Fixation	*nifT*	WP_011430645.1	*Synechococcus*_sp_JA-3-3Ab
Nitrogen Fixation	*nifU*	WP_011430652.1	*Synechococcus*_sp_JA-3-3Ab
Nitrogen Fixation	*nifV*	WP_011430647.1	*Synechococcus*_sp_JA-3-3Ab
Nitrogen Fixation	*nifW*	WP_011430666.1	*Synechococcus*_sp_JA-3-3Ab
Nitrogen Fixation	*nifX*	WP_011430663.1	*Synechococcus*_sp_JA-3-3Ab
Nitrogen Fixation	*nifZ*	WP_011430646.1	*Synechococcus*_sp_JA-3-3Ab
Metabolism	*nirA*	Ga0078150_11122	*Synechococcus* sp. 1G10
		WP_011429510.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011129330.1	*Synechococcus*_sp_WH8102
		WP_066908764.1	*Synechococcus*_sp__MIT9509
		WP_011827554.1	*Prochlorococcus_marinus*_MIT9303
		WP_011824662.1	*Prochlorococcus_marinus*_NATL1A
		WP_011242624.1	*Synechococcus_elongatus*_PCC_7942
Metabolism	*nirB*	WP_011242623.1	*Synechococcus_elongatus*_PCC_7942
Nitrate/Nitrite	*nrtA*	Ga0078150_11119	*Synechococcus* sp. 1G10
		WP_011429511.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011242625.1	*Synechococcus_elongatus*_PCC_7942
Nitrate/Nitrite	*nrtB*	Ga0078150_11118	*Synechococcus* sp. 1G10
		WP_011429512.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011242626.1	*Synechococcus_elongatus*_PCC_7942
Nitrate/Nitrite	*nrtC*	Ga0078150_11117	*Synechococcus* sp. 1G10
		WP_011429513.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011242627.1	*Synechococcus_elongatus*_PCC_7942
Nitrate/Nitrite	*nrtD*	Ga0078150_11116	*Synechococcus* sp. 1G10
		WP_011429514.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011242628.1	*Synechococcus_elongatus*_PCC_7942
Nitrate/Nitrite	*nrtP*	WP_049692707.1	*Synechococcus*_sp_WH8103
		WP_066908788.1	*Synechococcus*_sp__MIT9509
Nitrate/Nitrite	*nrtS*	WP_081422186.1	*Synechococcus_elongatus*_PCC_6301
Regulation	*ntcA*	Ga0078150_107159	*Synechococcus* sp. 1G10
		WP_011430628.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011127149.1	*Synechococcus*_sp_WH8102

Table B.8 continued from previous page

N Source	Gene	Accession Number	Species
Regulation	*ntcA*	WP_066904724.1	*Synechococcus*_sp__MIT9509
		WP_041374935.1	*Prochlorococcus*_*marinus*_MIT9303
		WP_011823097.1	*Prochlorococcus*_*marinus*_NATL1A
		WP_011243689.1	*Synechococcus*_*elongatus*_PCC_7942
Regulation	*ntcB*	WP_011429866.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011242622.1	*Synechococcus*_*elongatus*_PCC_7942
Peptides	*oppC*	WP_011243384.1	*Synechococcus*_*elongatus*_PCC_7942
Amino Acid Biosynthesis	*pheA*	Ga0078150_11268	*Synechococcus* sp. 1G10
		WP_011430306.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011128998.1	*Synechococcus*_sp_WH8102
		WP_066905003.1	*Synechococcus*_sp__MIT9509
		WP_041375214.1	*Prochlorococcus*_*marinus*_MIT9303
		WP_011824473.1	*Prochlorococcus*_*marinus*_NATL1A
		WP_011377776.1	*Synechococcus*_*elongatus*_PCC_7942
Glyphosate	*phnD*	Ga0078150_12146	*Synechococcus* sp. 1G10
		WP_011128035.1	*Synechococcus*_sp_WH8102
		WP_066905322.1	*Synechococcus*_sp__MIT9509
		WP_063395008.1	*Prochlorococcus*_*marinus*_MIT9303
		WP_011823444.1	*Prochlorococcus*_*marinus*_NATL1A
Regulation	*pipX*	Ga0078150_11082	*Synechococcus* sp. 1G10
		WP_011430860.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011127526.1	*Synechococcus*_sp_WH8102
		WP_066911974.1	*Synechococcus*_sp__MIT9509
		WP_011129584.1	*Prochlorococcus*_*marinus*_MIT9303
		WP_041700531.1	*Prochlorococcus*_*marinus*_NATL1A
		WP_011244342.1	*Synechococcus*_*elongatus*_PCC_7942
Amino Acid Biosynthesis	*proA*	Ga0078150_101135	*Synechococcus* sp. 1G10
		WP_011431592.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011128350.1	*Synechococcus*_sp_WH8102
		WP_066908316.1	*Synechococcus*_sp__MIT9509
		WP_011826472.1	*Prochlorococcus*_*marinus*_MIT9303
		WP_011823371.1	*Prochlorococcus*_*marinus*_NATL1A
		WP_011244143.1	*Synechococcus*_*elongatus*_PCC_7942
Amino Acid Biosynthesis	*proB*	Ga0078150_12310	*Synechococcus* sp. 1G10
		WP_011430991.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011127647.1	*Synechococcus*_sp_WH8102
		WP_066905491.1	*Synechococcus*_sp__MIT9509
		WP_011826365.1	*Prochlorococcus*_*marinus*_MIT9303

Table B.8 continued from previous page

N Source	Gene	Accession Number	Species
Amino Acid Biosynthesis	*proB*	WP_011823532.1	*Prochlorococcus_marinus*_NATL1A
		WP_011378361.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*proC*	Ga0078150_11079	*Synechococcus* sp. 1G10
		WP_011429750.1	*Synechococcus*_sp_JA-3-3Ab
		WP_042503088.1	*Synechococcus*_sp_WH8102
		WP_066912025.1	*Synechococcus*_sp__MIT9509
		WP_011826755.1	*Prochlorococcus_marinus*_MIT9303
		WP_011823199.1	*Prochlorococcus_marinus*_NATL1A
		WP_011378294.1	*Synechococcus_elongatus*_PCC_7942
Metabolism	*puo*	WP_020908634.1	*Rhodococcus_erythropolis*_PR4
Metabolism	*putA*	Ga0078150_11094	*Synechococcus* sp. 1G10
		WP_011128814.1	*Synechococcus*_sp_WH8102
		WP_066912011.1	*Synechococcus*_sp__MIT9509
		WP_011826780.1	*Prochlorococcus_marinus*_MIT9303
		WP_011823178.1	*Prochlorococcus_marinus*_NATL1A
		WP_011243342.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*serA*	Ga0078150_104117	*Synechococcus* sp. 1G10
		WP_011430204.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011127403.1	*Synechococcus*_sp_WH8102
		WP_066911197.1	*Synechococcus*_sp__MIT9509
		WP_011825199.1	*Prochlorococcus_marinus*_MIT9303
		WP_011824331.1	*Prochlorococcus_marinus*_NATL1A
		WP_039755951.1	*Synechococcus_elongatus*_PCC_7942
Metabolism	*speA*	Ga0078150_1358	*Synechococcus* sp. 1G10
		WP_011429050.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011129212.1	*Synechococcus*_sp_WH8102
		WP_066912242.1	*Synechococcus*_sp__MIT9509
		WP_011827403.1	*Prochlorococcus_marinus*_MIT9303
		WP_011822861.1	*Prochlorococcus_marinus*_NATL1A
		WP_049749326.1	*Synechococcus_elongatus*_PCC_7942
Metabolism	*speB*	Ga0078150_11431	*Synechococcus* sp. 1G10
		WP_011429747.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011129275.1	*Synechococcus*_sp_WH8102
		WP_066912234.1	*Synechococcus*_sp__MIT9509
		WP_011827519.1	*Prochlorococcus_marinus*_MIT9303
		WP_011824651.1	*Prochlorococcus_marinus*_NATL1A
Metabolism	*syNOS*	WP_006458277.1	*Synechococcus*_sp_PCC_7335
Amino Acid Biosynthesis	*thrA*	Ga0078150_102257	*Synechococcus* sp. 1G10

Table B.8 continued from previous page

N Source	Gene	Accession Number	Species
Amino Acid Biosynthesis	*thrA*	WP_011429972.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011127578.1	*Synechococcus*_sp_WH8102
		WP_066909980.1	*Synechococcus*_sp__MIT9509
		WP_011825546.1	*Prochlorococcus*_*marinus*_MIT9303
		WP_041700607.1	*Prochlorococcus*_*marinus*_NATL1A
		WP_011378311.1	*Synechococcus*_*elongatus*_PCC_7942
Amino Acid Biosynthesis	*thrB*	Ga0078150_101140	*Synechococcus* sp. 1G10
		WP_011430460.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011128345.1	*Synechococcus*_sp_WH8102
		WP_066908295.1	*Synechococcus*_sp__MIT9509
		WP_011826485.1	*Prochlorococcus*_*marinus*_MIT9303
		WP_011823375.1	*Prochlorococcus*_*marinus*_NATL1A
		WP_011242429.1	*Synechococcus*_*elongatus*_PCC_7942
Amino Acid Biosynthesis	*thrC1*	Ga0078150_12321	*Synechococcus* sp. 1G10
		WP_011429208.1	*Synechococcus*_sp_JA-3-3Ab
		WP_042504535.1	*Synechococcus*_sp_WH8102
		WP_066908920.1	*Synechococcus*_sp__MIT9509
		WP_011827591.1	*Prochlorococcus*_*marinus*_MIT9303
		WP_011824690.1	*Prochlorococcus*_*marinus*_NATL1A
		WP_011243808.1	*Synechococcus*_*elongatus*_PCC_7942
Amino Acid Biosynthesis	*thrC2*	Ga0078150_10594	*Synechococcus* sp. 1G10
		WP_011431249.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011244621.1	*Synechococcus*_*elongatus*_PCC_7942
Amino Acid Biosynthesis	*trpA*	Ga0078150_103124	*Synechococcus* sp. 1G10
		WP_011429138.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011128383.1	*Synechococcus*_sp_WH8102
		WP_066905676.1	*Synechococcus*_sp__MIT9509
		WP_041375164.1	*Prochlorococcus*_*marinus*_MIT9303
		WP_011823354.1	*Prochlorococcus*_*marinus*_NATL1A
		WP_011243417.1	*Synechococcus*_*elongatus*_PCC_7942
Amino Acid Biosynthesis	*trpB*	Ga0078150_11250	*Synechococcus* sp. 1G10
		WP_041438793.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011129133.1	*Synechococcus*_sp_WH8102
		WP_066909084.1	*Synechococcus*_sp__MIT9509
		WP_041375255.1	*Prochlorococcus*_*marinus*_MIT9303
		WP_011823024.1	*Prochlorococcus*_*marinus*_NATL1A
		WP_011244258.1	*Synechococcus*_*elongatus*_PCC_7942
Amino Acid Biosynthesis	*trpC*	Ga0078150_102158	*Synechococcus* sp. 1G10

Table B.8 continued from previous page

N Source	Gene	Accession Number	Species
Amino Acid Biosynthesis	*trpC*	WP_011431168.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011128493.1	*Synechococcus*_sp_WH8102
		WP_066906869.1	*Synechococcus*_sp__MIT9509
		WP_011826600.1	*Prochlorococcus_marinus*_MIT9303
		WP_011824273.1	*Prochlorococcus_marinus*_NATL1A
		WP_011242666.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*trpD*	Ga0078150_109138	*Synechococcus* sp. 1G10
		WP_011429527.1	*Synechococcus*_sp_JA-3-3Ab
		WP_042504263.1	*Synechococcus*_sp_WH8102
		WP_066906640.1	*Synechococcus*_sp__MIT9509
		WP_041374820.1	*Prochlorococcus_marinus*_MIT9303
		WP_011823621.1	*Prochlorococcus_marinus*_NATL1A
		WP_011378328.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*trpE*	Ga0078150_10680	*Synechococcus* sp. 1G10
		WP_011430950.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011128903.1	*Synechococcus*_sp_WH8102
		WP_066904381.1	*Synechococcus*_sp__MIT9509
		WP_011826875.1	*Prochlorococcus_marinus*_MIT9303
		WP_011824531.1	*Prochlorococcus_marinus*_NATL1A
		WP_011377825.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*trpF*	Ga0078150_10360	*Synechococcus* sp. 1G10
		WP_011429662.1	*Synechococcus*_sp_JA-3-3Ab
		WP_042503748.1	*Synechococcus*_sp_WH8102
		WP_066905776.1	*Synechococcus*_sp__MIT9509
		WP_011825451.1	*Prochlorococcus_marinus*_MIT9303
		WP_011823328.1	*Prochlorococcus_marinus*_NATL1A
		WP_011377568.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*trpG*	Ga0078150_106197	*Synechococcus* sp. 1G10
		WP_011431437.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011129168.1	*Synechococcus*_sp_WH8102
		WP_066909171.1	*Synechococcus*_sp__MIT9509
		WP_011827315.1	*Prochlorococcus_marinus*_MIT9303
		WP_011823042.1	*Prochlorococcus_marinus*_NATL1A
		WP_011243427.1	*Synechococcus_elongatus*_PCC_7942
Amino Acid Biosynthesis	*tyrA*	Ga0078150_10694	*Synechococcus* sp. 1G10
		WP_011431107.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011128916.1	*Synechococcus*_sp_WH8102
		WP_066904989.1	*Synechococcus*_sp__MIT9509

Table B.8 continued from previous page

N Source	Gene	Accession Number	Species
Amino Acid Biosynthesis	*tyrA*	WP_011826897.1	*Prochlorococcus_marinus*_MIT9303
		WP_011824521.1	*Prochlorococcus_marinus*_NATL1A
		WP_039755844.1	*Synechococcus_elongatus*_PCC_7942
Metabolism	*ureA*	Ga0078150_11737	*Synechococcus* sp. 1G10
		WP_011129300.1	*Synechococcus*_sp_WH8102
		WP_066908813.1	*Synechococcus*_sp__MIT9509
		WP_011827547.1	*Prochlorococcus_marinus*_MIT9303
		WP_011824456.1	*Prochlorococcus_marinus*_NATL1A
Metabolism	*ureB*	Ga0078150_11738	*Synechococcus* sp. 1G10
		WP_011129301.1	*Synechococcus*_sp_WH8102
		WP_066908810.1	*Synechococcus*_sp__MIT9509
		WP_011827548.1	*Prochlorococcus_marinus*_MIT9303
		WP_011295399.1	*Prochlorococcus_marinus*_NATL1A
Metabolism	*ureC*	Ga0078150_11739	*Synechococcus* sp. 1G10
		WP_011429503.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011129302.1	*Synechococcus*_sp_WH8102
		WP_066908807.1	*Synechococcus*_sp__MIT9509
		WP_011827549.1	*Prochlorococcus_marinus*_MIT9303
		WP_011824457.1	*Prochlorococcus_marinus*_NATL1A
Metabolism	*ureD*	Ga0078150_11736	*Synechococcus* sp. 1G10
		WP_011429645.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011129299.1	*Synechococcus*_sp_WH8102
		WP_066908816.1	*Synechococcus*_sp__MIT9509
		WP_011827546.1	*Prochlorococcus_marinus*_MIT9303
		WP_011824455.1	*Prochlorococcus_marinus*_NATL1A
Metabolism	*ureE*	Ga0078150_11735	*Synechococcus* sp. 1G10
		WP_011429504.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011129298.1	*Synechococcus*_sp_WH8102
		WP_066908961.1	*Synechococcus*_sp__MIT9509
		WP_063395030.1	*Prochlorococcus_marinus*_MIT9303
		WP_011824454.1	*Prochlorococcus_marinus*_NATL1A
Metabolism	*ureF*	Ga0078150_11734	*Synechococcus* sp. 1G10
		WP_011429505.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011129297.1	*Synechococcus*_sp_WH8102
		WP_066908820.1	*Synechococcus*_sp__MIT9509
		WP_011827544.1	*Prochlorococcus_marinus*_MIT9303
		WP_011824453.1	*Prochlorococcus_marinus*_NATL1A
Metabolism	*ureG*	Ga0078150_11733	*Synechococcus* sp. 1G10

Table B.8 continued from previous page

N Source	Gene	Accession Number	Species
Metabolism	*ureG*	WP_049749711.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011129296.1	*Synechococcus*_sp_WH8102
		WP_066908823.1	*Synechococcus*_sp__MIT9509
		WP_011827543.1	*Prochlorococcus*_*marinus*_MIT9303
		WP_011824452.1	*Prochlorococcus*_*marinus*_NATL1A
Metabolism	*ureJ*	Ga0078150_101169	*Synechococcus* sp. 1G10
		WP_049749734.1	*Synechococcus*_sp_JA-3-3Ab
		WP_042503880.1	*Synechococcus*_sp_WH8102
		WP_041374886.1	*Prochlorococcus*_*marinus*_MIT9303
		WP_041700597.1	*Prochlorococcus*_*marinus*_NATL1A
		WP_011242497.1	*Synechococcus*_*elongatus*_PCC_7942
Urea	*urtA*	Ga0078150_11732	*Synechococcus* sp. 1G10
		WP_011430744.1	*Synechococcus*_sp_JA-3-3Ab
		WP_042504522.1	*Synechococcus*_sp_WH8102
		WP_066908827.1	*Synechococcus*_sp__MIT9509
		WP_011827542.1	*Prochlorococcus*_*marinus*_MIT9303
		WP_011824451.1	*Prochlorococcus*_*marinus*_NATL1A
Urea	*urtB*	Ga0078150_11731	*Synechococcus* sp. 1G10
		WP_011430743.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011129294.1	*Synechococcus*_sp_WH8102
		WP_066908830.1	*Synechococcus*_sp__MIT9509
		WP_011827541.1	*Prochlorococcus*_*marinus*_MIT9303
		WP_011824450.1	*Prochlorococcus*_*marinus*_NATL1A
Urea	*urtC*	Ga0078150_11730	*Synechococcus* sp. 1G10
		WP_011430742.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011129293.1	*Synechococcus*_sp_WH8102
		WP_066908834.1	*Synechococcus*_sp__MIT9509
		WP_011827540.1	*Prochlorococcus*_*marinus*_MIT9303
		WP_011824449.1	*Prochlorococcus*_*marinus*_NATL1A
Urea	*urtD*	Ga0078150_11729	*Synechococcus* sp. 1G10
		WP_011430741.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011129292.1	*Synechococcus*_sp_WH8102
		WP_066908837.1	*Synechococcus*_sp__MIT9509
		WP_011827539.1	*Prochlorococcus*_*marinus*_MIT9303
		WP_011824448.1	*Prochlorococcus*_*marinus*_NATL1A
Urea	*urtE*	Ga0078150_11728	*Synechococcus* sp. 1G10
		WP_011430740.1	*Synechococcus*_sp_JA-3-3Ab
		WP_011129291.1	*Synechococcus*_sp_WH8102

Table B.8 continued from previous page

N Source	Gene	Accession Number	Species
Urea	*urtE*	WP_066908840.1	*Synechococcus*_sp__MIT9509
		WP_041375289.1	*Prochlorococcus_marinus*_MIT9303
		WP_011824447.1	*Prochlorococcus_marinus*_NATL1A

Table B.9: **Prevalence of inorganic nitrogen assmilatory genes in cyanobacteria**. F-Pcy: freshwater picocyanobacteria; SE-Cy: *Synechococcus elongatus* strains; T-Pcy: thermal picocyanobacteria; M-Pcy: marine picocyanobacteria; LF-Cy: larger freshwater cyanobacteria.

Group	**Genome**	***amt1***	***amt2***	***amt3***	***amt4***	***amtB***	***nrtABCD***	***nrtP***	***focA***	***nrtS***	***cynABD***	***cynS***	***cynH***	***nifD***	***nifK***	***fdxN***	***nifU***	***nifB***	***nifE***	***nifH***	***nifN***	***nifX***	***nifW***	***nifV***	***nifZ***	***nifT***
F-Pcy	*Cyanobium gracile* 6307	+	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
F-Pcy	*Cyanobium usitatum* Tous	+	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
F-Pcy	*Synechococcus* sp. 8F6	+	-	-	-	-	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
F-Pcy	*Synechococcus* sp. 6H9	+	-	-	-	-	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
F-Pcy	*Vulcanococcus limneticus* 3B3	+	-	-	-	+	+	-	-	-	-	+	-	+	+	-	+	+	+	+	+	-	-	+	+	+
F-Pcy	*Synechococcus* sp. BO8801	+	-	-	-	+	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
F-Pcy	*Synechococcus* sp. 1G10	+	-	-	-	+	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
F-Pcy	*Synechococcus* lacustris Tous	+	-	-	-	+	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
F-Pcy	*Synechococcus* sp. 4A10	+	-	-	-	+	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
F-Pcy	*Synechococcus* sp. MW101C3	+	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
F-Pcy	*Synechococcus* sp. FACHB-909	+	-	-	-	+	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
F-Pcy	*Synechococcus* sp. CCAP1479/10	+	-	-	-	+	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
F-Pcy	*Synechococcus* sp. CCAP1479/13	+	-	-	-	+	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
F-Pcy	*Synechococcus* sp. CCAP1479/9	+	-	-	-	+	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
F-Pcy	*Synechococcus* sp. CCY0621	+	-	-	-	+	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
F-Pcy	*Synechococcus* sp. CCY9618	+	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
F-Pcy	*Vulcanococcus limneticus* LL	+	-	-	-	+	+	-	-	-	-	+	-	+	+	-	+	+	+	+	+	-	-	+	+	+
SE-Cy	*Synechococcus* elongatus FACHB-1061	+	-	-	-	+	+	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SE-Cy	*Synechococcus* elongatus FACHB-242	+	-	-	-	+	+	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SE-Cy	*Synechococcus* elongatus PCC 6301	+	-	-	-	+	+	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SE-Cy	*Synechococcus* elongatus PCC 7942	+	-	-	-	+	+	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SE-Ccy	*Synechococcus* elongatus UTEX 2973	+	-	-	-	+	+	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SE-Cy	*Synechococcus* sp. PCC 6312	+	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SE-Cy	*Synechococcus* sp. PCC 7502	-	-	-	+	+	+	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SE-Cy	*Thermosynechococcus elongatus* BP-1	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SE-Cy	*Thermosynechococcus* sp. NK55a	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T-Pcy	*Synechococcus* sp. JA-2-3Ba(2-13)	+	-	-	-	-	+	-	-	-	-	-	-	+	+	-	+	+	+	+	+	+	+	+	+	+
T-Pcy	*Synechococcus* sp. JA-3-3Ab	+	-	-	-	-	+	-	-	-	-	-	-	+	+	-	+	+	+	+	+	+	+	+	+	+

Table B.9 continued from previous page

Group	Genome	*amt1*	*amt2*	*amt3*	*amt4*	*amtB*	*nrtABCD*	*nrtP*	*focA*	*nrtS*	*cynABD*	*cynS*	*cynH*	*nifD*	*nifK*	*fdxN*	*nifU*	*nifB*	*nifE*	*nifH*	*nifN*	*nifX*	*nifW*	*nifV*	*nifZ*	*nifT*
T-Pcy	*Synechococcus* sp. 60AY4M2	+	-	-	-	-	+	-	-	-	-	-	-	+	+	-	+	+	+	+	+	+	+	+	+	+
T-Pcy	*Synechococcus* sp. 63AY4M1	+	-	-	-	-	+	-	-	-	-	-	-	+	+	-	+	+	+	+	+	+	+	+	+	+
T-Pcy	*Synechococcus* sp. 63AY4M2	+	-	-	-	-	+	-	-	-	-	-	-	+	+	-	+	+	+	+	+	+	+	+	+	+
T-Pcy	*Synechococcus* sp. 65AY640	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T-Pcy	*Synechococcus* sp. 65AY6A5	+	-	-	-	-	+	-	-	-	-	-	-	+	+	-	+	+	+	+	+	+	+	+	+	+
T-Pcy	*Synechococcus* sp. 65AY6Li	+	-	-	-	-	+	-	-	-	-	-	-	+	+	-	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Cyanobium* sp. NIES-981	+	-	-	-	-	+	+	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Cyanobium* sp. PCC 7001	+	-	-	-	-	-	+	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus marinus* NATL2A	+	-	-	-	-	-	-	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus marinus* MIT9312	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus marinus* MIT9515	+	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus marinus* NATL1A	+	-	-	-	-	-	-	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus* sp. MIT0801	+	-	-	-	-	-	-	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus* sp. MIT1306	+	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus* sp. P1363	+	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus* sp. SS52	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus marinus* MIT9211	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus marinus* MIT9303	+	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus marinus* MIT9313	+	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus marinus* CCMP1375	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus marinus* AS9601	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus marinus* MIT9215	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus marinus* MIT9301	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus marinus* MIT9302	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus marinus* pastoris CCMP1986	+	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus* sp. MIT0604	+	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus* sp. RS50	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus* sp. RS01	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus* sp. RS04	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. NIES970	+	-	-	-	-	+	+	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. NKBG15041c	+	-	-	-	-	+	+	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table B.9 continued from previous page

Group	Genome	*amt1*	*amt2*	*amt3*	*amt4*	*amtB*	*nrtABCD*	*nrtP*	*focA*	*nrtS*	*cynABD*	*cynS*	*cynH*	*nifD*	*nifK*	*fdxN*	*nifU*	*nifB*	*nifE*	*nifH*	*nifN*	*nifX*	*nifW*	*nifV*	*nifZ*	*nifT*
M-Pcy	*Synechococcus* sp. PCC 7002	+	-	+	-	-	-	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. OG1	+	-	+	-	-	+	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. PCC 7003	+	-	+	-	-	-	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. PCC 7117	+	-	+	-	-	-	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. PCC 11901	+	-	+	-	-	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. NKBG042902	+	-	+	-	-	-	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. PCC 8807	+	-	+	-	-	-	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. PCC 73109	+	-	+	-	-	-	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. BDU 130192	+	-	+	-	-	-	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. CB0101	+	-	-	-	+	-	+	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. CC9311	+	-	-	-	-	-	+	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. CC9605	+	-	-	-	-	-	+	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. CC9902	+	-	-	-	-	-	+	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. KORDI-100	+	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. KORDI-49	+	-	-	-	-	-	+	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. KORDI-52	+	-	-	-	-	-	+	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. RCC307	+	-	-	-	-	-	+	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. BL107	+	-	-	-	-	-	+	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. BS55D	+	-	-	-	+	-	+	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. BS56D	+	-	-	-	+	-	+	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. CC9616	+	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. GEYO	+	-	-	-	+	-	+	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. HB1133	+	-	-	-	-	-	+	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. MITS9504	+	-	-	-	-	-	+	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. MITS9508	+	-	-	-	+	-	+	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. MITS9509	+	-	-	-	-	-	+	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. N26	+	-	-	-	-	-	+	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. PCC 7335	+	-	-	-	-	-	+	-	+	+	+	-	+	+	-	+	+	+	+	-	+	-	+	+	+
M-Pcy	*Synechococcus* sp. PCC 7336	+	-	-	-	-	-	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. RS9916	+	-	-	-	-	-	+	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. RS9917	+	-	-	-	-	-	-	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-

Table B.9 continued from previous page

Group	Genome	*amt1*	*amt2*	*amt3*	*amt4*	*amtB*	*nrtABCD*	*nrtP*	*focA*	*nrtS*	*cynABD*	*cynS*	*cynH*	*nifD*	*nifK*	*fdxN*	*nifU*	*nifB*	*nifE*	*nifH*	*nifN*	*nifX*	*nifW*	*nifV*	*nifZ*	*nifT*
M-Pcy	*Synechococcus* sp. RSCCF101	+	-	-	-	-	-	+	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. UW105	+	-	-	-	-	-	+	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. UW140	+	-	-	-	-	-	+	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. UW179A	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. UW179B	+	-	-	-	+	-	+	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. UW69	+	-	-	-	-	-	+	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. UW86	+	-	-	-	-	-	+	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. WH5701	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. WH7805	+	-	-	-	-	-	+	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. WH8016	+	-	-	-	+	-	+	+	-	-	+	+	-	-	-	-	-	-	+	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. XM-24	+	-	-	-	+	-	+	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. YX04-3	+	-	-	-	-	-	+	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. A15-127	+	-	-	-	-	-	+	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. A15-44	+	-	-	-	-	-	+	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. BMK-MC-1	+	-	-	-	-	-	+	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. LTW-R	+	-	-	-	+	-	+	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. NOUM97013	+	-	-	-	-	-	+	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. SYN20	+	-	-	-	-	-	+	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. SynAce01	+	-	-	-	-	-	+	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. WH7803	+	-	-	-	-	-	+	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. WH8020	+	-	-	-	-	-	+	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. WH8101	+	-	-	-	-	-	+	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. WH8102	+	-	-	-	-	-	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. WH8103	+	-	-	-	-	-	+	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. WH8109	+	-	-	-	-	-	+	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
LF-Cy	*Anabaena* sp. 90	-	-	-	+	+	+	+	-	-	-	-	-	+	+	+	+	+	+	-	+	+	-	+	+	+
LF-Cy	*Anabaena* sp. FACHB-1391	-	-	-	+	+	+	+	-	-	-	-	-	+	+	+	+	+	+	+	+	+	-	+	+	+
LF-Cy	*Anabaena* sp. WA102	-	-	-	+	+	+	-	-	-	-	-	-	+	+	+	+	+	+	-	+	+	-	+	+	+
LF-Cy	*Aphanizomenon flos-aquae* FACHB-1287	-	-	-	-	+	+	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	-	+	+	+
LF-Cy	*Aphanizomenon flos-aquae* LD13	-	-	-	-	+	+	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	-	+	+	+
LF-Cy	*Aphanizomenon flos-aquae* NIES-81	-	-	-	+	+	+	-	-	-	-	-	-	+	+	+	+	+	+	-	+	+	-	+	+	+

Table B.9 continued from previous page

Group	Genome	*amt1*	*amt2*	*amt3*	*amt4*	*amtB*	*nrtABCD*	*nrtP*	*focA*	*nrtS*	*cynABD*	*cynS*	*cynH*	*nifD*	*nifK*	*fdxN*	*nifU*	*nifB*	*nifE*	*nifH*	*nifN*	*nifX*	*nifW*	*nifV*	*nifZ*	*nifT*
LF-Cy	*Aphanizomenon* sp. FACHB-1416	-	-	-	-	+	+	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	-	+	+	+
LF-Cy	*Arthrospira platensis* C1	+	-	-	-	-	+	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LF-Cy	*Arthrospira platensis* FACHB-971	+	-	-	-	-	+	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LF-Cy	*Cylindrospermopsis raciborskii* MVCC14	+	-	-	-	+	+	-	-	-	-	+	-	+	+	+	+	+	+	+	+	+	-	+	+	+
LF-Cy	*Cylindrospermopsis raciborskii* CS-505	+	-	-	-	-	+	-	-	-	-	+	-	+	+	+	+	+	+	+	+	+	-	-	+	+
LF-Cy	*Cylindrospermopsis raciborskii* ITEP-A1	+	-	-	-	+	+	-	-	-	-	+	-	+	+	+	+	+	+	+	+	+	-	+	+	+
LF-Cy	*Cylindrospermopsis raciborskii* S07	+	-	-	-	-	+	-	-	-	-	+	-	+	+	+	+	+	+	+	+	+	-	-	+	+
LF-Cy	*Dolichospermum* sp. UHCC 0352	-	-	-	+	+	+	+	-	-	-	-	-	+	+	+	+	+	+	-	+	+	-	+	+	+
LF-Cy	*Dolichospermum compactum* NIES-806	-	-	-	+	+	+	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	-	+	+	+
LF-Cy	*Dolichospermum* sp. FACHB-1091	-	-	-	+	+	+	+	-	-	-	-	-	+	+	+	+	+	+	+	+	+	-	+	+	+
LF-Cy	*Geminocystis herdmanii* PCC 6308	+	-	+	-	-	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LF-Cy	*Geminocystis* sp. NIES-3708	+	-	+	-	-	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LF-Cy	*Geminocystis* sp. NIES-3709	+	-	+	-	+	+	-	-	+	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LF-Cy	*Gloeocapsa* sp. PCC 73106	+	-	-	-	-	+	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LF-Cy	*Gloeocapsa* sp. PCC 7428	+	-	-	-	-	+	-	-	+	+	+	-	-	-	-	-	-	-	-	-	+	-	-	-	+
LF-Cy	*Microcystis aeruginosa* KLA2	+	-	-	-	-	+	-	-	+	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LF-Cy	*Microcystis aeruginosa* NIES-2481	+	-	-	-	+	+	-	-	+	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LF-Cy	*Microcystis aeruginosa* PCC 7005	+	-	-	-	+	+	-	-	+	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LF-Cy	*Microcystis aeruginosa* PCC 7806SL	+	-	-	-	+	+	-	-	+	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LF-Cy	*Microcystis panniformis* FACHB-1757	+	-	-	-	+	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LF-Cy	*Microcystis* sp. 0824	+	-	-	-	+	+	-	-	+	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LF-Cy	*Microcystis* sp. MC19	+	-	-	-	+	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LF-Cy	*Nostoc paludosum* FACHB-159	+	-	-	+	+	+	+	-	-	-	+	-	+	+	+	+	+	+	+	+	+	-	+	+	+
LF-Cy	*Nostoc* sp. PCC 7107	+	-	-	-	+	+	+	-	-	-	-	-	+	+	+	+	+	+	+	+	+	-	+	+	+
LF-Cy	*Nostoc* sp. PCC 7120	+	-	-	+	+	+	-	-	+	-	+	-	+	+	+	+	+	+	+	+	-	-	+	+	+
LF-Cy	*Nostoc* sp. PCC 7524	+	-	-	+	+	+	-	-	+	-	+	-	+	-	+	+	+	+	+	+	+	-	+	+	+
LF-Cy	*Nostoc sphaeroides* Kutzing En	+	-	-	-	+	+	+	-	-	-	-	-	+	+	+	+	+	+	+	+	+	-	+	+	+
LF-Cy	*Planktothricoides* sp. FACHB-1261	+	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LF-Cy	*Planktothricoides* sp. FACHB-1370	+	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LF-Cy	*Planktothricoides* sp. SR001	+	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LF-Cy	*Planktothrix prolifica* NIVA-CYA 406	+	-	+	-	+	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table B.9 continued from previous page

Group	**Genome**	***amt1***	***amt2***	***amt3***	***amt4***	***amtB***	***nrtABCD***	***nrtP***	***focA***	***nrtS***	***cynABD***	***cynS***	***cynH***	***nifD***	***nifK***	***fdxN***	***nifU***	***nifB***	***nifE***	***nifH***	***nifN***	***nifX***	***nifW***	***nifV***	***nifZ***	***nifT***
LF-Cy	*Planktothrix agardhii* NIES-204	+	-	+	-	-	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LF-Cy	*Planktothrix agardhii* NIVA-CYA 1268	+	-	+	-	-	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LF-Cy	*Planktothrix rubescens* NIVA-CYA 407	+	-	+	-	-	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LF-Cy	*Planktothrix* sp. PCC 11201	-	-	+	-	+	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LF-Cy	*Pleurocapsa* sp. PCC 7319	+	-	-	-	-	-	+	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LF-Cy	*Pleurocapsa* sp. PCC 7327	+	-	-	-	-	+	-	-	-	-	-	-	+	+	-	+	+	+	+	+	-	-	+	+	+
LF-Cy	*Prochlorothrix hollandica* PCC 9006	+	-	+	-	-	+	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LF-Cy	*Pseudanabaena* sp. ABRG5-3	+	-	-	+	+	+	-	-	-	+	+	-	+	+	-	+	+	+	+	-	+	-	+	+	+
LF-Cy	*Pseudanabaena* sp. FACHB-723	+	-	-	-	-	+	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LF-Cy	*Pseudanabaena* sp. PCC 6802	+	-	-	-	+	+	-	-	-	+	+	-	+	+	-	+	+	+	+	+	+	-	+	+	+
LF-Cy	*Synechocystis* sp. FACHB-929	+	+	+	-	-	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LF-Cy	*Synechocystis* sp. PCC 6714	+	+	+	-	-	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LF-Cy	*Synechocystis* sp. PCC 6803	+	+	+	-	-	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LF-Cy	*Synechocystis* sp. CACIAM 05	+	+	+	-	-	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table B.10: **Prevalence of amino acid transporter genes in cyanobacteria**. F-Pcy: freshwater picocyanobacteria; SE-Cy: *Synechococcus elongatus* strains; T-Pcy: thermal picocyanobacteria; M-Pcy: marine picocyanobacteria; LF-Cy: larger freshwater cyanobacteria.

Group	**Genome**	***natA***	***natB***	***natC***	***natD***	***natE***	***natF***	***natG***	***natH***	***bgtA***	***bgtB***	***natI***	***natJ***	***natK***	***natL***	***natM***	***gtrA***	***gtrB***	***gtrC***	***gltS***	***dmeA***	***agcS***
F-Pcy	*Cyanobium gracile* 6307	-	-	-	-	-	+	+	+	+	-	+	+	+	+	+	-	-	-	+	-	-
F-Pcy	*Cyanobium usitatum* Tous	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-
F-Pcy	*Synechococcus* sp. 8F6	-	-	-	-	-	+	+	+	+	-	+	+	+	+	+	-	-	-	-	-	-
F-Pcy	*Synechococcus* sp. 6H9	-	-	-	-	-	+	+	+	+	-	+	+	+	+	+	-	-	-	-	-	-
F-Pcy	*Vulcanococcus limneticus* 3B3	-	-	-	-	-	+	+	+	+	-	+	+	+	+	+	-	-	-	+	-	-
F-Pcy	*Synechococcus* sp. BO8801	-	-	-	-	-	+	+	+	+	-	+	+	+	+	+	-	-	-	+	-	-
F-Pcy	*Synechococcus* sp. 1G10	-	-	-	-	-	+	+	+	+	-	+	+	+	+	+	-	-	-	+	-	-
F-Pcy	*Synechococcus* lacustris Tous	-	-	-	-	-	+	+	+	+	-	+	+	+	+	+	-	-	-	-	-	-
F-Pcy	*Synechococcus* sp. 4A10	-	-	-	-	-	+	+	+	+	-	+	+	+	+	+	-	-	-	+	-	-
F-Pcy	*Synechococcus* sp. MW101C3	-	-	-	-	-	+	+	+	+	-	+	+	+	+	+	-	-	-	+	-	+
F-Pcy	*Synechococcus* sp. FACHB-909	-	-	-	-	-	+	+	+	+	-	+	+	+	+	+	-	-	-	+	-	-
F-Pcy	*Synechococcus* sp. CCAP1479/10	-	-	-	-	-	+	+	+	+	-	+	+	+	+	+	-	-	-	+	-	-
F-Pcy	*Synechococcus* sp. CCAP1479/13	-	-	-	-	-	+	+	+	+	-	+	+	+	+	+	-	-	-	+	-	-
F-Pcy	*Synechococcus* sp. CCAP1479/9	-	-	-	-	-	+	+	+	+	-	+	+	+	+	+	-	-	-	+	-	-
F-Pcy	*Synechococcus* sp. CCY0621	-	-	-	-	-	+	+	+	+	-	+	+	+	+	+	-	-	-	+	-	-
F-Pcy	*Synechococcus* sp. CCY9618	-	-	-	-	-	+	+	+	+	-	+	+	+	+	+	-	-	-	-	-	-
F-Pcy	*Vulcanococcus limneticus* LL	-	-	-	-	-	+	+	+	+	-	+	+	+	+	+	-	-	-	+	-	-
SE-Cy	*Synechococcus* elongatus FACHB-1061	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	+	-	+	-
SE-Cy	*Synechococcus* elongatus FACHB-242	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	+	-	+	-
SE-Cy	*Synechococcus* elongatus PCC 6301	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	+	-	+	-
SE-Cy	*Synechococcus* elongatus PCC 7942	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	+	-	+	-
SE-Cy	*Synechococcus* elongatus UTEX 2973	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	+	-	+	-
SE-Cy	*Synechococcus* sp. PCC 6312	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	-	+	-	-
SE-Cy	*Synechococcus* sp. PCC 7502	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	+	-	-	-
SE-Cy	*Thermosynechococcus elongatus* BP-1	+	+	+	+	+	-	-	-	-	-	+	+	+	+	+	-	-	-	-	-	-
SE-Cy	*Thermosynechococcus* sp. NK55a	+	-	+	+	+	-	-	-	-	-	+	+	+	+	+	-	-	-	-	-	-
T-Pcy	*Synechococcus* sp. JA-2-3Ba(2-13)	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	+	-	-	-
T-Pcy	*Synechococcus* sp. JA-3-3Ab	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	+	-	-	-
T-Pcy	*Synechococcus* sp. 60AY4M2	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-

Table B.10 continued from previous page

Group	Genome	*natA*	*natB*	*natC*	*natD*	*natE*	*natF*	*natG*	*natH*	*bgtA*	*bgtB*	*natI*	*natJ*	*natK*	*natL*	*natM*	*gtrA*	*gtrB*	*gtrC*	*gltS*	*dmeA*	*agcS*
T-Pcy	*Synechococcus* sp. 63AY4M1	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-
T-Pcy	*Synechococcus* sp. 63AY4M2	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-
T-Pcy	*Synechococcus* sp. 65AY640	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-
T-Pcy	*Synechococcus* sp. 65AY6A5	-	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	+	-	-	-
T-Pcy	*Synechococcus* sp. 65AY6Li	-	+	+	+	+	+	+	+	+	-	-	-	-	+	-	-	-	+	-	-	-
M-Pcy	*Cyanobium* sp. NIES-981	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	+	+	-	-	-	+
M-Pcy	*Cyanobium* sp. PCC 7001	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	+	+	-	-	-	+
M-Pcy	*Prochlorococcus marinus* NATL2A	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus marinus* MIT9312	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus marinus* MIT9515	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus marinus* NATL1A	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus* sp. MIT0801	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus* sp. MIT1306	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	+	+	-	+	-	+
M-Pcy	*Prochlorococcus* sp. P1363	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	+	+	-	-	-	+
M-Pcy	*Prochlorococcus* sp. SS52	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+
M-Pcy	*Prochlorococcus marinus* MIT9211	-	-	-	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+
M-Pcy	*Prochlorococcus marinus* MIT9303	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	+	+	-	+	-	+
M-Pcy	*Prochlorococcus marinus* MIT9313	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	+	+	-	-	-	+
M-Pcy	*Prochlorococcus marinus* CCMP1375	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+
M-Pcy	*Prochlorococcus marinus* AS9601	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus marinus* MIT9215	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus marinus* MIT9301	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus marinus* MIT9302	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus marinus* pastoris CCMP1986	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus* sp. MIT0604	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus* sp. RS50	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus* sp. RS01	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Prochlorococcus* sp. RS04	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. NIES970	+	+	+	+	+	-	-	-	-	-	+	+	+	+	+	-	-	+	-	-	-
M-Pcy	*Synechococcus* sp. NKBG15041c	+	+	+	+	+	-	-	-	-	-	+	+	+	+	+	-	-	+	-	-	-
M-Pcy	*Synechococcus* sp. PCC 7002	+	+	+	+	-	-	-	-	-	-	+	+	+	+	+	-	-	+	-	-	-
M-Pcy	*Synechococcus* sp. OG1	+	+	+	-	+	-	-	-	-	-	+	+	+	+	+	-	-	+	-	-	-

Table B.10 continued from previous page

Group	Genome	*natA*	*natB*	*natC*	*natD*	*natE*	*natF*	*natG*	*natH*	*bgtA*	*bgtB*	*natI*	*natJ*	*natK*	*natL*	*natM*	*gtrA*	*gtrB*	*gtrC*	*gltS*	*dmeA*	*agcS*
M-Pcy	*Synechococcus* sp. PCC 7003	+	+	+	+	+	-	-	-	-	-	+	+	+	+	+	-	-	+	-	-	-
M-Pcy	*Synechococcus* sp. PCC 7117	+	+	+	+	+	-	-	-	-	-	+	+	+	+	+	-	-	+	-	-	-
M-Pcy	*Synechococcus* sp. PCC 11901	+	+	+	+	+	-	-	-	-	-	+	+	+	+	+	-	-	+	-	-	-
M-Pcy	*Synechococcus* sp. NKBG042902	+	+	+	+	+	-	-	-	-	-	+	+	+	+	+	-	-	+	-	-	-
M-Pcy	*Synechococcus* sp. PCC 8807	+	+	+	+	+	-	-	-	-	-	+	+	+	+	+	-	-	+	-	-	-
M-Pcy	*Synechococcus* sp. PCC 73109	+	+	+	+	+	-	-	-	-	-	+	+	+	+	+	-	-	+	-	-	-
M-Pcy	*Synechococcus* sp. BDU 130192	+	+	+	+	+	-	-	-	-	-	+	+	+	+	+	-	-	+	-	-	-
M-Pcy	*Synechococcus* sp. CB0101	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	-	-	-	-	-	+
M-Pcy	*Synechococcus* sp. CC9311	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	+	+	-	-	-	+
M-Pcy	*Synechococcus* sp. CC9605	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+
M-Pcy	*Synechococcus* sp. CC9902	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+
M-Pcy	*Synechococcus* sp. KORDI-100	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+
M-Pcy	*Synechococcus* sp. KORDI-49	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+
M-Pcy	*Synechococcus* sp. KORDI-52	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	+	+	-	-	-	+
M-Pcy	*Synechococcus* sp. RCC307	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	+
M-Pcy	*Synechococcus* sp. BL107	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	+	-	+
M-Pcy	*Synechococcus* sp. BS55D	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+
M-Pcy	*Synechococcus* sp. BS56D	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+
M-Pcy	*Synechococcus* sp. CC9616	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+
M-Pcy	*Synechococcus* sp. GEYO	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	+	+	-	-	-	+
M-Pcy	*Synechococcus* sp. HB1133	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+
M-Pcy	*Synechococcus* sp. MITS9504	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	+	+	-	-	-	+
M-Pcy	*Synechococcus* sp. MITS9508	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	+	+	-	-	-	+
M-Pcy	*Synechococcus* sp. MITS9509	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	+	+	-	-	-	+
M-Pcy	*Synechococcus* sp. N26	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+
M-Pcy	*Synechococcus* sp. PCC 7335	-	-	-	-	-	+	+	+	+	+	-	-	-	-	+	+	+	+	-	-	+
M-Pcy	*Synechococcus* sp. PCC 7336	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	+	+	+	+	-	-
M-Pcy	*Synechococcus* sp. RS9916	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	+	+	-	-	-	+
M-Pcy	*Synechococcus* sp. RS9917	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+
M-Pcy	*Synechococcus* sp. RSCCF101	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+
M-Pcy	*Synechococcus* sp. UW105	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	+	+	-	-	-	+
M-Pcy	*Synechococcus* sp. UW140	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	+	+	-	-	-	+

Table B.10 continued from previous page

Group	**Genome**	***natA***	***natB***	***natC***	***natD***	***natE***	***natF***	***natG***	***natH***	***bgtA***	***bgtB***	***natI***	***natJ***	***natK***	***natL***	***natM***	***gtrA***	***gtrB***	***gtrC***	***gltS***	***dmeA***	***agcS***
M-Pcy	*Synechococcus* sp. UW179A	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	+	+	-	-	-	+
M-Pcy	*Synechococcus* sp. UW179B	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	+	+	-	-	-	+
M-Pcy	*Synechococcus* sp. UW69	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+
M-Pcy	*Synechococcus* sp. UW86	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+
M-Pcy	*Synechococcus* sp. WH5701	-	-	-	-	-	+	+	+	+	-	+	+	+	+	+	-	-	-	+	-	-
M-Pcy	*Synechococcus* sp. WH7805	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+
M-Pcy	*Synechococcus* sp. WH8016	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	+	+	-	-	-	+
M-Pcy	*Synechococcus* sp. XM-24	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	+	+	-	-	-	+
M-Pcy	*Synechococcus* sp. YX04-3	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	+	-	+
M-Pcy	*Synechococcus* sp. A15-127	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+
M-Pcy	*Synechococcus* sp. A15-44	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+
M-Pcy	*Synechococcus* sp. BMK-MC-1	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+
M-Pcy	*Synechococcus* sp. LTW-R	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	+	+	-	-	-	+
M-Pcy	*Synechococcus* sp. NOUM97013	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+
M-Pcy	*Synechococcus* sp. SYN20	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	+	+	-	-	-	+
M-Pcy	*Synechococcus* sp. SynAce01	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. WH7803	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+
M-Pcy	*Synechococcus* sp. WH8020	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	+	+	-	-	-	+
M-Pcy	*Synechococcus* sp. WH8101	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+
M-Pcy	*Synechococcus* sp. WH8102	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	+	-	+
M-Pcy	*Synechococcus* sp. WH8103	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+
M-Pcy	*Synechococcus* sp. WH8109	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+
LF-Cy	*Anabaena* sp. 90	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	+	+	+	-	-	-
LF-Cy	*Anabaena* sp. FACHB-1391	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	-	-	-
LF-Cy	*Anabaena* sp. WA102	+	-	+	+	+	+	+	+	+	+	-	-	-	-	-	+	+	+	-	-	-
LF-Cy	*Aphanizomenon flos-aquae* FACHB-1287	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	+	+	+	-	-	-
LF-Cy	*Aphanizomenon flos-aquae* LD13	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	+	+	+	-	-	-
LF-Cy	*Aphanizomenon flos-aquae* NIES-81	+	+	+	-	+	+	+	+	+	+	-	-	-	-	-	+	+	+	-	-	-
LF-Cy	*Aphanizomenon* sp. FACHB-1416	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	+	+	+	-	-	-
LF-Cy	*Arthrospira platensis* C1	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	+	+	-	-	-	+
LF-Cy	*Arthrospira platensis* FACHB-971	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	+	+	-	-	-	+
LF-Cy	*Cylindrospermopsis raciborskii* MVCC14	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	+	+	-	-	-

Table B.10 continued from previous page

Group	Genome	*natA*	*natB*	*natC*	*natD*	*natE*	*natF*	*natG*	*natH*	*bgtA*	*bgtB*	*natI*	*natJ*	*natK*	*natL*	*natM*	*gtrA*	*gtrB*	*gtrC*	*gltS*	*dmeA*	*agcS*
LF-Cy	*Cylindrospermopsis raciborskii* CS-505	+	+	+	-	+	+	+	+	+	-	-	-	-	-	-	+	+	+	-	-	-
LF-Cy	*Cylindrospermopsis raciborskii* ITEP-A1	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	+	+	+	-	-	-
LF-Cy	*Cylindrospermopsis raciborskii* S07	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	+	+	+	-	-	-
LF-Cy	*Dolichospermum* sp. UHCC 0352	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	+	+	-	-	-
LF-Cy	*Dolichospermum compactum* NIES-806	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	+	+	+	-	-	-
LF-Cy	*Dolichospermum* sp. FACHB-1091	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	+	+	+	-	-	-
LF-Cy	*Geminocystis herdmanii* PCC 6308	+	+	+	+	+	-	-	-	+	+	-	-	-	-	-	+	+	+	+	-	-
LF-Cy	*Geminocystis* sp. NIES-3708	+	+	+	+	+	-	-	-	+	+	-	-	-	-	-	+	+	+	+	-	-
LF-Cy	*Geminocystis* sp. NIES-3709	+	+	+	+	+	-	-	-	+	+	-	-	-	-	-	+	+	+	+	-	-
LF-Cy	*Gloeocapsa* sp. PCC 73106	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	+	+	+	-	-	+
LF-Cy	*Gloeocapsa* sp. PCC 7428	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	-	+	+
LF-Cy	*Microcystis aeruginosa* KLA2	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	-	-
LF-Cy	*Microcystis aeruginosa* NIES-2481	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	-	-	-
LF-Cy	*Microcystis aeruginosa* PCC 7005	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	-	-	-
LF-Cy	*Microcystis aeruginosa* PCC 7806SL	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	+	+	-	-
LF-Cy	*Microcystis panniformis* FACHB-1757	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	-	-	-
LF-Cy	*Microcystis* sp. 0824	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	-	-	-
LF-Cy	*Microcystis* sp. MC19	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	-	-
LF-Cy	*Nostoc paludosum* FACHB-159	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	-	-	-
LF-Cy	*Nostoc* sp. PCC 7107	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	+	+	+	-	-	-
LF-Cy	*Nostoc* sp. PCC 7120	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-
LF-Cy	*Nostoc* sp. PCC 7524	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	-	-	-
LF-Cy	*Nostoc sphaeroides* Kutzing En	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	-	-	+	-	-	-
LF-Cy	*Planktothricoides* sp. FACHB-1261	-	-	-	-	-	+	+	+	+	-	+	+	+	+	+	-	-	+	-	-	+
LF-Cy	*Planktothricoides* sp. FACHB-1370	-	-	-	-	-	+	+	+	+	-	+	+	+	+	+	-	-	+	-	-	-
LF-Cy	*Planktothricoides* sp. SR001	-	-	-	-	-	+	+	+	+	-	+	+	+	+	+	-	-	+	-	-	+
LF-Cy	*Planktothrix prolifica* NIVA-CYA 406	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	-	-	-	+	-	-
LF-Cy	*Planktothrix agardhii* NIES-204	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-
LF-Cy	*Planktothrix agardhii* NIVA-CYA 1268	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-
LF-Cy	*Planktothrix rubescens* NIVA-CYA 407	-	-	-	-	-	+	+	-	+	+	+	+	+	+	+	-	-	-	-	-	-
LF-Cy	*Planktothrix* sp. PCC 11201	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	-	-	-	+	-	-
LF-Cy	*Pleurocapsa* sp. PCC 7319	+	+	+	+	+	+	+	+	+	+	-	-	-	-	+	+	+	+	-	+	+

Table B.10 continued from previous page

Group	Genome	*natA*	*natB*	*natC*	*natD*	*natE*	*natF*	*natG*	*natH*	*bgtA*	*bgtB*	*natI*	*natJ*	*natK*	*natL*	*natM*	*gtrA*	*gtrB*	*gtrC*	*gltS*	*dmeA*	*agcS*
LF-Cy	*Pleurocapsa* sp. PCC 7327	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	+	+	+	-	-	+
LF-Cy	*Prochlorothrix hollandica* PCC 9006	-	+	-	-	-	+	+	+	+	-	-	-	-	-	-	+	+	-	-	-	+
LF-Cy	*Pseudanabaena* sp. ABRG5-3	+	+	-	+	+	+	+	+	+	-	+	+	+	+	+	-	-	+	-	-	-
LF-Cy	*Pseudanabaena* sp. FACHB-723	+	+	-	+	+	+	+	+	+	-	+	+	+	+	+	-	-	+	-	-	-
LF-Cy	*Pseudanabaena* sp. PCC 6802	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	+	-	-	-
LF-Cy	*Synechocystis* sp. FACHB-929	+	+	+	+	+	-	-	-	+	+	-	-	-	-	-	+	+	+	+	-	-
LF-Cy	*Synechocystis* sp. PCC 6714	+	+	+	+	+	-	-	-	+	+	-	-	-	-	-	+	+	+	+	-	-
LF-Cy	*Synechocystis* sp. PCC 6803	+	+	+	+	+	-	-	-	+	+	-	-	-	-	-	+	+	+	+	-	-
LF-Cy	*Synechocystis* sp. CACIAM 05	+	+	+	+	+	-	-	-	+	+	-	-	-	-	-	+	+	+	+	-	-

Table B.11: **Prevalence of organic and novel nitrogen assimilatory genes in cyanobacteria**. F-Pcy: freshwater picocyanobacteria; SE-Cy: *Synechococcus elongatus* strains; T-Pcy: thermal picocyanobacteria; M-Pcy: marine picocyanobacteria; LF-Cy: larger freshwater cyanobacteria.

Group	**Genome**	***urtA***	***urtB***	***urtC***	***urtD***	***urtE***	***ureA***	***ureB***	***ureC***	***ureD***	***ureE***	***ureF***	***ureG***	***ureJ***	***oppC***	***dppA***	***dppB***	***dppC***	***dppD***	***dppF***	***chdA***	***chiA***	***choA***	***phnD***	***amiA***
F-Pcy	*Cyanobium gracile* 6307	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	+	-	+	+	-
F-Pcy	*Cyanobium usitatum* Tous	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	-	-	+	+	-
F-Pcy	*Synechococcus* sp. 8F6	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	+	-	+	+	-
F-Pcy	*Synechococcus* sp. 6H9	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	+	-	+	+	-
F-Pcy	*Vulcanococcus limneticus* 3B3	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	-	-	+	+	-
F-Pcy	*Synechococcus* sp. BO8801	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	+	-	+	+	-
F-Pcy	*Synechococcus* sp. 1G10	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	+	-	+	+	-
F-Pcy	*Synechococcus* lacustris Tous	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	-	-	+	+	-
F-Pcy	*Synechococcus* sp. 4A10	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	+	-	+	+	-
F-Pcy	*Synechococcus* sp. MW101C3	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	-	+	+	-
F-Pcy	*Synechococcus* sp. FACHB-909	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	+	-	+	+	-
F-Pcy	*Synechococcus* sp. CCAP1479/10	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	+	-	+	+	-
F-Pcy	*Synechococcus* sp. CCAP1479/13	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	+	-	+	+	-
F-Pcy	*Synechococcus* sp. CCAP1479/9	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	+	-	+	+	-
F-Pcy	*Synechococcus* sp. CCY0621	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	+	-	+	+	-
F-Pcy	*Synechococcus* sp. CCY9618	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	+	-	+	+	-
F-Pcy	*Vulcanococcus limneticus* LL	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	-	-	+	+	-
SE-Cy	*Synechococcus* elongatus FACHB-1061	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	-	-	+	+	-	+
SE-Cy	*Synechococcus* elongatus FACHB-242	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	-	-	+	+	-	+
SE-Cy	*Synechococcus* elongatus PCC 6301	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	-	-	+	+	-	+
SE-Cy	*Synechococcus* elongatus PCC 7942	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	-	-	+	+	-	+
SE-Cy	*Synechococcus* elongatus UTEX 2973	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	-	-	+	+	-	+
SE-Cy	*Synechococcus* sp. PCC 6312	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	-	+	-	-
SE-Cy	*Synechococcus* sp. PCC 7502	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	-	+	-	-
SE-Cy	*Thermosynechococcus elongatus* BP-1	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	-	+	-	-
SE-Cy	*Thermosynechococcus* sp. NK55a	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	-	+	-	-
T-Pcy	*Synechococcus* sp. JA-2-3Ba(2-13)	+	+	+	+	-	+	+	+	+	+	+	+	+	+	-	-	-	+	-	+	-	+	-	-
T-Pcy	*Synechococcus* sp. JA-3-3Ab	+	+	+	+	-	+	+	+	+	+	+	+	+	+	-	-	-	+	-	+	-	-	-	-
T-Pcy	*Synechococcus* sp. 60AY4M2	+	+	+	+	-	-	+	-	+	-	-	-	+	+	-	-	-	+	-	+	-	-	-	-

Table B.11 continued from previous page

Group	Genome	*urtA*	*urtB*	*urtC*	*urtD*	*urtE*	*ureA*	*ureB*	*ureC*	*ureD*	*ureE*	*ureF*	*ureG*	*ureJ*	*oppC*	*dppA*	*dppB*	*dppC*	*dppD*	*dppF*	*chdA*	*chiA*	*choA*	*phnD*	*amiA*
T-Pcy	*Synechococcus* sp. 63AY4M1	+	+	+	+	-	-	+	-	+	-	-	-	+	+	-	-	-	+	-	+	-	-	-	-
T-Pcy	*Synechococcus* sp. 63AY4M2	+	+	+	+	-	-	+	-	+	-	-	-	+	+	-	-	-	+	-	+	-	-	-	-
T-Pcy	*Synechococcus* sp. 65AY640	+	+	+	+	-	-	+	-	+	-	-	-	+	+	-	-	-	+	-	+	-	-	-	-
T-Pcy	*Synechococcus* sp. 65AY6A5	+	+	+	+	-	-	+	-	+	-	-	-	+	+	-	-	-	+	-	+	-	-	-	-
T-Pcy	*Synechococcus* sp. 65AY6Li	+	+	+	+	-	+	+	+	+	+	+	+	+	+	-	-	-	+	-	+	-	-	-	-
M-Pcy	*Cyanobium* sp. NIES-981	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Cyanobium* sp. PCC 7001	-	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	+	-	+	+	-
M-Pcy	*Prochlorococcus marinus* NATL2A	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Prochlorococcus marinus* MIT9312	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	-	-	-	+	-
M-Pcy	*Prochlorococcus marinus* MIT9515	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	-	-	-	+	-
M-Pcy	*Prochlorococcus marinus* NATL1A	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Prochlorococcus* sp. MIT0801	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Prochlorococcus* sp. MIT1306	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	+	+	+	-
M-Pcy	*Prochlorococcus* sp. P1363	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	-	+	-	-
M-Pcy	*Prochlorococcus* sp. SS52	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Prochlorococcus marinus* MIT9211	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Prochlorococcus marinus* MIT9303	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	+	+	+	-
M-Pcy	*Prochlorococcus marinus* MIT9313	+	+	+	+	+	+	+	+	+	-	+	+	+	-	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Prochlorococcus marinus* CCMP1375	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Prochlorococcus marinus* AS9601	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	-	-	-	+	-
M-Pcy	*Prochlorococcus marinus* MIT9215	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	-	-	-	+	-
M-Pcy	*Prochlorococcus marinus* MIT9301	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	-	-	+	-
M-Pcy	*Prochlorococcus marinus* MIT9302	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	-	-	-	+	-
M-Pcy	*Prochlorococcus marinus* pastoris CCMP1986	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	-	-	+	-
M-Pcy	*Prochlorococcus* sp. MIT0604	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	-	-	+	-
M-Pcy	*Prochlorococcus* sp. RS50	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	-	-	-	+	-
M-Pcy	*Prochlorococcus* sp. RS01	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	-	-	-	+	-
M-Pcy	*Prochlorococcus* sp. RS04	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	-	-	-	+	-
M-Pcy	*Synechococcus* sp. NIES970	+	+	+	+	-	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Synechococcus* sp. NKBG15041c	+	+	-	+	-	+	+	+	-	+	-	+	-	+	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Synechococcus* sp. PCC 7002	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Synechococcus* sp. OG1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	+	-	-

Table B.11 continued from previous page

Group	Genome	*urtA*	*urtB*	*urtC*	*urtD*	*urtE*	*ureA*	*ureB*	*ureC*	*ureD*	*ureE*	*ureF*	*ureG*	*ureJ*	*oppC*	*dppA*	*dppB*	*dppC*	*dppD*	*dppF*	*chdA*	*chiA*	*choA*	*phnD*	*amiA*
M-Pcy	*Synechococcus* sp. PCC 7003	-	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Synechococcus* sp. PCC 7117	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	-	-	+	+	-
M-Pcy	*Synechococcus* sp. PCC 11901	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	-	+	-	-
M-Pcy	*Synechococcus* sp. NKBG042902	+	+	+	-	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Synechococcus* sp. PCC 8807	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Synechococcus* sp. PCC 73109	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Synechococcus* sp. BDU 130192	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Synechococcus* sp. CB0101	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	-	-	-	+	-
M-Pcy	*Synechococcus* sp. CC9311	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Synechococcus* sp. CC9605	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Synechococcus* sp. CC9902	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Synechococcus* sp. KORDI-100	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	+	+	+	-
M-Pcy	*Synechococcus* sp. KORDI-49	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	+	+	+	-
M-Pcy	*Synechococcus* sp. KORDI-52	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	+	+	+	-
M-Pcy	*Synechococcus* sp. RCC307	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	+	+	+	-
M-Pcy	*Synechococcus* sp. BL107	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	-	-	+	-
M-Pcy	*Synechococcus* sp. BS55D	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Synechococcus* sp. BS56D	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Synechococcus* sp. CC9616	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Synechococcus* sp. GEYO	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Synechococcus* sp. HB1133	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	-	+	-	-
M-Pcy	*Synechococcus* sp. MITS9504	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	-	+	+	+	-
M-Pcy	*Synechococcus* sp. MITS9508	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Synechococcus* sp. MITS9509	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	-	+	+	+	-
M-Pcy	*Synechococcus* sp. N26	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Synechococcus* sp. PCC 7335	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	+	+	+	+	-	-	+	+	+
M-Pcy	*Synechococcus* sp. PCC 7336	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Synechococcus* sp. RS9916	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	+	+	+	-
M-Pcy	*Synechococcus* sp. RS9917	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	+	+	+	-
M-Pcy	*Synechococcus* sp. RSCCF101	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	-	+	+	+	-
M-Pcy	*Synechococcus* sp. UW105	+	+	-	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	+	+	+	-
M-Pcy	*Synechococcus* sp. UW140	+	+	-	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	+	+	+	-

Table B.11 continued from previous page

Group	Genome	*urtA*	*urtB*	*urtC*	*urtD*	*urtE*	*ureA*	*ureB*	*ureC*	*ureD*	*ureE*	*ureF*	*ureG*	*ureJ*	*oppC*	*dppA*	*dppB*	*dppC*	*dppD*	*dppF*	*chdA*	*chiA*	*choA*	*phnD*	*amiA*
M-Pcy	*Synechococcus* sp. UW179A	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	-	+	+	+	-
M-Pcy	*Synechococcus* sp. UW179B	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	+	+	+	-
M-Pcy	*Synechococcus* sp. UW69	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	+	+	+	-
M-Pcy	*Synechococcus* sp. UW86	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Synechococcus* sp. WH5701	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Synechococcus* sp. WH7805	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	+	+	+	-
M-Pcy	*Synechococcus* sp. WH8016	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	+	+	+	-
M-Pcy	*Synechococcus* sp. XM-24	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Synechococcus* sp. YX04-3	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Synechococcus* sp. A15-127	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Synechococcus* sp. A15-44	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	+	+	+	-
M-Pcy	*Synechococcus* sp. BMK-MC-1	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	+	+	+	-
M-Pcy	*Synechococcus* sp. LTW-R	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Synechococcus* sp. NOUM97013	+	+	-	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Synechococcus* sp. SYN20	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	+	+	+	-
M-Pcy	*Synechococcus* sp. SynAce01	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	+	-	+	+	-
M-Pcy	*Synechococcus* sp. WH7803	+	-	-	-	-	-	-	-	-	-	-	-	+	-	+	+	+	+	-	-	+	+	+	-
M-Pcy	*Synechococcus* sp. WH8020	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	+	+	+	-
M-Pcy	*Synechococcus* sp. WH8101	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	+	+	+	-
M-Pcy	*Synechococcus* sp. WH8102	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	-	+	+	-
M-Pcy	*Synechococcus* sp. WH8103	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	-	-	+	-
M-Pcy	*Synechococcus* sp. WH8109	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-	-	+	+	-
LF-Cy	*Anabaena* sp. 90	-	-	-	-	-	+	+	+	+	+	+	+	-	+	+	+	+	+	+	-	-	+	-	-
LF-Cy	*Anabaena* sp. FACHB-1391	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	-	-	+	-	-
LF-Cy	*Anabaena* sp. WA102	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	-	-	+	-	-
LF-Cy	*Aphanizomenon flos-aquae* FACHB-1287	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	-	-	+	-	-
LF-Cy	*Aphanizomenon flos-aquae* LD13	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	-	-	+	-	-
LF-Cy	*Aphanizomenon flos-aquae* NIES-81	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	-	+	-	-
LF-Cy	*Aphanizomenon* sp. FACHB-1416	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	-	-	+	-	-
LF-Cy	*Arthrospira platensis* C1	+	+	-	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	-	+	-	-
LF-Cy	*Arthrospira platensis* FACHB-971	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	-	+	-	-
LF-Cy	*Cylindrospermopsis raciborskii* MVCC14	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	-	+	-	-

Table B.11 continued from previous page

Group	Genome	*urtA*	*urtB*	*urtC*	*urtD*	*urtE*	*ureA*	*ureB*	*ureC*	*ureD*	*ureE*	*ureF*	*ureG*	*ureJ*	*oppC*	*dppA*	*dppB*	*dppC*	*dppD*	*dppF*	*chdA*	*chiA*	*choA*	*phnD*	*amiA*
LF-Cy	*Cylindrospermopsis raciborskii* CS-505	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	-	+	+	-
LF-Cy	*Cylindrospermopsis raciborskii* ITEP-A1	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	-	+	+	-
LF-Cy	*Cylindrospermopsis raciborskii* S07	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	-	+	-	-
LF-Cy	*Dolichospermum* sp. UHCC 0352	-	-	-	-	-	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	-	+	-	-
LF-Cy	*Dolichospermum compactum* NIES-806	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	+	-	+	-	-
LF-Cy	*Dolichospermum* sp. FACHB-1091	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	-	+	-	-
LF-Cy	*Geminocystis herdmanii* PCC 6308	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	-	+	-	-
LF-Cy	*Geminocystis* sp. NIES-3708	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	-	+	-	-
LF-Cy	*Geminocystis* sp. NIES-3709	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	+	-	-
LF-Cy	*Gloeocapsa* sp. PCC 73106	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	+	-	-	-	+	-	-
LF-Cy	*Gloeocapsa* sp. PCC 7428	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	-	+
LF-Cy	*Microcystis aeruginosa* KLA2	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	+	-	-
LF-Cy	*Microcystis aeruginosa* NIES-2481	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	+	-	-
LF-Cy	*Microcystis aeruginosa* PCC 7005	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	+	-	-
LF-Cy	*Microcystis aeruginosa* PCC 7806SL	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	+	-	-
LF-Cy	*Microcystis panniformis* FACHB-1757	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	+	-	-
LF-Cy	*Microcystis* sp. 0824	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	+	-	-
LF-Cy	*Microcystis* sp. MC19	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	+	-	-
LF-Cy	*Nostoc paludosum* FACHB-159	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	+	-	+	+	-
LF-Cy	*Nostoc* sp. PCC 7107	-	-	-	-	-	+	+	+	+	+	+	+	-	+	+	+	+	+	-	+	-	+	-	-
LF-Cy	*Nostoc* sp. PCC 7120	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	+	-	+	+	+
LF-Cy	*Nostoc* sp. PCC 7524	-	-	-	-	-	+	+	+	+	+	+	+	-	+	+	+	+	+	-	+	-	+	+	+
LF-Cy	*Nostoc sphaeroides* Kutzing En	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	+	-	+	-	+
LF-Cy	*Planktothricoides* sp. FACHB-1261	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	-	+	-	+
LF-Cy	*Planktothricoides* sp. FACHB-1370	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	-	+	-	+
LF-Cy	*Planktothricoides* sp. SR001	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	-	-	-	+	-	+
LF-Cy	*Planktothrix prolifica* NIVA-CYA 406	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	-	-
LF-Cy	*Planktothrix agardhii* NIES-204	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	-	-
LF-Cy	*Planktothrix agardhii* NIVA-CYA 1268	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	-	-
LF-Cy	*Planktothrix rubescens* NIVA-CYA 407	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	-	-
LF-Cy	*Planktothrix* sp. PCC 11201	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	-	-
LF-Cy	*Pleurocapsa* sp. PCC 7319	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	-	-	-	+	-	-

Table B.11 continued from previous page

Group	**Genome**	***urtA***	***urtB***	***urtC***	***urtD***	***urtE***	***ureA***	***ureB***	***ureC***	***ureD***	***ureE***	***ureF***	***ureG***	***ureJ***	***oppC***	***dppA***	***dppB***	***dppC***	***dppD***	***dppF***	***chdA***	***chiA***	***choA***	***phnD***	***amiA***
LF-Cy	*Pleurocapsa* sp. PCC 7327	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	-	+	-	-
LF-Cy	*Prochlorothrix hollandica* PCC 9006	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	-	-	+	+	-
LF-Cy	*Pseudanabaena* sp. ABRG5-3	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	+	-	-
LF-Cy	*Pseudanabaena* sp. FACHB-723	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	+	-	-
LF-Cy	*Pseudanabaena* sp. PCC 6802	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	-	+	+	-
LF-Cy	*Synechocystis* sp. FACHB-929	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	+	-	-	+	+	-	-
LF-Cy	*Synechocystis* sp. PCC 6714	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	+	-	-	-	+	-	-
LF-Cy	*Synechocystis* sp. PCC 6803	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	+	-	-	+	+	-	-
LF-Cy	*Synechocystis* sp. CACIAM 05	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	+	-	-	+	+	-	-

Table B.12: **Prevalence of nitrogen metabolism genes in cyanobacteria**. F-Pcy: freshwater picocyanobacteria; SE-Cy: *Synechococcus elongatus* strains; T-Pcy: thermal picocyanobacteria; M-Pcy: marine picocyanobacteria; LF-Cy: larger freshwater cyanobacteria.

Group	Genome	***glnA***	***glnN***	***glsF***	***gltD***	***gifA***	***gifB***	***gdhA***	***narB***	***nirA***	***nirB***	***carAB***	***narM***	***cphA***	***cphB***	***iadA***	***nblA***	***speA***	***speB***	***aguA***	***aguB***	***puo***	***putA***	***arcA***	***arcC***	***aoxA***	***syNOS***
F-Pcy	*Cyanobium gracile* 6307	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	+	-
F-Pcy	*Cyanobium usitatum* Tous	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
F-Pcy	*Synechococcus* sp. 8F6	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
F-Pcy	*Synechococcus* sp. 6H9	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
F-Pcy	*Vulcanococcus limneticus* 3B3	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	+	-
F-Pcy	*Synechococcus* sp. BO8801	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	+	+	-	-	+	-
F-Pcy	*Synechococcus* sp. 1G10	+	+	+	-	-	-	-	+	+	-	+	-	+	-	-	-	+	+	-	-	-	+	-	-	+	-
F-Pcy	*Synechococcus* lacustris Tous	-	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	+	-
F-Pcy	*Synechococcus* sp. 4A10	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	+	-
F-Pcy	*Synechococcus* sp. MW101C3	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	+	+
F-Pcy	*Synechococcus* sp. FACHB-909	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	+	+	-	-	+	-
F-Pcy	*Synechococcus* sp. CCAP1479/10	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	+	-
F-Pcy	*Synechococcus* sp. CCAP1479/13	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	+	-
F-Pcy	*Synechococcus* sp. CCAP1479/9	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	+	-
F-Pcy	*Synechococcus* sp. CCY0621	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	+	-
F-Pcy	*Synechococcus* sp. CCY9618	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	+	-
F-Pcy	*Vulcanococcus limneticus* LL	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	+	-
SE-Cy	*Synechococcus* elongatus FACHB-1061	+	+	+	-	+	-	-	+	+	+	+	+	-	-	-	+	+	-	+	+	+	+	-	-	-	-
SE-Cy	*Synechococcus* elongatus FACHB-242	+	+	+	-	+	-	-	+	+	+	+	+	-	-	-	+	+	-	+	+	+	+	-	-	-	-
SE-Cy	*Synechococcus* elongatus PCC 6301	+	+	+	-	+	-	-	+	+	+	+	+	-	-	-	+	+	-	+	+	+	+	-	-	-	-
SE-Cy	*Synechococcus* elongatus PCC 7942	+	+	+	-	+	-	-	+	+	+	+	+	-	-	-	+	+	-	+	+	+	+	-	-	-	-
SE-Cy	*Synechococcus* elongatus UTEX 2973	+	+	+	-	+	-	-	+	+	+	+	+	-	-	-	+	+	-	+	+	+	+	-	-	-	-
SE-Cy	*Synechococcus* sp. PCC 6312	+	-	+	-	-	-	-	+	+	-	+	-	+	+	+	+	+	-	-	+	-	-	+	-	-	-
SE-Cy	*Synechococcus* sp. PCC 7502	+	+	+	-	-	-	-	+	+	-	+	-	+	+	+	-	+	-	-	+	-	+	+	-	-	-
SE-Cy	*Thermosynechococcus elongatus* BP-1	+	-	+	-	+	-	-	+	+	-	+	+	+	+	+	-	+	-	-	+	-	-	+	-	-	-
SE-Cy	*Thermosynechococcus* sp. NK55a	+	-	+	-	+	-	-	+	+	-	+	-	+	+	+	-	+	-	-	+	-	-	+	-	-	-
T-Pcy	*Synechococcus* sp. JA-2-3Ba(2-13)	+	-	+	-	+	-	-	+	+	-	+	-	+	+	+	-	+	+	-	-	-	-	+	-	-	-
T-Pcy	*Synechococcus* sp. JA-3-3Ab	+	-	+	-	+	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-
T-Pcy	*Synechococcus* sp. 60AY4M2	+	-	+	-	+	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-

Table B.12 continued from previous page

Group	Genome	*glnA*	*glnN*	*glsF*	*gltD*	*gifA*	*gifB*	*gdhA*	*narB*	*nirA*	*nirB*	*carAB*	*narM*	*cphA*	*cphB*	*iadA*	*nblA*	*speA*	*speB*	*aguA*	*aguB*	*puo*	*putA*	*arcA*	*arcC*	*aoxA*	*syNOS*
T-Pcy	*Synechococcus* sp. 63AY4M1	+	-	+	-	+	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-
T-Pcy	*Synechococcus* sp. 63AY4M2	+	-	+	-	+	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-
T-Pcy	*Synechococcus* sp. 65AY640	+	-	+	-	+	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-
T-Pcy	*Synechococcus* sp. 65AY6A5	+	-	+	-	+	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-
T-Pcy	*Synechococcus* sp. 65AY6Li	+	-	+	-	+	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-
M-Pcy	*Cyanobium* sp. NIES-981	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	+	-
M-Pcy	*Cyanobium* sp. PCC 7001	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Prochlorococcus marinus* NATL2A	+	-	+	-	-	-	-	-	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Prochlorococcus marinus* MIT9312	+	-	+	-	-	-	-	-	-	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Prochlorococcus marinus* MIT9515	+	-	+	-	-	-	-	-	-	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Prochlorococcus marinus* NATL1A	+	-	+	-	-	-	-	-	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Prochlorococcus* sp. MIT0801	+	-	+	-	-	-	-	-	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Prochlorococcus* sp. MIT1306	+	-	+	-	-	-	+	-	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Prochlorococcus* sp. P1363	+	-	+	-	-	-	+	-	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Prochlorococcus* sp. SS52	+	-	+	-	-	-	-	-	-	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Prochlorococcus marinus* MIT9211	+	-	+	-	-	-	-	-	-	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Prochlorococcus marinus* MIT9303	+	-	+	-	-	-	+	-	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Prochlorococcus marinus* MIT9313	+	-	+	-	-	-	+	-	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Prochlorococcus marinus* CCMP1375	+	-	+	-	-	-	-	-	-	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Prochlorococcus marinus* AS9601	+	-	+	-	-	-	-	-	-	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Prochlorococcus marinus* MIT9215	+	-	+	-	-	-	-	-	-	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Prochlorococcus marinus* MIT9301	+	-	+	-	-	-	-	-	-	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Prochlorococcus marinus* MIT9302	+	-	+	-	-	-	-	-	-	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Prochlorococcus marinus* pastoris CCMP1986	+	-	+	-	-	-	-	-	-	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Prochlorococcus* sp. MIT0604	+	-	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Prochlorococcus* sp. RS50	+	-	+	-	-	-	-	-	-	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Prochlorococcus* sp. RS01	+	-	+	-	-	-	-	-	-	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Prochlorococcus* sp. RS04	+	-	+	-	-	-	-	-	-	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. NIES970	+	+	+	-	+	-	-	+	+	-	+	+	+	+	+	+	+	+	-	-	-	-	+	-	-	-
M-Pcy	*Synechococcus* sp. NKBG15041c	+	+	+	-	+	-	-	+	+	-	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-
M-Pcy	*Synechococcus* sp. PCC 7002	+	+	+	-	+	-	-	+	+	-	+	+	+	+	+	+	+	+	-	-	-	-	+	-	-	-

Table B.12 continued from previous page

Group	Genome	*glnA*	*glnN*	*glsF*	*gltD*	*gifA*	*gifB*	*gdhA*	*narB*	*nirA*	*nirB*	*carAB*	*narM*	*cphA*	*cphB*	*iadA*	*nblA*	*speA*	*speB*	*aguA*	*aguB*	*puo*	*putA*	*arcA*	*arcC*	*aoxA*	*syNOS*
M-Pcy	*Synechococcus* sp. OG1	+	+	+	-	+	-	-	+	+	-	+	+	+	+	+	-	+	+	-	-	-	-	+	-	-	-
M-Pcy	*Synechococcus* sp. PCC 7003	+	+	+	-	+	-	-	+	+	-	+	+	+	+	+	+	+	+	-	-	-	-	+	-	-	-
M-Pcy	*Synechococcus* sp. PCC 7117	+	+	+	-	+	-	-	+	+	-	+	+	+	+	+	+	+	+	-	-	-	-	+	-	-	-
M-Pcy	*Synechococcus* sp. PCC 11901	+	+	+	-	+	-	-	+	+	-	+	+	+	+	+	-	+	+	-	-	-	-	+	-	-	-
M-Pcy	*Synechococcus* sp. NKBG042902	+	+	+	-	+	-	-	+	+	-	+	+	+	+	+	+	+	+	-	-	-	-	+	-	-	-
M-Pcy	*Synechococcus* sp. PCC 8807	+	+	+	-	+	-	-	+	+	-	+	+	+	+	+	+	+	+	-	-	-	-	+	-	-	-
M-Pcy	*Synechococcus* sp. PCC 73109	+	+	+	-	+	-	-	+	+	-	+	+	+	+	+	+	+	+	-	-	-	-	+	-	-	-
M-Pcy	*Synechococcus* sp. BDU 130192	+	+	+	-	+	-	-	+	+	-	+	+	+	+	+	+	+	+	-	-	-	-	+	-	-	-
M-Pcy	*Synechococcus* sp. CB0101	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	+	-
M-Pcy	*Synechococcus* sp. CC9311	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. CC9605	+	-	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	+	-
M-Pcy	*Synechococcus* sp. CC9902	+	-	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. KORDI-100	+	-	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	+	-
M-Pcy	*Synechococcus* sp. KORDI-49	+	-	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. KORDI-52	+	-	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. RCC307	+	-	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. BL107	+	-	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. BS55D	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. BS56D	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. CC9616	+	-	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	+	-
M-Pcy	*Synechococcus* sp. GEYO	+	-	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. HB1133	+	-	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. MITS9504	+	-	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. MITS9508	+	-	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. MITS9509	+	-	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. N26	+	-	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. PCC 7335	-	+	+	+	-	-	+	+	+	-	+	+	+	+	-	+	+	-	-	-	-	+	+	-	-	+
M-Pcy	*Synechococcus* sp. PCC 7336	+	-	+	+	+	-	-	+	+	-	+	-	-	-	-	+	+	+	-	-	-	+	+	-	-	-
M-Pcy	*Synechococcus* sp. RS9916	+	-	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. RS9917	+	+	+	-	-	-	-	-	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. RSCCF101	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-

Table B.12 continued from previous page

Group	Genome	*glnA*	*glnN*	*glsF*	*gltD*	*gifA*	*gifB*	*gdhA*	*narB*	*nirA*	*nirB*	*carAB*	*narM*	*cphA*	*cphB*	*iadA*	*nblA*	*speA*	*speB*	*aguA*	*aguB*	*puo*	*putA*	*arcA*	*arcC*	*aoxA*	*syNOS*
M-Pcy	*Synechococcus* sp. UW105	+	-	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. UW140	+	-	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. UW179A	+	-	+	-	-	-	+	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. UW179B	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. UW69	+	-	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. UW86	+	-	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. WH5701	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	+	-
M-Pcy	*Synechococcus* sp. WH7805	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. WH8016	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. XM-24	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. YX04-3	+	-	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. A15-127	+	-	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. A15-44	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. BMK-MC-1	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. LTW-R	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. NOUM97013	+	-	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. SYN20	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. SynAce01	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	+	-
M-Pcy	*Synechococcus* sp. WH7803	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. WH8020	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. WH8101	+	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	+	-
M-Pcy	*Synechococcus* sp. WH8102	+	-	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. WH8103	+	-	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
M-Pcy	*Synechococcus* sp. WH8109	+	-	+	-	-	-	-	+	+	-	+	-	-	-	-	-	+	+	-	-	-	+	-	-	-	-
LF-Cy	*Anabaena* sp. 90	+	-	+	-	+	-	+	+	+	-	+	-	+	+	+	+	+	-	-	-	-	+	+	-	-	+
LF-Cy	*Anabaena* sp. FACHB-1391	+	-	+	-	+	-	+	+	+	-	+	-	+	+	+	+	+	-	-	-	-	+	+	-	-	-
LF-Cy	*Anabaena* sp. WA102	+	-	+	-	+	-	+	+	+	-	+	-	+	+	+	-	+	-	-	-	-	+	+	-	-	-
LF-Cy	*Aphanizomenon flos-aquae* FACHB-1287	+	-	+	-	+	-	+	+	+	-	+	-	+	+	+	-	+	-	-	-	-	+	+	-	-	-
LF-Cy	*Aphanizomenon flos-aquae* LD13	+	-	+	-	+	-	+	+	+	-	+	-	+	+	+	-	+	-	-	-	-	+	+	-	-	-
LF-Cy	*Aphanizomenon flos-aquae* NIES-81	+	-	+	-	+	-	+	+	+	-	+	-	+	+	+	-	+	-	-	-	-	+	+	-	-	-
LF-Cy	*Aphanizomenon* sp. FACHB-1416	+	-	+	-	+	-	+	+	+	-	+	-	+	+	+	-	+	-	-	-	-	+	+	-	-	-

Table B.12 continued from previous page

Group	Genome	*glnA*	*glnN*	*glsF*	*gltD*	*gifA*	*gifB*	*gdhA*	*narB*	*nirA*	*nirB*	*carAB*	*narM*	*cphA*	*cphB*	*iadA*	*nblA*	*speA*	*speB*	*aguA*	*aguB*	*puo*	*putA*	*arcA*	*arcC*	*aoxA*	*syNOS*
LF-Cy	*Arthrospira platensis* C1	+	-	+	-	+	-	+	+	+	-	+	-	+	+	+	+	+	+	-	-	-	+	+	-	-	-
LF-Cy	*Arthrospira platensis* FACHB-971	+	-	+	-	+	-	+	+	+	-	+	-	+	+	+	+	+	+	-	-	-	+	+	-	-	-
LF-Cy	*Cylindrospermopsis raciborskii* MVCC14	+	-	+	-	+	-	-	+	+	-	+	-	+	+	+	-	-	-	-	-	-	+	-	-	-	-
LF-Cy	*Cylindrospermopsis raciborskii* CS-505	+	-	+	-	+	-	-	+	+	-	+	-	+	+	+	-	+	-	-	-	-	+	+	-	-	-
LF-Cy	*Cylindrospermopsis raciborskii* ITEP-A1	+	-	+	-	+	-	-	+	+	-	+	-	+	+	+	-	+	-	-	-	-	+	+	-	-	-
LF-Cy	*Cylindrospermopsis raciborskii* S07	+	-	+	-	+	-	-	+	+	-	+	-	+	+	+	-	+	-	-	-	-	+	+	-	-	-
LF-Cy	*Dolichospermum* sp. UHCC 0352	+	-	+	-	+	-	+	+	+	-	+	-	+	+	+	-	+	-	-	-	-	+	+	-	-	+
LF-Cy	*Dolichospermum compactum* NIES-806	+	-	+	-	+	-	+	+	+	-	+	-	+	+	+	-	+	-	-	-	-	+	+	-	-	-
LF-Cy	*Dolichospermum* sp. FACHB-1091	+	-	+	-	+	-	+	+	+	-	+	-	+	+	+	+	+	-	-	-	-	+	+	-	-	-
LF-Cy	*Geminocystis herdmanii* PCC 6308	+	+	+	+	+	-	-	+	+	-	+	-	+	+	+	+	+	+	-	-	-	+	+	-	-	-
LF-Cy	*Geminocystis* sp. NIES-3708	+	+	+	+	+	-	-	+	+	-	+	-	+	+	+	+	+	-	-	-	-	+	+	-	+	-
LF-Cy	*Geminocystis* sp. NIES-3709	+	+	+	+	+	-	-	+	+	-	+	-	+	+	+	+	+	+	-	-	-	-	+	-	+	-
LF-Cy	*Gloeocapsa* sp. PCC 73106	+	+	+	+	+	-	-	+	+	-	+	+	+	+	+	+	+	+	-	-	-	+	+	-	+	-
LF-Cy	*Gloeocapsa* sp. PCC 7428	+	-	+	-	+	-	+	+	+	-	+	-	+	+	+	-	+	+	-	-	-	+	+	-	-	-
LF-Cy	*Microcystis aeruginosa* KLA2	+	+	+	+	+	-	-	+	+	-	+	-	+	+	+	-	+	+	-	-	-	+	+	-	+	-
LF-Cy	*Microcystis aeruginosa* NIES-2481	+	+	+	+	+	-	-	+	+	-	+	-	+	+	+	+	+	+	-	-	-	+	+	-	+	-
LF-Cy	*Microcystis aeruginosa* PCC 7005	+	+	+	+	-	-	-	+	+	-	+	-	+	+	+	+	+	+	-	-	-	+	+	-	+	-
LF-Cy	*Microcystis aeruginosa* PCC 7806SL	+	+	+	+	-	-	-	+	+	-	+	-	+	+	+	+	+	+	-	-	-	+	+	-	+	-
LF-Cy	*Microcystis panniformis* FACHB-1757	+	+	+	-	+	-	-	+	+	-	+	-	-	+	+	+	+	+	-	-	-	+	+	-	+	-
LF-Cy	*Microcystis* sp. 0824	+	+	+	+	+	-	-	+	+	-	+	-	+	+	+	+	+	+	-	-	-	+	+	-	+	-
LF-Cy	*Microcystis* sp. MC19	+	+	+	+	+	-	-	+	+	-	+	-	+	+	+	+	+	+	-	-	-	+	+	-	+	-
LF-Cy	*Nostoc paludosum* FACHB-159	+	-	+	-	+	-	+	+	+	-	+	-	+	+	+	+	+	-	-	-	-	+	+	-	-	-
LF-Cy	*Nostoc* sp. PCC 7107	+	-	+	-	+	-	+	+	+	-	+	-	+	+	+	-	+	-	-	-	-	+	+	-	-	+
LF-Cy	*Nostoc* sp. PCC 7120	+	-	+	-	+	-	+	+	+	-	+	-	+	+	+	+	+	-	-	-	-	+	+	-	-	-
LF-Cy	*Nostoc* sp. PCC 7524	+	-	+	-	+	-	+	+	+	-	+	-	+	+	+	-	+	-	-	-	-	+	+	-	-	-
LF-Cy	*Nostoc sphaeroides* Kutzing En	+	-	+	-	+	-	+	+	+	-	+	-	+	+	+	-	+	-	-	-	-	+	+	-	-	-
LF-Cy	*Planktothricoides* sp. FACHB-1261	+	-	+	+	+	-	+	+	+	-	+	-	+	+	+	+	+	-	-	+	-	+	+	-	-	-
LF-Cy	*Planktothricoides* sp. FACHB-1370	+	-	+	+	+	-	+	+	+	-	+	-	+	+	+	+	+	-	-	+	-	+	+	-	-	-
LF-Cy	*Planktothricoides* sp. SR001	+	-	+	+	+	-	+	+	+	-	+	-	+	+	+	+	+	-	-	+	-	+	+	-	-	-
LF-Cy	*Planktothrix prolifica* NIVA-CYA 406	+	-	+	+	+	-	+	+	+	-	+	-	+	+	+	-	+	+	-	-	-	+	+	-	+	-
LF-Cy	*Planktothrix agardhii* NIES-204	+	-	+	+	+	-	+	+	+	-	+	-	+	+	+	+	+	+	-	-	-	+	+	-	+	-

Table B.12 continued from previous page

Group	Genome	*glnA*	*glnN*	*glsF*	*gltD*	*gifA*	*gifB*	*gdhA*	*narB*	*nirA*	*nirB*	*carAB*	*narM*	*cphA*	*cphB*	*iadA*	*nblA*	*speA*	*speB*	*aguA*	*aguB*	*puo*	*putA*	*arcA*	*arcC*	*aoxA*	*syNOS*
LF-Cy	*Planktothrix agardhii* NIVA-CYA 1268	+	-	+	+	+	-	+	+	+	-	+	-	+	+	+	+	+	+	-	-	-	+	+	-	+	-
LF-Cy	*Planktothrix rubescens* NIVA-CYA 407	+	-	+	+	+	-	+	+	+	-	+	-	+	+	+	+	+	+	-	-	-	+	+	-	+	-
LF-Cy	*Planktothrix* sp. PCC 11201	+	-	+	+	+	-	+	+	+	-	+	-	+	+	+	-	+	+	-	-	-	+	+	-	-	+
LF-Cy	*Pleurocapsa* sp. PCC 7319	+	+	+	+	+	-	-	+	+	-	+	+	+	+	+	-	+	-	-	-	-	-	+	-	-	-
LF-Cy	*Pleurocapsa* sp. PCC 7327	+	-	+	+	+	-	-	+	+	-	+	+	+	+	+	+	+	+	-	-	+	-	+	-	-	-
LF-Cy	*Prochlorothrix hollandica* PCC 9006	-	+	+	+	-	-	-	+	+	-	+	-	+	+	+	-	+	-	-	-	-	+	+	-	-	-
LF-Cy	*Pseudanabaena* sp. ABRG5-3	-	+	+	+	-	-	-	+	+	-	+	-	+	+	+	-	+	-	-	+	-	+	+	-	-	-
LF-Cy	*Pseudanabaena* sp. FACHB-723	-	+	+	+	-	-	-	+	+	-	+	+	+	+	+	-	+	-	-	+	-	+	+	-	-	-
LF-Cy	*Pseudanabaena* sp. PCC 6802	+	-	+	+	-	-	-	+	+	-	+	-	+	+	+	-	+	-	-	+	-	+	+	-	-	-
LF-Cy	*Synechocystis* sp. FACHB-929	+	+	+	+	+	+	-	+	+	-	+	-	+	+	+	+	+	+	-	-	-	+	+	+	+	-
LF-Cy	*Synechocystis* sp. PCC 6714	+	+	+	+	+	+	-	+	+	-	+	-	+	+	+	+	+	+	-	-	-	+	+	+	+	-
LF-Cy	*Synechocystis* sp. PCC 6803	+	+	+	+	+	+	-	+	+	-	+	-	+	+	+	+	+	+	-	-	-	+	+	+	+	-
LF-Cy	*Synechocystis* sp. CACIAM 05	+	+	+	+	+	+	-	+	+	-	+	-	+	+	+	+	+	+	-	-	-	+	+	+	+	-

Table B.13: **Prevalence of nitrogen assimilatory regulation genes in cyanobacteria**. F-Pcy: freshwater picocyanobacteria; SE-Cy: *Synechococcus elongatus* strains; T-Pcy: thermal picocyanobacteria; M-Pcy: marine picocyanobacteria; LF-Cy: larger freshwater cyanobacteria.

Group	Genome	*ntcA*	*ntcB*	*glnB*	*icd*	*pipX*	*spkE*
F-Pcy	*Cyanobium gracile* 6307	+	-	+	+	+	-
F-Pcy	*Cyanobium usitatum* Tous	+	-	+	+	+	-
F-Pcy	*Synechococcus* sp. 8F6	+	-	+	+	+	-
F-Pcy	*Synechococcus* sp. 6H9	+	-	+	+	+	-
F-Pcy	*Vulcanococcus limneticus* 3B3	+	-	+	+	+	-
F-Pcy	*Synechococcus* sp. BO8801	+	-	+	+	+	-
F-Pcy	*Synechococcus* sp. 1G10	+	-	+	+	+	-
F-Pcy	*Synechococcus* lacustris Tous	+	-	+	+	+	-
F-Pcy	*Synechococcus* sp. 4A10	+	-	+	+	+	-
F-Pcy	*Synechococcus* sp. MW101C3	+	-	+	+	+	-
F-Pcy	*Synechococcus* sp. FACHB-909	+	-	+	+	+	-
F-Pcy	*Synechococcus* sp. CCAP1479/10	+	-	+	+	+	-
F-Pcy	*Synechococcus* sp. CCAP1479/13	+	-	+	+	+	-
F-Pcy	*Synechococcus* sp. CCAP1479/9	+	-	+	+	+	-
F-Pcy	*Synechococcus* sp. CCY0621	+	-	+	+	+	-
F-Pcy	*Synechococcus* sp. CCY9618	+	-	+	+	+	-
F-Pcy	*Vulcanococcus limneticus* LL	+	-	+	+	+	-
SE-Cy	*Synechococcus* elongatus FACHB-1061	+	+	+	+	+	-
SE-Cy	*Synechococcus* elongatus FACHB-242	+	+	+	+	+	-
SE-Cy	*Synechococcus* elongatus PCC 6301	+	+	+	+	+	-
SE-Cy	*Synechococcus* elongatus PCC 7942	+	+	+	+	+	-
SE-Cy	*Synechococcus* elongatus UTEX 2973	+	+	+	+	+	-
SE-Cy	*Synechococcus* sp. PCC 6312	+	+	+	+	+	-
SE-Cy	*Synechococcus* sp. PCC 7502	+	+	+	+	+	-
SE-Cy	*Thermosynechococcus elongatus* BP-1	+	+	+	+	+	-
SE-Cy	*Thermosynechococcus* sp. NK55a	+	+	+	+	+	-
T-Pcy	*Synechococcus* sp. JA-2-3Ba(2-13)	+	+	+	+	+	-
T-Pcy	*Synechococcus* sp. JA-3-3Ab	+	+	+	+	+	-
T-Pcy	*Synechococcus* sp. 60AY4M2	+	+	+	+	+	-
T-Pcy	*Synechococcus* sp. 63AY4M1	+	+	+	+	+	-
T-Pcy	*Synechococcus* sp. 63AY4M2	+	+	+	+	+	-
T-Pcy	*Synechococcus* sp. 65AY640	+	+	+	+	+	-
T-Pcy	*Synechococcus* sp. 65AY6A5	+	+	+	+	+	-
T-Pcy	*Synechococcus* sp. 65AY6Li	+	+	+	+	+	-
M-Pcy	*Cyanobium* sp. NIES-981	+	-	+	+	+	-
M-Pcy	*Cyanobium* sp. PCC 7001	+	-	+	+	+	-
M-Pcy	*Prochlorococcus marinus* NATL2A	+	-	+	+	+	-
M-Pcy	*Prochlorococcus marinus* MIT9312	+	-	+	+	+	-
M-Pcy	*Prochlorococcus marinus* MIT9515	+	-	+	+	+	-
M-Pcy	*Prochlorococcus marinus* NATL1A	+	-	+	+	+	-
M-Pcy	*Prochlorococcus* sp. MIT0801	+	-	+	+	+	-
M-Pcy	*Prochlorococcus* sp. MIT1306	+	-	+	+	+	-

Table B.13 continued from previous page

Group	**Genome**	***ntcA***	***ntcB***	***glnB***	***icd***	***pipX***	***spkE***
M-Pcy	*Prochlorococcus* sp. P1363	+	-	+	+	+	-
M-Pcy	*Prochlorococcus* sp. SS52	+	-	+	+	+	-
M-Pcy	*Prochlorococcus marinus* MIT9211	+	-	+	+	+	-
M-Pcy	*Prochlorococcus marinus* MIT9303	+	-	+	+	+	-
M-Pcy	*Prochlorococcus marinus* MIT9313	+	-	+	+	+	-
M-Pcy	*Prochlorococcus marinus* CCMP1375	+	-	+	+	+	-
M-Pcy	*Prochlorococcus marinus* AS9601	+	-	+	+	-	-
M-Pcy	*Prochlorococcus marinus* MIT9215	+	-	+	+	+	-
M-Pcy	*Prochlorococcus marinus* MIT9301	+	-	+	+	+	-
M-Pcy	*Prochlorococcus marinus* MIT9302	+	-	+	+	+	-
M-Pcy	*Prochlorococcus marinus* pastoris CCMP1986	+	-	+	+	+	-
M-Pcy	*Prochlorococcus* sp. MIT0604	+	-	+	+	+	-
M-Pcy	*Prochlorococcus* sp. RS50	+	-	+	+	+	-
M-Pcy	*Prochlorococcus* sp. RS01	+	-	+	+	+	-
M-Pcy	*Prochlorococcus* sp. RS04	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. NIES970	+	+	+	-	+	-
M-Pcy	*Synechococcus* sp. NKBG15041c	+	+	+	-	+	-
M-Pcy	*Synechococcus* sp. PCC 7002	+	+	+	-	+	-
M-Pcy	*Synechococcus* sp. OG1	+	+	+	-	+	-
M-Pcy	*Synechococcus* sp. PCC 7003	+	+	+	-	+	-
M-Pcy	*Synechococcus* sp. PCC 7117	+	+	+	-	+	-
M-Pcy	*Synechococcus* sp. PCC 11901	+	+	+	-	+	-
M-Pcy	*Synechococcus* sp. NKBG042902	-	+	+	-	+	-
M-Pcy	*Synechococcus* sp. PCC 8807	+	+	+	-	+	-
M-Pcy	*Synechococcus* sp. PCC 73109	+	+	+	-	+	-
M-Pcy	*Synechococcus* sp. BDU 130192	+	+	+	-	+	-
M-Pcy	*Synechococcus* sp. CB0101	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. CC9311	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. CC9605	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. CC9902	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. KORDI-100	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. KORDI-49	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. KORDI-52	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. RCC307	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. BL107	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. BS55D	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. BS56D	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. CC9616	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. GEYO	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. HB1133	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. MITS9504	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. MITS9508	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. MITS9509	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. N26	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. PCC 7335	+	+	+	-	+	-
M-Pcy	*Synechococcus* sp. PCC 7336	+	+	+	+	+	-
M-Pcy	*Synechococcus* sp. RS9916	+	-	+	+	+	-

Table B.13 continued from previous page

Group	Genome	*ntcA*	*ntcB*	*glnB*	*icd*	*pipX*	*spkE*
M-Pcy	*Synechococcus* sp. RS9917	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. RSCCF101	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. UW105	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. UW140	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. UW179A	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. UW179B	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. UW69	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. UW86	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. WH5701	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. WH7805	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. WH8016	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. XM-24	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. YX04-3	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. A15-127	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. A15-44	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. BMK-MC-1	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. LTW-R	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. NOUM97013	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. SYN20	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. SynAce01	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. WH7803	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. WH8020	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. WH8101	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. WH8102	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. WH8103	+	-	+	+	+	-
M-Pcy	*Synechococcus* sp. WH8109	+	-	+	+	+	-
LF-Cy	*Anabaena* sp. 90	+	+	+	+	+	-
LF-Cy	*Anabaena* sp. FACHB-1391	+	+	+	+	+	-
LF-Cy	*Anabaena* sp. WA102	+	-	+	+	+	+
LF-Cy	*Aphanizomenon flos-aquae* FACHB-1287	+	+	+	+	+	+
LF-Cy	*Aphanizomenon flos-aquae* LD13	+	+	+	+	+	+
LF-Cy	*Aphanizomenon flos-aquae* NIES-81	+	+	+	+	+	-
LF-Cy	*Aphanizomenon* sp. FACHB-1416	+	+	+	+	+	+
LF-Cy	*Arthrospira platensis* C1	+	+	+	+	+	-
LF-Cy	*Arthrospira platensis* FACHB-971	+	+	+	+	+	-
LF-Cy	*Cylindrospermopsis raciborskii* MVCC14	+	+	+	+	+	+
LF-Cy	*Cylindrospermopsis raciborskii* CS-505	+	+	+	+	+	+
LF-Cy	*Cylindrospermopsis raciborskii* ITEP-A1	+	+	+	+	+	+
LF-Cy	*Cylindrospermopsis raciborskii* S07	+	+	+	+	+	+
LF-Cy	*Dolichospermum* sp. UHCC 0352	+	+	+	+	+	-
LF-Cy	*Dolichospermum compactum* NIES-806	+	+	+	+	+	-
LF-Cy	*Dolichospermum* sp. FACHB-1091	+	+	+	+	+	-
LF-Cy	*Geminocystis herdmanii* PCC 6308	+	+	+	+	+	-
LF-Cy	*Geminocystis* sp. NIES-3708	+	+	+	+	+	-
LF-Cy	*Geminocystis* sp. NIES-3709	+	+	+	+	+	-
LF-Cy	*Gloeocapsa* sp. PCC 73106	+	+	+	+	+	-
LF-Cy	*Gloeocapsa* sp. PCC 7428	+	+	+	+	+	+

Table B.13 continued from previous page

Group	Genome	*ntcA*	*ntcB*	*glnB*	*icd*	*pipX*	*spkE*
LF-Cy	*Microcystis aeruginosa* KLA2	+	+	+	+	+	-
LF-Cy	*Microcystis aeruginosa* NIES-2481	+	+	+	+	+	-
LF-Cy	*Microcystis aeruginosa* PCC 7005	+	+	+	+	+	-
LF-Cy	*Microcystis aeruginosa* PCC 7806SL	+	+	+	+	+	-
LF-Cy	*Microcystis panniformis* FACHB-1757	+	+	+	+	+	-
LF-Cy	*Microcystis* sp. 0824	+	+	+	+	+	-
LF-Cy	*Microcystis* sp. MC19	+	+	+	+	+	-
LF-Cy	*Nostoc paludosum* FACHB-159	+	+	+	+	+	-
LF-Cy	*Nostoc* sp. PCC 7107	+	+	+	+	+	+
LF-Cy	*Nostoc* sp. PCC 7120	+	+	+	+	+	+
LF-Cy	*Nostoc* sp. PCC 7524	+	+	+	+	+	-
LF-Cy	*Nostoc sphaeroides* Kutzing En	+	+	+	+	+	-
LF-Cy	*Planktothricoides* sp. FACHB-1261	+	+	+	+	+	-
LF-Cy	*Planktothricoides* sp. FACHB-1370	+	+	+	+	+	-
LF-Cy	*Planktothricoides* sp. SR001	+	+	+	+	+	-
LF-Cy	*Planktothrix prolifica* NIVA-CYA 406	+	+	+	+	+	-
LF-Cy	*Planktothrix agardhii* NIES-204	+	+	+	+	+	+
LF-Cy	*Planktothrix agardhii* NIVA-CYA 1268	+	+	+	+	+	+
LF-Cy	*Planktothrix rubescens* NIVA-CYA 407	+	+	+	+	+	-
LF-Cy	*Planktothrix* sp. PCC 11201	+	+	+	+	+	+
LF-Cy	*Pleurocapsa* sp. PCC 7319	+	+	+	+	+	-
LF-Cy	*Pleurocapsa* sp. PCC 7327	+	+	+	+	+	-
LF-Cy	*Prochlorothrix hollandica* PCC 9006	+	+	+	+	+	-
LF-Cy	*Pseudanabaena* sp. ABRG5-3	+	+	+	-	+	-
LF-Cy	*Pseudanabaena* sp. FACHB-723	+	+	+	-	+	-
LF-Cy	*Pseudanabaena* sp. PCC 6802	+	+	+	-	+	-
LF-Cy	*Synechocystis* sp. FACHB-929	+	+	+	+	+	+
LF-Cy	*Synechocystis* sp. PCC 6714	+	+	+	+	+	+
LF-Cy	*Synechocystis* sp. PCC 6803	+	+	+	+	+	+
LF-Cy	*Synechocystis* sp. CACIAM 05	+	+	+	+	+	+

Table B.14: **Prevalence of amino acid biosynthesis genes in cyanobacteria**. F-Pcy: freshwater picocyanobacteria; SE-Cy: *Synechococcus elongatus* strains; T-Pcy: thermal picocyanobacteria; M-Pcy: marine picocyanobacteria; LF-Cy: larger freshwater cyanobacteria.

Group	**Genome**	**Ala**	**Arg**	**Asn-*asnB***	**Asn-*aspS*/*gatABC***	**Asp**	**Cys**	**Gln**	**Glu**	**Gly**	**His**	**Ile**	**Leu**	**Lys**	**Met**	**Phe**	**Pro**	**Ser**	**Thr**	**Trp**	**Tyr**	**Val**
F-Pcy	*Cyanobium gracile* 6307	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
F-Pcy	*Cyanobium usitatum* Tous	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
F-Pcy	*Synechococcus* sp. 8F6	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
F-Pcy	*Synechococcus* sp. 6H9	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
F-Pcy	*Vulcanococcus limneticus* 3B3	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
F-Pcy	*Synechococcus* sp. BO8801	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
F-Pcy	*Synechococcus* sp. 1G10	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
F-Pcy	*Synechococcus* lacustris Tous	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
F-Pcy	*Synechococcus* sp. 4A10	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
F-Pcy	*Synechococcus* sp. MW101C3	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
F-Pcy	*Synechococcus* sp. FACHB-909	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
F-Pcy	*Synechococcus* sp. CCAP1479/10	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
F-Pcy	*Synechococcus* sp. CCAP1479/13	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
F-Pcy	*Synechococcus* sp. CCAP1479/9	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
F-Pcy	*Synechococcus* sp. CCY0621	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
F-Pcy	*Synechococcus* sp. CCY9618	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
F-Pcy	*Vulcanococcus limneticus* LL	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
SE-Cy	*Synechococcus* elongatus FACHB-1061	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
SE-Cy	*Synechococcus* elongatus FACHB-242	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
SE-Cy	*Synechococcus* elongatus PCC 6301	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
SE-Cy	*Synechococcus* elongatus PCC 7942	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
SE-Cy	*Synechococcus* elongatus UTEX 2973	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
SE-Cy	*Synechococcus* sp. PCC 6312	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
SE-Cy	*Synechococcus* sp. PCC 7502	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
SE-Cy	*Thermosynechococcus elongatus* BP-1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
SE-Cy	*Thermosynechococcus* sp. NK55a	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

Table B.14 continued from previous page

Group	Genome	Ala	Arg	Asn-*asnB*	Asn-*aspS*/*gatABC*	Asp	Cys	Gln	Glu	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val
T-Pcy	*Synechococcus* sp. JA-2-3Ba(2-13)	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
T-Pcy	*Synechococcus* sp. JA-3-3Ab	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
T-Pcy	*Synechococcus* sp. 60AY4M2	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
T-Pcy	*Synechococcus* sp. 63AY4M1	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
T-Pcy	*Synechococcus* sp. 63AY4M2	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
T-Pcy	*Synechococcus* sp. 65AY640	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
T-Pcy	*Synechococcus* sp. 65AY6A5	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
T-Pcy	*Synechococcus* sp. 65AY6Li	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Cyanobium* sp. NIES-981	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Cyanobium* sp. PCC 7001	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Prochlorococcus marinus* NATL2A	-	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Prochlorococcus marinus* MIT9312	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Prochlorococcus marinus* MIT9515	-	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Prochlorococcus marinus* NATL1A	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Prochlorococcus* sp. MIT0801	-	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Prochlorococcus* sp. MIT1306	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Prochlorococcus* sp. P1363	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Prochlorococcus* sp. SS52	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Prochlorococcus marinus* MIT9211	-	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Prochlorococcus marinus* MIT9303	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Prochlorococcus marinus* MIT9313	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Prochlorococcus marinus* CCMP1375	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Prochlorococcus marinus* AS9601	-	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Prochlorococcus marinus* MIT9215	-	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Prochlorococcus marinus* MIT9301	-	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Prochlorococcus marinus* MIT9302	-	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Prochlorococcus marinus* pastoris CCMP1986	-	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Prochlorococcus* sp. MIT0604	-	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Prochlorococcus* sp. RS50	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

Table B.14 continued from previous page

Group	Genome	Ala	Arg	Asn-*asnB*	Asn-*aspS*/*gatABC*	Asp	Cys	Gln	Glu	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val
M-Pcy	*Prochlorococcus* sp. RS01	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Prochlorococcus* sp. RS04	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. NIES970	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. NKBG15041c	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. PCC 7002	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. OG1	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. PCC 7003	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. PCC 7117	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. PCC 11901	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. NKBG042902	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. PCC 8807	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. PCC 73109	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. BDU 130192	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. CB0101	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. CC9311	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. CC9605	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. CC9902	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. KORDI-100	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. KORDI-49	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. KORDI-52	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. RCC307	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. BL107	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. BS55D	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. BS56D	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. CC9616	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. GEYO	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. HB1133	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. MITS9504	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. MITS9508	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

Table B.14 continued from previous page

Group	Genome	Ala	Arg	Asn-*asnB*	Asn-*aspS/gatABC*	Asp	Cys	Gln	Glu	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val
M-Pcy	*Synechococcus* sp. MITS9509	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. N26	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. PCC 7335	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. PCC 7336	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. RS9916	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. RS9917	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. RSCCF101	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. UW105	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. UW140	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. UW179A	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. UW179B	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. UW69	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. UW86	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. WH5701	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. WH7805	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. WH8016	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. XM-24	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. YX04-3	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. A15-127	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. A15-44	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. BMK-MC-1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. LTW-R	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. NOUM97013	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. SYN20	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. SynAce01	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. WH7803	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. WH8020	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. WH8101	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. WH8102	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

Table B.14 continued from previous page

Group	Genome	Ala	Arg	Asn-*asnB*	Asn-*aspS*/*gatABC*	Asp	Cys	Gln	Glu	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val
M-Pcy	*Synechococcus* sp. WH8103	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
M-Pcy	*Synechococcus* sp. WH8109	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Anabaena* sp. 90	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Anabaena* sp. FACHB-1391	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Anabaena* sp. WA102	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Aphanizomenon flos-aquae* FACHB-1287	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Aphanizomenon flos-aquae* LD13	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Aphanizomenon flos-aquae* NIES-81	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Aphanizomenon* sp. FACHB-1416	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Arthrospira platensis* C1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Arthrospira platensis* FACHB-971	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Cylindrospermopsis raciborskii* MVCC14	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+
LF-Cy	*Cylindrospermopsis raciborskii* CS-505	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Cylindrospermopsis raciborskii* ITEP-A1	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Cylindrospermopsis raciborskii* S07	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Dolichospermum* sp. UHCC 0352	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Dolichospermum compactum* NIES-806	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Dolichospermum* sp. FACHB-1091	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Geminocystis herdmanii* PCC 6308	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Geminocystis* sp. NIES-3708	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Geminocystis* sp. NIES-3709	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Gloeocapsa* sp. PCC 73106	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Gloeocapsa* sp. PCC 7428	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Microcystis aeruginosa* KLA2	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Microcystis aeruginosa* NIES-2481	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Microcystis aeruginosa* PCC 7005	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Microcystis aeruginosa* PCC 7806SL	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Microcystis panniformis* FACHB-1757	+	+	-	+	+	+	+	+	-	+	-	+	+	+	+	+	+	+	+	+	-
LF-Cy	*Microcystis* sp. 0824	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

Table B.14 continued from previous page

Group	Genome	Ala	Arg	Asn-*asnB*	Asn-*aspS*/*gatABC*	Asp	Cys	Gln	Glu	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val
LF-Cy	*Microcystis* sp. MC19	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Nostoc paludosum* FACHB-159	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Nostoc* sp. PCC 7107	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Nostoc* sp. PCC 7120	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Nostoc* sp. PCC 7524	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Nostoc sphaeroides* Kutzing En	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+
LF-Cy	*Planktothricoides* sp. FACHB-1261	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Planktothricoides* sp. FACHB-1370	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Planktothricoides* sp. SR001	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Planktothrix prolifica* NIVA-CYA 406	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Planktothrix agardhii* NIES-204	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Planktothrix agardhii* NIVA-CYA 1268	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Planktothrix rubescens* NIVA-CYA 407	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Planktothrix* sp. PCC 11201	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Pleurocapsa* sp. PCC 7319	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Pleurocapsa* sp. PCC 7327	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Prochlorothrix hollandica* PCC 9006	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Pseudanabaena* sp. ABRG5-3	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Pseudanabaena* sp. FACHB-723	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Pseudanabaena* sp. PCC 6802	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Synechocystis* sp. FACHB-929	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Synechocystis* sp. PCC 6714	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Synechocystis* sp. PCC 6803	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
LF-Cy	*Synechocystis* sp. CACIAM 05	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

Table B.15: **FDR-adjusted p-values from hypergeometric test analyses for gene enrichment between population groups.** Populations are categorised as follows: 1) F-Pcy; 2) SE-Cy; 3) T-Pcy; 4) M-Pcy; 5) LF-Cy; A) M-Pcy *Syn/Pro Synechococcus*; B) M-Pcy *β-Synechococcus*; C) M-Pcy *Prochlorococcus.*

Gene	**1 - 2**	**1 - 3**	**1 - 4**	**1 - A**	**1 - B**	**1 - C**	**1 - 5**	**A - B**	**A - C**	**B - C**
ntcA	1.000000	1.000000	1.000000	1.000000	0.983572	1.000000	0.000000	0.584790	1.000000	0.913167
ntcB	0.000003	0.000007	0.253477	1.000000	0.000000	1.000000	1.000000	0.000000	1.000000	0.000000
glnB	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
icd	1.000000	1.000000	0.299500	1.000000	0.000000	1.000000	1.000000	0.000000	1.000000	0.000000
pipX	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	0.009812	1.000000	0.858120	1.000000
spkE	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	0.120890	1.000000	1.000000	1.000000
amt1	0.913167	1.000000	1.000000	1.000000	1.000000	1.000000	0.858120	1.000000	1.000000	1.000000
amt2	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	0.062325	1.000000	1.000000	1.000000
amt3	1.000000	1.000000	0.521435	1.000000	0.000052	1.000000	0.120890	0.000000	1.000000	0.000014
amt4	0.913167	1.000000	1.000000	1.000000	1.000000	1.000000	0.241108	1.000000	1.000000	1.000000
amtB	0.959666	0.000863	0.000000	0.000042	0.000106	0.000007	0.635807	0.422446	0.336543	1.000000
nrtA	1.000000	1.000000	0.000000	0.000000	0.001884	0.000000	0.971256	0.290514	1.000000	0.371062
nrtB	1.000000	1.000000	0.000000	0.000000	1.000000	0.000000	0.812373	0.000000	1.000000	0.000002
nrtC	0.758594	0.008018	0.000001	0.000000	0.943478	0.000001	0.769685	0.000000	1.000000	0.000002
nrtD	1.000000	1.000000	0.000000	0.000000	0.000417	0.000000	0.000000	0.713834	0.652879	0.371062
nrtP	1.000000	0.000007	0.021857	0.000000	0.000057	1.000000	0.731293	1.000000	0.000000	0.000013
focA	1.000000	1.000000	0.000014	0.000000	0.000052	0.038372	0.166130	0.872266	0.002113	0.043992
nrtS	1.000000	1.000000	0.299500	1.000000	0.000000	1.000000	0.088437	0.000000	1.000000	0.000000
cynA	0.008647	1.000000	0.110836	0.603907	0.000000	1.000000	0.002060	0.000000	0.839611	0.000001
cynB	0.008647	1.000000	0.110836	0.603907	0.000000	1.000000	0.001340	0.000000	0.839611	0.000001
cynD	0.008647	0.872266	0.021857	0.731293	0.000025	0.839611	0.000267	0.000006	0.755367	0.000068
cynS	0.031185	1.000000	0.000016	0.328612	0.388217	0.002587	0.000000	1.000000	0.000000	0.000068
cynH	0.858120	0.002400	0.012087	0.000012	1.000000	1.000000	0.000000	0.000548	0.000001	1.000000
urtA	1.000000	0.000007	0.000000	1.000000	0.983572	0.282879	0.858120	0.584790	0.031185	0.881367
urtB	0.008647	1.000000	0.967103	1.000000	1.000000	0.282879	0.668274	1.000000	0.106945	0.521435
urtC	0.008647	1.000000	0.521435	0.808785	0.983572	0.282879	0.858120	1.000000	0.472230	0.881367
urtD	0.008647	1.000000	0.858120	1.000000	0.983572	0.282879	0.521435	0.862108	0.106945	0.881367
urtE	0.008647	0.000007	0.521435	1.000000	0.189562	0.282879	0.120890	0.148785	0.220063	0.839611
natA	0.008647	0.320426	0.000000	1.000000	0.000000	1.000000	0.000000	0.000000	1.000000	0.000000
natB	0.000003	0.000007	0.366374	1.000000	0.000000	1.000000	0.000000	0.000000	1.000000	0.000000
natC	0.000039	0.000007	0.366374	1.000000	0.000000	1.000000	0.000000	0.000000	1.000000	0.000000
natD	0.000003	0.000007	0.440232	1.000000	0.000007	1.000000	0.000000	0.000000	1.000000	0.000002
natE	0.000003	0.000007	0.440232	1.000000	0.000007	1.000000	0.000000	0.000000	1.000000	0.000002
natF	0.000003	0.023735	0.000000	1.000000	0.000000	0.003774	0.396544	0.000000	0.000312	0.012087
natG	0.371062	0.023735	0.035160	1.000000	0.000000	0.008591	0.298117	0.000000	0.001146	0.006343
natH	0.371062	0.023735	0.035160	1.000000	0.000000	0.008591	1.000000	0.000000	0.001146	0.006343
bgtA	0.371062	0.023735	0.035160	1.000000	0.000000	0.008591	0.021367	0.000000	0.001146	0.006343
bgtB	0.371062	0.000007	0.000000	0.986522	1.000000	1.000000	0.009940	1.000000	1.000000	1.000000
natI	1.000000	1.000000	0.589169	0.000000	1.000000	0.000000	0.000206	0.000000	1.000000	0.000000
natJ	0.913167	0.000007	0.000000	0.000000	1.000000	0.000000	0.000206	0.000000	1.000000	0.000000
natK	1.000000	0.000007	0.000000	0.000000	1.000000	0.000000	0.000123	0.000000	1.000000	0.000000
natL	1.000000	0.000105	0.000000	0.000000	1.000000	0.000000	0.000334	0.000000	1.000000	0.000000

Table B.15 continued from previous page

Gene	1 - 2	1 - 3	1 - 4	1 - A	1 - B	1 - C	1 - 5	A - B	A - C	B - C
natM	1.000000	0.000007	0.000000	0.000000	1.000000	0.000000	0.000546	0.000000	0.879155	0.000000
gtrA	1.000000	0.000007	0.000001	0.004022	1.000000	0.282879	0.000123	0.031185	0.195012	0.521435
gtrB	1.000000	1.000000	0.017051	0.004022	1.000000	0.282879	0.000000	0.031185	0.195012	0.521435
gtrC	1.000000	0.000007	0.253477	1.000000	0.000000	1.000000	0.088437	0.000000	1.000000	0.000000
gltS	0.001955	1.000000	0.732736	0.000046	0.001136	0.000765	0.000001	0.899685	0.894401	1.000000
dmeA	0.023199	0.005864	0.000000	1.000000	1.000000	1.000000	0.227376	1.000000	1.000000	1.000000
agcS	0.008647	1.000000	0.000002	0.000000	1.000000	0.153929	0.000000	0.000000	0.000004	0.132932
oppC	1.000000	0.000054	0.589169	1.000000	0.000000	1.000000	0.668274	0.000000	1.000000	0.000000
dppA	1.000000	0.000007	1.000000	1.000000	1.000000	1.000000	0.000000	1.000000	1.000000	1.000000
dppB	0.000003	1.000000	0.000000	1.000000	1.000000	1.000000	0.668274	1.000000	1.000000	1.000000
dppC	1.000000	0.000007	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
dppD	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	0.000002	1.000000	1.000000	1.000000
dppF	1.000000	0.000007	0.000000	1.000000	0.983572	1.000000	1.000000	0.862108	1.000000	0.913167
glnA	1.000000	1.000000	1.000000	0.986522	1.000000	1.000000	0.007005	1.000000	1.000000	1.000000
glnN	1.000000	1.000000	0.002149	0.000074	1.000000	0.000000	0.731293	0.001986	0.000684	0.000000
glsF	1.000000	1.000000	0.540521	1.000000	1.000000	1.000000	0.001340	1.000000	1.000000	1.000000
gltD	0.125156	0.000007	0.000000	1.000000	1.000000	1.000000	0.521435	1.000000	1.000000	1.000000
gifA	1.000000	1.000000	0.000000	1.000000	0.000000	1.000000	0.858120	0.000000	1.000000	0.000000
gifB	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	0.000334	1.000000	1.000000	1.000000
gdhA	0.000325	1.000000	0.858120	1.000000	1.000000	0.282879	0.000000	1.000000	0.220063	0.521435
narB	1.000000	0.000007	0.000000	1.000000	1.000000	0.000000	0.000000	1.000000	0.000000	0.000001
nirA	1.000000	0.000007	0.000000	1.000000	1.000000	0.000228	0.000000	1.000000	0.000000	0.002999
nirB	1.000000	0.000007	0.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
ureA	1.000000	0.005210	0.967103	1.000000	1.000000	0.282879	0.000000	1.000000	0.106945	0.521435
ureB	0.008647	0.000007	0.000000	1.000000	1.000000	0.282879	1.000000	1.000000	0.106945	0.521435
ureC	0.008647	0.005210	0.967103	1.000000	1.000000	0.282879	1.000000	1.000000	0.106945	0.521435
ureD	0.008647	1.000000	0.858120	1.000000	0.983572	0.282879	1.000000	0.862108	0.106945	0.881367
ureE	0.008647	0.005210	0.732736	1.000000	1.000000	0.153929	1.000000	1.000000	0.091830	0.344852
ureF	0.008647	0.005210	0.858120	1.000000	0.983572	0.282879	1.000000	0.862108	0.106945	0.881367
ureG	0.008647	0.005210	0.967103	1.000000	1.000000	0.282879	0.000206	1.000000	0.106945	0.521435
ureJ	0.008647	1.000000	0.008279	0.000042	0.971256	0.026621	1.000000	0.000939	0.227376	0.080634
carA	0.008647	1.000000	1.000000	1.000000	1.000000	1.000000	0.000000	1.000000	1.000000	1.000000
carB	0.205065	0.000863	0.000000	1.000000	1.000000	1.000000	0.000000	1.000000	1.000000	1.000000
narM	1.000000	0.000007	0.000000	1.000000	0.000000	1.000000	1.000000	0.000000	1.000000	0.000000
cphA	1.000000	0.000105	0.000000	0.986522	0.000005	1.000000	0.000000	0.000000	1.000000	0.000000
cphB	0.001955	0.872266	0.299500	1.000000	0.000000	1.000000	0.000000	0.000000	1.000000	0.000000
iadA	0.125862	1.000000	0.738434	1.000000	0.000000	1.000000	0.000123	0.000000	1.000000	0.000000
nblA	0.035185	1.000000	0.366374	1.000000	0.000052	1.000000	0.000000	0.000002	1.000000	0.000014
speA	0.035185	0.000007	0.000000	1.000000	1.000000	1.000000	0.001340	1.000000	1.000000	1.000000
speB	0.001955	0.000007	0.000000	1.000000	1.000000	1.000000	0.000000	1.000000	1.000000	1.000000
aguA	1.000000	0.000007	0.000000	1.000000	1.000000	1.000000	0.000000	1.000000	1.000000	1.000000
aguB	0.000003	0.000007	0.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
puo	0.008647	1.000000	1.000000	0.245830	0.930580	0.584790	0.000000	1.000000	1.000000	1.000000
putA	0.000003	1.000000	0.000000	1.000000	0.000000	1.000000	0.000000	0.000000	1.000000	0.000000
arcA	0.105180	1.000000	0.839611	1.000000	0.000007	1.000000	0.000000	0.000000	1.000000	0.000002
arcC	0.125156	0.000007	0.000000	1.000000	1.000000	1.000000	0.009812	1.000000	1.000000	1.000000

Table B.15 continued from previous page

Gene	1 - 2	1 - 3	1 - 4	1 - A	1 - B	1 - C	1 - 5	A - B	A - C	B - C
aoxA	0.035185	1.000000	0.612460	0.000051	0.000373	0.000004	0.858120	0.625686	0.238425	1.000000
syNOS	1.000000	1.000000	1.000000	0.986522	1.000000	1.000000	0.055433	1.000000	1.000000	1.000000
nifD	0.001262	0.930580	0.000000	0.492311	0.930580	0.584790	0.021346	1.000000	1.000000	1.000000
nifK	1.000000	0.000721	0.831894	0.492311	0.930580	0.584790	0.134810	1.000000	1.000000	1.000000
fdxN	1.000000	1.000000	0.112108	1.000000	1.000000	1.000000	0.000000	1.000000	1.000000	1.000000
nifU	1.000000	0.000250	0.000000	1.000000	1.000000	1.000000	0.002060	1.000000	1.000000	1.000000
nifB	1.000000	0.000105	1.000000	0.492311	0.930580	0.584790	0.000073	1.000000	1.000000	1.000000
nifE	1.000000	0.872266	0.000000	0.492311	0.930580	0.584790	0.058087	1.000000	1.000000	1.000000
nifH	1.000000	0.002734	0.260858	0.492311	0.930580	0.584790	0.174300	1.000000	1.000000	1.000000
nifN	1.000000	0.002734	0.403485	0.684993	0.930580	0.584790	0.079300	1.000000	1.000000	1.000000
nifX	1.000000	0.002734	0.112108	0.245830	0.930580	0.584790	0.079300	1.000000	1.000000	1.000000
nifW	1.000000	0.002734	0.260858	1.000000	1.000000	1.000000	0.221307	1.000000	1.000000	1.000000
nifV	1.000000	0.002734	0.112108	1.000000	1.000000	1.000000	0.004735	1.000000	1.000000	1.000000
nifZ	1.000000	0.000105	1.000000	0.492311	0.930580	0.584790	0.002060	1.000000	1.000000	1.000000
nifT	1.000000	0.000105	1.000000	0.492311	0.930580	0.584790	0.041914	1.000000	1.000000	1.000000
chdA	1.000000	0.002734	0.260858	0.492311	0.930580	0.584790	0.221701	1.000000	1.000000	1.000000
chiA	1.000000	0.000250	0.403485	0.000000	0.001136	0.000016	0.784091	1.000000	1.000000	1.000000
choA	1.000000	1.000000	0.285990	0.001038	1.000000	0.849253	0.003030	0.011758	0.009328	1.000000
phnD	0.003261	0.039549	0.429168	0.983572	1.000000	0.001590	0.396544	1.000000	0.000277	0.012087
amiA	0.008647	1.000000	0.000000	1.000000	0.473986	1.000000	0.000000	0.290514	1.000000	0.687005
ald	1.000000	0.000007	0.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
sufS	0.000003	1.000000	0.135776	1.000000	1.000000	0.000024	0.000000	1.000000	0.000000	0.000555
ilvB	0.008647	0.000007	0.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
ilvN	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
ilvC	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
ilvD	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	0.000043	1.000000	1.000000	1.000000
livE	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
leuA	1.000000	1.000000	1.000000	1.000000	0.983572	1.000000	0.858120	0.862108	1.000000	0.913167
leuD	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
leuC	1.000000	1.000000	1.000000	1.000000	0.983572	1.000000	1.000000	0.584790	1.000000	0.913167
leuB	0.913167	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
metH	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
pheA	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
trpE	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
trpG	1.000000	0.872266	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
trpD	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
trpF	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
trpC	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
trpA	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
trpB	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
tyrA	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	0.000206	1.000000	1.000000	1.000000
asnB	1.000000	0.000007	0.000678	0.388455	0.007160	0.282140	1.000000	0.000040	0.006745	0.132932
asnS	1.000000	0.000007	0.000000	1.000000	0.000000	1.000000	0.000124	0.000000	1.000000	0.000000
aspS	0.260858	0.153929	0.000010	1.000000	1.000000	1.000000	0.000000	1.000000	1.000000	1.000000
gatA	0.000003	0.000007	0.000000	1.000000	0.473986	1.000000	1.000000	0.459108	1.000000	0.687005
gatB	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000

Table B.15 continued from previous page

Gene	1 - 2	1 - 3	1 - 4	1 - A	1 - B	1 - C	1 - 5	A - B	A - C	B - C
gatC	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
cysE	1.000000	1.000000	1.000000	0.245830	0.930580	0.584790	1.000000	1.000000	1.000000	1.000000
cysM1	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	0.221307	1.000000	1.000000	1.000000
cysM2	1.000000	0.000250	0.022048	0.003468	1.000000	0.000730	1.000000	0.006175	0.357217	0.001343
serA	0.913167	1.000000	1.000000	1.000000	1.000000	1.000000	0.971256	1.000000	1.000000	1.000000
glyA	1.000000	1.000000	0.540521	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
aspC	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
lysC	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
asd	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
thrA	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
thrB	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
thrC1	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
thrC2	1.000000	1.000000	0.000000	0.000000	1.000000	0.000000	1.000000	0.000003	0.161165	0.000000
argJ	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
argB	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
argC	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
argD	1.000000	1.000000	1.000000	0.983572	1.000000	1.000000	1.000000	1.000000	0.879155	1.000000
argF	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
argG	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
argH	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
kprS	1.000000	1.000000	1.000000	1.000000	0.983572	1.000000	1.000000	0.584790	1.000000	0.913167
hisZ	1.000000	0.872266	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
hisG	1.000000	1.000000	1.000000	1.000000	0.983572	1.000000	1.000000	0.584790	1.000000	0.913167
hisIE	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
hisA	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	0.858120	1.000000
hisF	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
hisH	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
hisB	0.913167	1.000000	1.000000	1.000000	1.000000	1.000000	0.021367	1.000000	1.000000	1.000000
hisC	1.000000	0.872266	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
hisD	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
dapA	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
dapB	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
aspB	1.000000	1.000000	1.000000	0.769150	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
dapF	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	0.731293	1.000000	0.858120	1.000000
lysA	1.000000	1.000000	0.540521	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
proB	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
proA	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
proC	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000

Table B.16: **AAT substrate preferences**

Transporter	Favoured Substrates	Other Substrates	References
N-I	P, F, L	A, E, G, H, M, N, Q, S, T	[253, 256, 257, 259]
N-II	D	A, E, G, H, M, N, Q, S, T	[254, 256, 258]
Bgt	K, R	H, Q	[254, 259]
N-III	G	A, E, F, L, P, Q	[255]
Gtr	E		[259]
GltS	E		[259]
DmeA	D, E, Q		[258]
AgcS	G, N, Q	A, C, M, S	[260]

Table B.17: **Brands and purity of nutrients utilised for growth assays.** Chitin purity was not determined by supplier.

Nutrient	Brand	Purity
Alanine	Thermo Scientific Alfa Aesar	99%
Arginine	Thermo Scientific Alfa Aesar	98%
Asparagine	MP Biomedicals	99%
Aspartic Acid	Thermo Scientific Acros	97%
Cysteine	Sigma Aldrich	97%
Glutamic Acid	Thermo Scientific Alfa Aesar	98%
Glutamine	Thermo Scientific Acros	98%
Glycine	Thermo Scientific Acros	99%
Histidine	MP Biomedicals	99%
Isoleucine	Fisher Bioreagents	98%
Leucine	Merck	98%
Lysine	Thermo Scientific Alfa Aesar	99%
Methionine	Thermo Scientific Acros	98%
Phenylalanine	Thermo Scientific Acros	98%
Proline	Thermo Scientific Alfa Aesar	99%
Serine	Thermo Scientific Alfa Aesar	99%
Threonine	Thermo Scientific Acros	98%
Tryptophan	Thermo Scientific Alfa Aesar	99%
Tyrosine	Thermo Scientific Alfa Aesar	99%
Valine	Thermo Scientific Acros	99%
Chitin	Merck	ND
Glyphosate	Sigma Aldrich	96%
Ammonium	VWR	98%
Nitrate	VWR	99%
Urea	VWR	99%

Table B.18: **FDR-adjusted p-values from t-test analyses for growth rates of *Synechococcus* sp. CCY9618 and *Synechococcus* sp. CCAP1479/10 grown on 25 different nitrogen sources.** Significant differences ($q < .05$) are highlighted in green. High concentration of N is 250 mg N L^{-1}; Low concentration of N is 1 mg N L^{-1}. High concentration of tyrosine was reduced to 25 mg N L^{-1} due to poor solubility.

	High concentration significance between species			Low concentration significance between species			Significance between CCY9618 concentrations			Significance between CCAP1479/10 concentrations		
Substrate	**df**	**t**	**p adjusted (q)**	**df**	**t**	**p adjusted (q)**	**df**	**t**	**p adjusted (q)**	**df**	**t**	**p adjusted (q)**
Alanine	4	5.5731	0.044973	4	1.9961	0.293771	4	4.7543	0.071942	4	4.4875	0.078319
Arginine	4	0.1781	0.929592	4	2.1527	0.266955	4	2.1612	0.268274	4	0.7142	0.702908
Asparagine	4	2.6286	0.2134	4	0.2244	0.923994	4	4.0866	0.088182	4	4.6383	0.072377
Aspartate	4	3.4157	0.13381	4	1.6124	0.37603	4	0.7591	0.684024	4	3.2655	0.142729
Cysteine	-	-	-	-	-	-	-	-	-	-	-	-
Glutamate	4	1.7278	0.3589	4	1.0703	0.57172	4	2.1373	0.264159	4	10.3534	0.010778
Glutamine	4	0.0275	0.989602	4	1.2594	0.491941	4	2.5097	0.228989	4	0.734	0.697846
Glycine	4	0.9542	0.601858	4	0.8847	0.636171	4	2.3116	0.237143	4	2.6785	0.210357
Histidine	4	0.1118	0.955815	4	0.8006	0.667874	4	1.4886	0.413083	4	0.4083	0.837771
Isoleucine	4	1.0533	0.578054	4	0.334	0.872076	4	2.443	0.233458	4	0.8038	0.675381
Leucine	4	1.1924	0.5044	4	0.5715	0.753577	4	2.4641	0.232131	4	2.6035	0.214837
Lysine	4	0.726	0.69895	4	0.6915	0.700659	4	2.2658	0.245638	4	1.6278	0.373189
Methionine	-	-	-	-	-	-	4	0.3698	0.858656	-	-	-
Phenylalanine	4	0.2633	0.908303	4	0.8317	0.664744	4	0.8782	0.635905	4	1.6587	0.367747
Proline	4	2.1243	0.264522	4	2.1435	0.265942	4	3.0384	0.16975	4	4.6638	0.074496
Serine	4	0.2324	0.92806	4	0.2271	0.927075	4	1.4681	0.41904	4	1.3593	0.449494
Threonine	-	-	-	-	-	-	-	-	-	-	-	-
Tryptophan	4	1.8361	0.335788	4	0.2865	0.894896	4	1.6781	0.363427	4	3.0155	0.165743
Tyrosine	-	-	-	-	-	-	-	-	-	-	-	-
Valine	4	0.2055	0.923353	4	0.7838	0.67546	4	0.3476	0.866146	4	0.6058	0.741697
Chitin	-	-	-	-	-	-	-	-	-	-	-	-
Glyphosate	-	-	-	-	-	-	-	-	-	-	-	-
Nitrate	4	1.2127	0.50131	4	0.6227	0.733579	4	0.8986	0.631177	4	1.7224	0.352948
Ammonium	4	1.0517	0.574338	4	0.2208	0.92161	4	1.3656	0.45045	4	1.7681	0.346461
Urea	4	0.972	0.60406	4	0.1919	0.923871	4	1.1945	0.507633	4	0.1125	0.960457
No N	-	-	-	-	-	-	-	-	-	-	-	-

Table B.19: **FDR-adjusted p-values from t-test analyses for lag phases of *Synechococcus* sp. CCY9618 and *Synechococcus* sp. CCAP1479/10 grown on 25 different nitrogen sources.** Significant differences ($q < .05$) are highlighted in green. High concentration of N is 250 mg N L^{-1}; Low concentration of N is 1 mg N L^{-1}. High concentration of tyrosine was reduced to 25 mg N L^{-1} due to poor solubility.

	High concentration significance between species			**Low concentration significance between species**			**Significance between CCY9618 concentrations**			**Significance between CCAP1479/10 concentrations**		
Substrate	**df**	**t**	**p adjusted (q)**	**df**	**t**	**p adjusted (q)**	**df**	**t**	**p adjusted (q)**	**df**	**t**	**p adjusted (q)**
Alanine	4	0.993	0.599333	4	0.2206	0.916513	4	0.4624	0.809708	4	1.2151	0.5044
Arginine	4	4.1487	0.08949	4	0.9865	0.598876	4	0.1738	0.927896	4	2.3696	0.240623
Asparagine	4	3.8162	0.098573	4	0.2923	0.895367	4	1.7263	0.355444	4	6.7262	0.025526
Aspartate	4	19.691	0.00485	4	0.3273	0.872196	4	2.4047	0.235344	4	7.9573	0.018107
Cysteine	-	-	-	-	-	-	-	-	-	-	-	-
Glutamate	4	4.0307	0.087023	4	0.0739	0.964486	4	3.2863	0.143371	4	23.0671	0.00388
Glutamine	4	3.3322	0.14065	4	0.1106	0.951534	4	1.3748	0.449931	4	12.7004	0.005543
Glycine	4	7.1542	0.021556	4	0.6911	0.696156	4	0.5233	0.781472	4	1.0077	0.594185
Histidine	4	8.6029	0.013857	4	1.1205	0.543869	4	7.4283	0.020541	4	9.827	0.010582
Isoleucine	4	4.0765	0.086159	4	1.8859	0.325134	4	0.7124	0.694628	4	4.2202	0.0873
Leucine	4	2.6726	0.207804	4	2.4274	0.233447	4	2.6023	0.211284	4	2.7515	0.203106
Lysine	4	3.9514	0.090533	4	1.4931	0.41512	4	0.0717	0.961163	4	0.3734	0.860934
Methionine	-	-	-	-	-	-	4	1.2327	0.502989	-	-	-
Phenylalanine	4	10.5373	0.0097	4	1.6424	0.37092	4	2.3149	0.239855	4	7.6414	0.0194
Proline	4	11.8716	0.007275	4	5.4262	0.047235	4	4.1035	0.089725	4	31.3152	0.003233
Serine	4	0.4346	0.82685	4	1.3155	0.464702	4	1.3986	0.44168	4	1.4568	0.420461
Threonine	-	-	-	-	-	-	-	-	-	-	-	-
Tryptophan	4	2.9596	0.171711	4	9.7797	0.0097	4	2.8669	0.1843	4	24.8494	0.0097
Tyrosine	-	-	-	-	-	-	-	-	-	-	-	-
Valine	4	6.3501	0.03104	4	0.7783	0.674642	4	0.8038	0.670521	4	40.3658	0.006467
Chitin	-	-	-	-	-	-	-	-	-	-	-	-
Glyphosate	-	-	-	-	-	-	-	-	-	-	-	-
Nitrate	4	1.7737	0.348276	4	0.1051	0.950703	4	1.4362	0.42661	4	4.2757	0.086297
Ammonium	4	3.5478	0.121505	4	9.1616	0.011938	4	6.068	0.034181	4	18.9968	0.0194
Urea	4	2.6898	0.212236	4	0.8082	0.67725	4	2.0139	0.291766	4	2.3253	0.244622
No N	-	-	-	-	-	-	-	-	-	-	-	-

Table B.20: **Summary of quantified proteins**. UPN: Unique peptide number; PSM: Peptide Spectrum Matches. TMT was used to quantitatively determine the proteomic response when grown on AA N-substrate compared to NO_3^- and N-starvation. All proteins were detected in triplicate in all conditions.

Protein Accession	**Coverage (%)**	**UPN**	**PSM**	**Gene**	**COG**	**Description**
Ga0436389_001_101384_101671	37.89	3	5	-	S	hypothetical protein
Ga0436389_001_105813_106616	7.49	1	2	-	S	Glutamine amidotransferases class-II
Ga0436389_001_107869_108948	33.15	11	35	*sbpA*	P	sulfate ABC transporter substrate-binding protein
Ga0436389_001_1594_2124	19.32	4	8	*ilvN*	E	acetolactate synthase small subunit
Ga0436389_001_17070_17834	60.63	9	32	-	S	hypothetical protein
Ga0436389_001_20756_21772	16.57	4	7	-	O	O-linked N-acetylglucosamine transferase SPINDLY family
Ga0436389_001_265_831	22.34	4	5	*ycf4*	U	photosystem I assembly protein Ycf4
Ga0436389_001_34324_36078	11.30	6	11	-	V	Glycosyltransferase
Ga0436389_001_36075_37373	10.42	4	8	-	M	Glycosyltransferase
Ga0436389_001_38108_38986	8.22	2	3	-	S	polysaccharide pyruvyl transferase family protein
Ga0436389_001_39028_39897	26.64	7	12	-	S	beta-1,6-N-acetylglucosaminyltransferase
Ga0436389_001_42666_43829	28.68	8	22	*rfbB*	M	dTDP-glucose 4,6-dehydratase
Ga0436389_001_43859_44758	18.06	3	4	-	M	dTDP-4-dehydrorhamnose reductase
Ga0436389_001_48625_50067	13.75	5	10	*gltX*	J	glutamate–tRNA ligase
Ga0436389_001_50405_50593	14.52	1	2	-	S	hypothetical protein
Ga0436389_001_50769_51245	36.71	5	30	*rplS*	J	50S ribosomal protein L19
Ga0436389_001_51713_52573	6.29	2	7	*map*	E	type I methionyl aminopeptidase
Ga0436389_001_53396_54502	7.34	2	4	*pta*	C	phosphotransacetylase family protein
Ga0436389_001_54542_55054	33.53	6	26	-	LU	DNA recombination-mediator protein A
Ga0436389_001_55630_56544	38.49	9	33	-	O	paraslipin
Ga0436389_001_57116_57607	43.56	6	15	-	S	YajQ family cyclic di-GMP-binding protein
Ga0436389_001_6633_6902	17.98	2	4	*infA*	J	translation initiation factor IF-1
Ga0436389_001_77097_77483	8.59	1	7	-	S	hypothetical protein
Ga0436389_001_834_1565	38.27	8	22	*ppiB*	M	peptidylprolyl isomerase
Ga0436389_001_85235_86230	8.46	3	5	-	P	zinc ABC transporter substrate-binding protein
Ga0436389_001_8994_10151	37.14	10	16	*pntA*	C	NAD(P) transhydrogenase subunit alpha
Ga0436389_001_97188_98384	6.53	2	4	*natI*	E	ABC transporter substrate-binding protein

Table B.20 continued from previous page

Protein Accession	Coverage (%)	UPN	PSM	Gene	COG	Description
Ga0436389_001_98381_99709	33.26	11	23	*hemL*	H	glutamate-1-semialdehyde 2,1-aminomutase
Ga0436389_002_8000_9232	16.10	6	14	-	C	NAD(P)/FAD-dependent oxidoreductase
Ga0436389_003_13722_14702	28.83	5	15	*cysK*	E	cysteine synthase A
Ga0436389_003_14740_15174	14.58	2	3	-	S	hypothetical protein
Ga0436389_003_15567_16925	21.46	7	21	*gor*	C	glutathione-disulfide reductase
Ga0436389_003_22524_23453	14.24	4	7	-	T	family 2A encapsulin nanocompartment shell protein
Ga0436389_003_25618_26139	18.50	2	5	*cpcB*	C	phycocyanin subunit beta
Ga0436389_003_33834_34154	75.47	6	67	-	K	P-II family nitrogen regulator
Ga0436389_003_35798_36316	72.67	5	203	*cpcB*	C	phycocyanin subunit beta
Ga0436389_003_38378_38656	28.26	2	6	-	S	DUF427 domain-containing protein
Ga0436389_003_40168_40587	5.76	1	3	-	P	tellurite resistance TerB family protein
Ga0436389_003_41071_43239	30.06	17	143	*glnN*	S	glutamine synthetase III
Ga0436389_003_45449_46516	16.06	4	5	-	C	alkene reductase
Ga0436389_003_56891_57901	19.05	5	8	-	P	ABC transporter substrate-binding protein
Ga0436389_003_63117_63968	16.25	4	6	-	S	Polyphosphate kinase 2 (PPK2)
Ga0436389_003_83842_84417	8.38	1	4	-	P	ferritin
Ga0436389_004_29599_30195	11.11	2	2	*psbP*	S	photosystem II reaction center PsbP
Ga0436389_004_36659_37756	25.48	8	14	-	T	sensor domain-containing diguanylate cyclase
Ga0436389_004_4451_6448	9.02	6	8	*cmpC*	P	ABC transporter substrate-binding protein
Ga0436389_004_46165_47994	1.31	1	2	-	V	ABC transporter
Ga0436389_004_47991_48395	69.40	5	70	*rbpD*	S	RNA-binding protein
Ga0436389_004_49277_51235	8.74	4	5	*deaD*	F	DEAD/DEAH box helicase
Ga0436389_004_62053_62373	20.75	2	6	*ycf54*	S	DUF2488 family protein
Ga0436389_004_64037_64612	34.55	7	29	-	S	hypothetical protein
Ga0436389_004_64917_65240	32.71	3	8	*ycf64*	C	Grx4 family monothiol glutaredoxin
Ga0436389_004_65878_66549	19.73	3	5	-	K	response regulator transcription factor
Ga0436389_004_71443_71565	27.50	1	5	*psbY*	U	photosystem II protein Y
Ga0436389_004_73565_74782	7.90	1	11	-	M	carbohydrate porin
Ga0436389_004_7569_8897	23.98	10	32	*cmpA*	P	ABC transporter substrate-binding protein
Ga0436389_005_16182_17213	8.16	2	3	-	S	hypothetical protein
Ga0436389_005_17805_19529	5.23	2	3	-	M	tandem-95 repeat protein

Table B.20 continued from previous page

Protein Accession	Coverage (%)	UPN	PSM	Gene	COG	Description
Ga0436389_005_2951_3700	24.90	3	11	-	S	DUF305 domain-containing protein
Ga0436389_005_618_1670	21.14	6	7	-	S	hypothetical protein
Ga0436389_006_102407_103084	29.33	5	16	-	S	hypothetical protein
Ga0436389_006_109737_112418	15.23	12	18	*alaS*	J	alanine-tRNA ligase
Ga0436389_006_116837_119929	1.75	2	7	*hepA*	KL	DEAD/DEAH box helicase
Ga0436389_006_125743_127530	10.76	6	11	*dfa1*	C	diflavin flavoprotein A
Ga0436389_006_127530_129344	8.94	5	7	*dfa3*	C	flavin reductase
Ga0436389_006_132039_134975	3.78	3	8	*gcvP*	E	aminomethyl-transferring glycine dehydrogenase
Ga0436389_006_139713_140174	50.98	8	23	*rpl9*	J	50S ribosomal protein L9
Ga0436389_006_143205_145160	3.53	2	4	*gidA*	D	tRNA uridine-5-carboxymethylaminomethyl(34) synthesis enzyme
Ga0436389_006_147182_147703	34.10	5	14	-	S	DUF1269 domain-containing protein
Ga0436389_006_149189_149611	25.00	3	7	-	S	hypothetical protein
Ga0436389_006_157336_160278	0.92	1	2	-	V	peptidase domain-containing ABC transporter
Ga0436389_006_169346_172165	14.06	12	19	*valS*	J	valine−tRNA ligase
Ga0436389_006_177279_178604	21.32	7	21	-	G	ScyD/ScyE family protein
Ga0436389_006_179061_179315	36.90	2	15	-	S	hypothetical protein
Ga0436389_006_182501_182869	54.92	8	44	-	K	AbrB family transcriptional regulator
Ga0436389_006_211053_212597	2.33	1	2	-	S	metallophosphoesterase
Ga0436389_006_212614_213261	30.70	7	17	-	C	NAD(P)H-dependent oxidoreductase
Ga0436389_006_214742_215242	17.47	2	6	*moaB*	H	molybdenum cofactor biosynthesis protein B
Ga0436389_006_217480_217971	11.66	2	3	-	FJ	deaminase
Ga0436389_006_2213_3172	40.75	9	40	-	ET	transporter substrate-binding domain-containing protein
Ga0436389_006_222551_223216	9.05	2	4	-	IQ	SDR family NAD(P)-dependent oxidoreductase
Ga0436389_006_224431_225378	38.73	10	33	*phnD*	P	putative selenate ABC transporter substrate-binding protein
Ga0436389_006_22757_24919	1.25	1	2	-	G	glycogen debranching protein GlgX
Ga0436389_006_230698_232128	12.61	5	7	-	S	L,D-transpeptidase family protein
Ga0436389_006_232805_235027	16.35	11	20	*katG*	P	catalase/peroxidase HPI
Ga0436389_006_235033_236019	17.99	5	6	-	H	hypothetical protein
Ga0436389_006_238562_238903	9.73	1	2	-	S	YciI family protein
Ga0436389_006_247274_249397	5.66	3	4	-	E	transglutaminase family protein
Ga0436389_006_249428_251854	12.87	10	19	*glgP*	G	glycogen/starch/alpha-glucan phosphorylase

Table B.20 continued from previous page

Protein Accession	Coverage (%)	UPN	PSM	Gene	COG	Description
Ga0436389_006_251960_252859	8.36	2	5	-	E	transglutaminase family protein
Ga0436389_006_252844_253824	19.02	5	7	-	S	alpha-E domain-containing protein
Ga0436389_006_253818_255257	15.03	6	12	-	S	circularly permuted type 2 ATP-grasp protein
Ga0436389_006_255986_256747	20.16	3	9	-	S	hypothetical protein
Ga0436389_006_256732_258456	13.41	7	12	*ureC*	E	urease subunit alpha
Ga0436389_006_258795_259097	29.00	3	4	*ureA*	E	urease subunit gamma
Ga0436389_006_26936_27742	32.46	8	26	-	IQ	glucose 1-dehydrogenase
Ga0436389_006_270480_271775	68.68	25	282	*urtA*	E	urea ABC transporter substrate-binding protein
Ga0436389_006_274310_275056	43.95	9	28	*urtD*	S	urea ABC transporter ATP-binding protein UrtD
Ga0436389_006_275088_275813	51.45	12	36	*urtE*	E	urea ABC transporter ATP-binding subunit UrtE
Ga0436389_006_276270_277475	40.40	13	50	-	C	acetamidase/formamidase family protein
Ga0436389_006_277642_278151	20.71	3	5	-	S	hypothetical protein
Ga0436389_006_281210_281758	12.64	2	5	-	S	ecotin family protein
Ga0436389_006_282399_284078	10.55	5	14	*pyrG*	F	CTP synthase
Ga0436389_006_284265_284750	29.19	4	18	*dps*	P	DNA starvation/stationary phase protection protein
Ga0436389_006_284840_286663	32.29	17	44	*aspS*	J	aspartate–tRNA ligase
Ga0436389_006_286749_287864	24.53	8	11	*gcvT*	E	glycine cleavage system aminomethyltransferase GcvT
Ga0436389_006_289696_290589	21.21	5	7	*speE*	H	polyamine aminopropyltransferase
Ga0436389_006_294469_294783	64.42	6	28	-	E	Belongs to the P(II) protein family
Ga0436389_006_33675_35189	9.72	5	9	-	E	GMC oxidoreductase
Ga0436389_006_3692_4618	33.12	7	11	-	ET	transporter substrate-binding domain-containing protein
Ga0436389_006_37954_39603	1.82	1	2	-	E	FAD-dependent oxidoreductase
Ga0436389_006_47155_47538	11.81	1	3	*usp*	T	Universal stress protein
Ga0436389_006_47679_48884	56.61	22	109	*pgk*	F	phosphoglycerate kinase
Ga0436389_006_48929_49336	17.78	3	7	-	S	hypothetical protein
Ga0436389_006_51533_52327	17.42	3	11	-	S	hypothetical protein
Ga0436389_006_59218_61458	30.56	16	56	*gspD*	NU	general secretion pathway protein D
Ga0436389_006_62459_63187	29.75	6	9	-	S	PilN domain-containing protein
Ga0436389_006_63184_64002	31.99	5	11	-	NU	hypothetical protein
Ga0436389_006_66146_66901	7.57	2	4	-	M	DUF3747 domain-containing protein
Ga0436389_006_67191_67577	36.72	6	57	*rplL*	J	50S ribosomal protein L7/L12

Table B.20 continued from previous page

Protein Accession	Coverage (%)	UPN	PSM	Gene	COG	Description
Ga0436389_006_67639_68166	48.00	9	40	*rplJ*	J	50S ribosomal protein L10
Ga0436389_006_69080_69787	42.13	12	51	*rpl1*	J	50S ribosomal protein L1
Ga0436389_006_69877_70302	45.39	7	39	*rplK*	J	50S ribosomal protein L11
Ga0436389_006_70462_71172	44.49	9	23	*nusG*	K	transcription termination/antitermination protein NusG
Ga0436389_006_71515_74388	6.06	3	11	*clpB2*	O	AAA family ATPase
Ga0436389_006_74880_76175	25.99	7	28	*eno*	H	phosphopyruvate hydratase
Ga0436389_006_76194_77891	4.96	3	6	-	S	AarF/ABC1/UbiB kinase family protein
Ga0436389_006_79714_80949	10.22	4	8	*argJ*	E	arginine biosynthesis bifunctional protein ArgJ
Ga0436389_006_84420_85865	42.41	18	31	*gatB*	J	Asp-tRNA(Asn)/Glu-tRNA(Gln) amidotransferase subunit GatB
Ga0436389_006_89006_89674	10.36	2	4	-	PT	FecR domain-containing protein
Ga0436389_006_89745_90203	30.92	5	17	*ndk*	F	nucleoside-diphosphate kinase
Ga0436389_006_92519_94498	7.13	5	6	*speA*	H	biosynthetic arginine decarboxylase
Ga0436389_006_98226_98930	17.95	4	9	-	S	SIMPL domain-containing protein
Ga0436389_007_103059_104093	10.76	3	4	-	GM	SDR family NAD(P)-dependent oxidoreductase
Ga0436389_007_106454_107632	4.08	2	2	-	M	glycosyltransferase
Ga0436389_007_11941_12342	36.09	5	16	*rps4*	J	30S ribosomal protein S4
Ga0436389_007_126216_127262	16.67	4	5	*galE*	M	UDP-glucose 4-epimerase GalE
Ga0436389_007_128358_129482	29.14	9	23	*gmd*	H	GDP-mannose 4,6-dehydratase
Ga0436389_007_132187_132612	68.79	9	28	*rpsP*	J	30S ribosomal protein S16
Ga0436389_007_132659_134155	9.64	5	8	*ffh*	U	signal recognition particle protein
Ga0436389_007_13269_14798	9.23	4	6	*murE*	M	UDP-N-acetylmuramyl-tripeptide synthetase
Ga0436389_007_136249_137358	23.58	7	19	*pdhA*	C	pyruvate dehydrogenase E1 component subunit alpha
Ga0436389_007_143700_144581	14.68	4	7	-	S	hypothetical protein
Ga0436389_007_150004_151392	15.80	7	22	*psbC*	P	photosystem II reaction center protein CP43
Ga0436389_007_151376_151486	86.11	1	4	*psbD*	C	Photosystem II DII subunit, Q(A) protein
Ga0436389_007_15768_16574	19.03	4	10	-	M	mechanosensitive ion channel family protein
Ga0436389_007_16624_17823	5.76	1	2	*cefD*	E	aminotransferase class V-fold PLP-dependent enzyme
Ga0436389_007_20391_21530	13.98	4	8	*yidC*	U	membrane protein insertase YidC
Ga0436389_007_21606_22052	10.14	1	3	-	J	PH domain-containing protein
Ga0436389_007_22652_23269	22.93	5	9	-	S	DUF2808 domain-containing protein
Ga0436389_007_23747_24559	12.96	4	6	*sppA*	OU	signal peptide peptidase SppA

Table B.20 continued from previous page

Protein Accession	Coverage (%)	UPN	PSM	Gene	COG	Description
Ga0436389_007_34915_35502	47.18	7	28	*hpf*	J	ribosome-associated translation inhibitor RaiA
Ga0436389_007_36310_38259	7.40	4	8	*fadD*	I	AMP-binding protein
Ga0436389_007_38319_38762	63.27	11	33	-	S	YlqD family protein
Ga0436389_007_38809_40152	26.62	11	31	*pdhC*	C	2-oxo acid dehydrogenase subunit E2
Ga0436389_007_41358_41957	44.22	7	57	*sodB*	C	superoxide dismutase
Ga0436389_007_42027_42698	37.22	8	37	*fkpA*	M	FKBP-type peptidyl-prolyl cis-trans isomerase
Ga0436389_007_42736_43077	40.71	4	6	-	S	hypothetical protein
Ga0436389_007_45787_46485	14.66	3	4	-	S	hypothetical protein
Ga0436389_007_47651_48175	21.26	2	6	-	S	hypothetical protein
Ga0436389_007_49464_52028	29.27	19	57	-	P	arylsulfatase
Ga0436389_007_53434_54396	40.00	9	19	*argF*	E	ornithine carbamoyltransferase
Ga0436389_007_57631_59523	35.71	18	47	*ftsH3*	O	ATP-dependent zinc metalloprotease FtsH
Ga0436389_007_60363_60617	20.24	1	3	-	S	Nif11-like leader peptide family RiPP precursor
Ga0436389_007_6310_6993	11.01	1	4	-	S	DUF2764 family protein
Ga0436389_007_63822_66329	12.22	8	10	-	G	bile acid beta-glucosidase
Ga0436389_007_7050_8879	4.76	2	2	-	C	V-type ATP synthase subunit A
Ga0436389_007_73340_74347	13.13	4	4	*pheS*	J	phenylalanine–tRNA ligase subunit alpha
Ga0436389_007_76955_78070	11.32	4	7	*thiE*	H	thiamine phosphate synthase
Ga0436389_007_85979_87067	22.38	7	23	-	M	polysaccharide export protein
Ga0436389_007_87566_89926	38.80	24	73	*epsB*	DM	polysaccharide biosynthesis tyrosine autokinase
Ga0436389_007_97943_99694	5.32	3	5	-	O	carbamoyltransferase
Ga0436389_008_100726_101124	34.09	3	17	*rplU*	J	50S ribosomal protein L21
Ga0436389_008_101180_101446	29.55	2	10	*rpmA*	J	50S ribosomal protein L27
Ga0436389_008_103159_103935	14.73	3	5	-	K	YebC/PmpR family DNA-binding transcriptional regulator protein
Ga0436389_008_107232_107804	36.84	5	26	*psbV*	C	cytochrome c-550
Ga0436389_008_108347_108646	48.48	3	67	*petF*	C	2Fe-2S iron-sulfur cluster-binding protein
Ga0436389_008_10887_14177	38.69	34	112	*rpoB*	K	DNA-directed RNA polymerase subunit beta
Ga0436389_008_109754_111340	40.34	17	55	*serA*	E	phosphoglycerate dehydrogenase
Ga0436389_008_113114_114532	11.02	4	7	*murD*	D	UDP-N-acetylmuramoyl-L-alanine–D-glutamate ligase
Ga0436389_008_116021_116248	30.67	2	6	*rpoZ*	K	DNA-directed RNA polymerase subunit omega
Ga0436389_008_118044_118871	17.45	4	5	-	S	hypothetical protein

Table B.20 continued from previous page

Protein Accession	Coverage (%)	UPN	PSM	Gene	COG	Description
Ga0436389_008_118912_119130	66.67	3	5	*ftrV*	S	ferredoxin-thioredoxin reductase variable chain
Ga0436389_008_121728_122231	23.95	3	5	-	O	peptide-methionine (R)-S-oxide reductase MsrB
Ga0436389_008_124196_125833	63.85	29	206	*groEL*	O	chaperonin GroEL
Ga0436389_008_125887_126198	54.37	5	57	*groS*	O	co-chaperone GroES
Ga0436389_008_126447_127913	58.20	20	195	*atpD*	C	F0F1 ATP synthase subunit beta
Ga0436389_008_137011_137478	49.03	8	39	*ycf65*	J	30S ribosomal protein PSRP-3
Ga0436389_008_14236_16140	32.33	22	95	*rpoC1*	K	DNA-directed RNA polymerase subunit gamma
Ga0436389_008_144427_145575	16.75	5	12	*ald*	C	alanine dehydrogenase
Ga0436389_008_146464_147237	8.95	2	4	-	S	hypothetical protein
Ga0436389_008_159598_160773	17.39	5	8	-	S	hypothetical protein
Ga0436389_008_16214_20329	40.19	48	145	*rpoC2*	K	DNA-directed RNA polymerase subunit beta’
Ga0436389_008_163251_164024	64.20	13	68	*pspA*	KT	PspA/IM30 family protein
Ga0436389_008_164060_164410	27.59	3	10	-	O	thioredoxin domain-containing protein
Ga0436389_008_167769_168452	6.61	1	2	*lolD*	V	ABC transporter ATP-binding protein
Ga0436389_008_170222_173014	21.51	17	33	*topA*	L	type I DNA topoisomerase
Ga0436389_008_179502_179948	60.81	6	9	-	K	AbrB family transcriptional regulator
Ga0436389_008_186580_187602	3.24	1	2	*ccmA*	V	ABC transporter ATP-binding protein
Ga0436389_008_189786_191453	47.39	22	95	*groL2*	O	chaperonin GroEL
Ga0436389_008_191842_192594	57.60	10	24	*fabG*	IQ	3-oxoacyl-[acyl-carrier-protein] reductase
Ga0436389_008_193810_194652	15.36	3	8	-	M	lipopolysaccharide heptosyltransferase family protein
Ga0436389_008_194738_195439	7.30	2	2	*ispD*	I	2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase
Ga0436389_008_200112_200912	22.18	4	5	-	S	helix-turn-helix domain-containing protein
Ga0436389_008_204171_205112	10.22	3	24	*petA*	C	apocytochrome f
Ga0436389_008_20510_21568	4.55	2	4	*rlmN*	J	23S rRNA (adenine(2503)-C(2))-methyltransferase RlmN
Ga0436389_008_205169_205729	18.82	3	10	*petC*	C	cytochrome b6-f complex iron-sulfur subunit
Ga0436389_008_209555_210064	46.75	6	43	*psaF*	S	Photosystem I reaction center subunit III
Ga0436389_008_212849_213967	12.10	4	8	*nfrC*	G	UDP-N-acetylglucosamine 2-epimerase (non-hydrolyzing)
Ga0436389_008_215246_216337	24.52	7	9	*ychF*	J	redox-regulated ATPase YchF
Ga0436389_008_221859_222092	15.58	1	3	-	S	hypothetical protein
Ga0436389_008_229434_230771	12.36	5	5	*thrA*	E	homoserine dehydrogenase
Ga0436389_008_233101_234186	33.24	10	31	*chlI*	H	magnesium chelatase ATPase subunit I

Table B.20 continued from previous page

Protein Accession	Coverage (%)	UPN	PSM	Gene	COG	Description
Ga0436389_008_234183_235682	20.64	6	13	-	N	DUF3370 domain-containing protein
Ga0436389_008_235682_236746	16.10	4	5	*ampC*	V	serine hydrolase
Ga0436389_008_242405_242728	51.40	5	51	*trxA*	O	thioredoxin
Ga0436389_008_242971_244134	57.36	16	78	*guaB*	F	GuaB3 family IMP dehydrogenase-related protein
Ga0436389_008_246097_248760	21.08	17	24	*gyrA*	L	DNA gyrase subunit A
Ga0436389_008_253221_253568	53.91	4	7	-	S	hypothetical protein
Ga0436389_008_254467_255165	31.90	6	11	-	S	hypothetical protein
Ga0436389_008_3_554	23.37	7	17	-	S	translation initiation factor IF-2
Ga0436389_008_41003_41635	8.10	2	2	-	S	hypothetical protein
Ga0436389_008_46935_48230	19.26	7	11	*purB*	F	adenylosuccinate lyase
Ga0436389_008_49509_49847	70.54	8	58	*glnB*	K	P-II family nitrogen regulator
Ga0436389_008_5382_6095	24.89	5	15	*rpiA*	G	ribose-5-phosphate isomerase RpiA
Ga0436389_008_53909_54109	39.39	3	13	*apcC*	U	phycobilisome linker polypeptide
Ga0436389_008_54156_54644	76.54	8	195	*apcB*	C	allophycocyanin subunit beta
Ga0436389_008_54688_55173	63.98	9	125	*apcA*	C	allophycocyanin subunit alpha apoprotein
Ga0436389_008_55558_58584	50.50	40	156	*apcE*	G	phycobilisome rod-core linker polypeptide
Ga0436389_008_58734_59969	10.22	4	6	-	H	methyltransferase
Ga0436389_008_61457_61672	9.86	1	7	*atpE*	C	ATP synthase F0 subunit C
Ga0436389_008_61773_62255	30.63	5	16	*atpG*	U	F0F1 ATP synthase subunit B’
Ga0436389_008_62252_62779	49.71	8	36	*atpF*	C	F0F1 ATP synthase subunit B
Ga0436389_008_62779_63327	44.51	6	25	*atpD*	C	F0F1 ATP synthase subunit delta
Ga0436389_008_63390_64907	46.34	23	159	*atpA*	C	F0F1 ATP synthase subunit alpha
Ga0436389_008_64921_65871	27.85	9	31	*atpC*	C	F0F1 ATP synthase subunit gamma
Ga0436389_008_71122_71487	15.70	2	3	*petF1*	C	2Fe-2S iron-sulfur cluster binding domain-containing protein
Ga0436389_008_72755_73261	8.93	2	6	-	S	DUF1993 domain-containing protein
Ga0436389_008_73445_74017	78.42	14	124	-	O	peroxiredoxin
Ga0436389_008_78195_79499	13.36	4	7	*tyrS*	J	tyrosine–tRNA ligase
Ga0436389_008_8027_9430	11.78	5	7	*hisD*	E	histidinol dehydrogenase
Ga0436389_008_83561_85066	18.36	7	22	*pepA*	E	leucyl aminopeptidase
Ga0436389_008_87560_88366	32.09	6	12	*lpxA*	I	acyl-ACP–UDP-N-acetylglucosamine O-acyltransferase
Ga0436389_008_88386_88907	43.35	6	11	*fabZ*	I	3-hydroxyacyl-ACP dehydratase FabZ

Table B.20 continued from previous page

Protein Accession	Coverage (%)	UPN	PSM	Gene	COG	Description
Ga0436389_008_88912_89784	20.00	4	5	*lpxC*	M	UDP-3-O-acyl-N-acetylglucosamine deacetylase
Ga0436389_008_89784_92189	30.59	21	55	*IAP75*	M	BamA/TamA family outer membrane protein
Ga0436389_008_904_2343	26.93	12	30	*nusA*	K	transcription termination factor NusA
Ga0436389_008_92233_92991	19.84	5	6	*purC*	F	phosphoribosylaminoimidazolesuccinocarboxamide synthase
Ga0436389_008_92988_93992	9.28	2	4	-	S	hypothetical protein
Ga0436389_008_94299_95657	27.65	11	15	*purD*	F	phosphoribosylamine−glycine ligase
Ga0436389_008_9533_9844	41.75	4	11	*rpsT*	J	30S ribosomal protein S20
Ga0436389_008_97722_99221	17.64	8	15	*kaiC*	T	circadian clock protein KaiC
Ga0436389_008_99359_99718	36.97	5	7	*kaiB*	T	circadian clock protein KaiB
Ga0436389_015_2539_2748	68.12	2	10	*psaE*	U	photosystem I reaction center subunit IV
Ga0436389_015_5500_5874	9.68	2	2	-	M	LysM peptidoglycan-binding domain-containing protein
Ga0436389_016_16760_17248	25.93	4	10	*ribH*	H	6,7-dimethyl-8-ribityllumazine synthase
Ga0436389_016_17245_18153	18.54	5	12	*rfbA*	H	glucose-1-phosphate thymidylyltransferase RfbA
Ga0436389_016_18150_18749	17.09	3	6	-	G	dTDP-4-dehydrorhamnose 3,5-epimerase
Ga0436389_016_19892_21154	5.24	2	3	-	GM	ABC transporter ATP-binding protein
Ga0436389_016_25417_26532	14.02	3	4	-	Q	class I SAM-dependent methyltransferase
Ga0436389_016_9559_11370	9.12	5	8	*lysC*	E	aspartate kinase
Ga0436389_017_1370_2116	18.55	3	3	-	V	hypothetical protein
Ga0436389_017_13730_14719	26.14	6	14	-	K	GntR family transcriptional regulator
Ga0436389_017_15679_16260	35.23	8	22	*infC*	J	translation initiation factor IF-3
Ga0436389_017_17466_19424	22.09	13	17	*gyrB*	L	DNA topoisomerase (ATP-hydrolyzing) subunit B
Ga0436389_017_19761_20240	30.82	3	15	*btuE*	O	glutathione peroxidase
Ga0436389_017_9594_12428	24.47	22	44	*secA*	U	preprotein translocase subunit SecA
Ga0436389_018_1_504	7.74	1	3	-	S	iron uptake porin
Ga0436389_019_10490_11986	28.31	12	16	*lysS*	J	lysine−tRNA ligase
Ga0436389_019_13170_13655	12.42	2	3	-	S	hypothetical protein
Ga0436389_019_13652_13888	39.74	2	7	-	S	hypothetical protein
Ga0436389_019_16179_17207	26.02	7	11	*cysM1*	E	cysteine synthase A
Ga0436389_019_17204_19351	8.81	5	5	-	KLT	serine/threonine protein kinase
Ga0436389_019_2836_4215	31.15	13	41	*ahcY*	H	adenosylhomocysteinase
Ga0436389_019_5093_5488	27.48	4	9	*ssb*	L	single-stranded DNA-binding protein

Table B.20 continued from previous page

Protein Accession	Coverage (%)	UPN	PSM	Gene	COG	Description
Ga0436389_019_5648_6694	44.54	13	50	*mreB*	D	rod shape-determining protein
Ga0436389_020_13321_14094	20.62	4	8	*rfbF*	JM	glucose-1-phosphate cytidylyltransferase
Ga0436389_020_17589_18644	19.94	5	9	-	GM	SDR family NAD(P)-dependent oxidoreductase
Ga0436389_020_23707_25512	5.32	3	4	-	EH	thiamine pyrophosphate-binding protein
Ga0436389_020_25561_26364	26.97	8	15	-	G	aldolase/citrate lyase family protein
Ga0436389_020_27372_28523	8.62	3	3	-	H	methyltransferase domain-containing protein
Ga0436389_020_28999_29982	5.20	2	3	-	S	glycosyltransferase
Ga0436389_020_34055_36073	43.45	20	125	*tktA*	H	transketolase
Ga0436389_020_36149_37396	15.66	5	8	*fabF*	I	beta-ketoacyl-ACP synthase II
Ga0436389_020_37796_38041	85.19	7	61	*psaC*	C	photosystem I iron-sulfur center protein PsaC
Ga0436389_020_38119_40017	12.82	7	12	*glmS*	M	glutamine–fructose-6-phosphate transaminase
Ga0436389_020_41281_42729	21.16	9	14	*manC*	M	mannose-1-phosphate guanylyltransferase/mannose-6-phosphate isomerase
Ga0436389_020_44896_45663	27.84	7	11	-	S	hypothetical protein
Ga0436389_020_45810_46403	8.63	1	4	-	S	hypothetical protein
Ga0436389_020_48116_48628	54.12	7	15	*glgP*	F	glycogen/starch/alpha-glucan phosphorylase
Ga0436389_020_4902_6290	21.65	9	15	*thiC*	H	phosphomethylpyrimidine synthase ThiC
Ga0436389_021_1_360	23.33	2	6	-	G	glycogen/starch/alpha-glucan phosphorylase
Ga0436389_022_100421_100738	10.48	1	2	*tatA*	U	TatA/E family twin arginine-targeting protein translocase
Ga0436389_022_103393_104133	9.35	2	5	*ntcA*	K	global nitrogen regulator NtcA
Ga0436389_022_104201_104953	11.60	2	4	*rph*	J	ribonuclease PH
Ga0436389_022_123591_124127	36.52	5	11	-	S	hypothetical protein
Ga0436389_022_124148_124717	63.49	8	17	-	S	hypothetical protein
Ga0436389_022_132852_134576	13.94	8	16	-	S	patatin-like phospholipase family protein
Ga0436389_022_135558_136343	49.81	13	43	-	MV	N-acetylmuramoyl-L-alanine amidase
Ga0436389_022_136489_136995	41.07	5	20	-	S	hypothetical protein
Ga0436389_022_139886_141661	22.17	12	18	-	M	peptidoglycan-binding protein
Ga0436389_022_1512_2375	15.33	3	5	*galM*	G	galactose mutarotase
Ga0436389_022_157260_160280	3.88	2	2	-	L	SMC family ATPase
Ga0436389_022_16464_17726	12.86	6	10	-	M	hypothetical protein
Ga0436389_022_177048_177530	21.25	4	8	-	S	MSMEG_0572 family nitrogen starvation response protein
Ga0436389_022_187481_188332	9.19	2	4	-	S	hypothetical protein

Table B.20 continued from previous page

Protein Accession	Coverage (%)	UPN	PSM	Gene	COG	Description
Ga0436389_022_205356_205733	17.60	2	5	-	S	hypothetical protein
Ga0436389_022_220013_220420	36.30	4	9	-	S	DUF1348 family protein
Ga0436389_022_223953_226880	11.90	9	17	*ileS*	J	isoleucine–tRNA ligase
Ga0436389_022_22543_23955	33.62	15	32	-	C	nucleotide sugar dehydrogenase
Ga0436389_022_227772_229340	26.82	10	18	-	S	alpha/beta hydrolase
Ga0436389_022_230442_230735	61.86	4	6	*gatC*	J	Asp-tRNA(Asn)/Glu-tRNA(Gln) amidotransferase subunit GatC
Ga0436389_022_235409_235606	66.15	3	12	-	S	hypothetical protein
Ga0436389_022_23956_24897	21.41	6	10	*rfbB*	GM	SDR family oxidoreductase
Ga0436389_022_243328_244350	13.82	4	7	-	S	type I-U CRISPR-associated RAMP protein Csb1/Cas7u
Ga0436389_022_257979_258380	23.31	3	6	-	S	hypothetical protein
Ga0436389_022_258493_259305	74.81	14	90	*psbO*	S	photosystem II manganese-stabilizing polypeptide
Ga0436389_022_259490_260659	24.16	8	15	*sat*	P	sulfate adenylyltransferase
Ga0436389_022_260726_262570	28.18	13	34	*ftsH2*	O	ATP-dependent zinc metalloprotease FtsH3
Ga0436389_022_263199_264287	21.27	5	7	*aroC*	E	chorismate synthase
Ga0436389_022_28222_29262	18.50	5	15	*ycf48*	S	photosynthesis system II assembly factor Ycf48
Ga0436389_022_29262_29672	22.79	2	3	*rub*	C	Rubredoxin
Ga0436389_022_30192_30947	34.66	6	15	*ndhK*	C	NADH-quinone oxidoreductase subunit NuoB
Ga0436389_022_30940_31545	45.27	6	16	*ndhJ*	C	NAD(P)H-quinone oxidoreductase subunit J
Ga0436389_022_38803_39825	10.59	3	3	*ycf22*	Q	MCE family protein
Ga0436389_022_43048_43608	26.88	3	9	-	L	NUDIX hydrolase
Ga0436389_022_48384_49166	23.85	5	17	*fabI*	I	enoyl-ACP reductase FabI
Ga0436389_022_50756_51397	8.92	3	4	*hisB*	E	imidazoleglycerol-phosphate dehydratase HisB
Ga0436389_022_59508_60977	12.07	5	12	-	G	phosphoglucomutase/phosphomannomutase family protein
Ga0436389_022_62594_63559	20.87	4	6	-	S	folate-binding protein
Ga0436389_022_63597_64190	15.23	2	8	*pyrE*	F	orotate phosphoribosyltransferase
Ga0436389_022_6410_7840	33.40	15	28	*icd*	C	NADP-dependent isocitrate dehydrogenase
Ga0436389_022_67292_68284	9.09	3	5	*mviM*	S	Gfo/Idh/MocA family oxidoreductase
Ga0436389_022_77174_78718	30.93	12	25	*purH*	F	Bifunctional purine biosynthesis protein PurH
Ga0436389_022_80105_81313	7.96	3	5	*ispH*	IM	4-hydroxy-3-methylbut-2-enyl diphosphate reductase
Ga0436389_022_81418_82803	7.59	3	11	*amt1*	P	Ammonium Transporter Family
Ga0436389_022_8440_9153	27.85	5	12	*ho1*	C	heme oxygenase (biliverdin-producing)

Table B.20 continued from previous page

Protein Accession	Coverage (%)	UPN	PSM	Gene	COG	Description
Ga0436389_022_86912_88222	44.27	15	44	*glyA*	E	serine hydroxymethyltransferase
Ga0436389_022_91266_92708	13.54	6	12	*leuC*	H	3-isopropylmalate dehydratase large subunit
Ga0436389_022_92705_93337	14.76	3	5	*leuD*	E	3-isopropylmalate dehydratase small subunit
Ga0436389_022_93334_96531	6.76	6	8	-	NU	DUF3769 domain-containing protein
Ga0436389_022_96927_99914	1.01	1	2	*ams1*	G	alpha-mannosidase
Ga0436389_023_3_1748	20.10	8	27	-	Q	hypothetical protein
Ga0436389_023_5570_6466	29.53	7	11	-	S	hypothetical protein
Ga0436389_024_13006_13851	12.81	2	5	*pheA*	E	Prephenate dehydratase
Ga0436389_024_15652_15972	66.98	7	25	*rpsJ*	J	30S ribosomal protein S10
Ga0436389_024_16103_17332	69.93	21	246	*tuf*	J	elongation factor Tu
Ga0436389_024_17375_19450	46.16	27	119	*fusA*	J	elongation factor G
Ga0436389_024_19563_20033	50.00	7	34	*rps7*	J	30S ribosomal protein S7
Ga0436389_024_20096_20470	18.55	3	11	*rpsL*	J	30S ribosomal protein S12
Ga0436389_024_2079_4040	9.49	4	11	-	Q	PFAM Hemolysin-type calcium-binding repeat (2 copies)
Ga0436389_024_23133_27662	14.05	17	34	*glsF*	E	glutamate synthase large subunit
Ga0436389_024_32741_35002	8.63	7	25	*psaA*	C	photosystem I core protein PsaA
Ga0436389_024_35131_37245	12.36	7	20	*psaB*	C	photosystem I core protein PsaB
Ga0436389_024_37846_43668	1.80	3	4	-	S	esterase-like activity of phytase family protein
Ga0436389_024_4925_5278	8.55	1	3	*clpS*	S	ATP-dependent Clp protease adapter ClpS
Ga0436389_024_5314_6540	20.83	7	12	*dapL*	H	LL-diaminopimelate aminotransferase
Ga0436389_024_63385_63681	21.43	3	19	*rpmE*	J	50S ribosomal protein L31
Ga0436389_024_63690_64097	21.48	3	11	*rps9*	J	30S ribosomal protein S9
Ga0436389_024_64094_64546	44.67	5	19	*rplM*	J	50S ribosomal protein L13
Ga0436389_024_65535_65885	37.07	5	12	*rplQ*	J	50S ribosomal protein L17
Ga0436389_024_65932_66870	63.14	14	59	*rpoA*	K	DNA-directed RNA polymerase subunit alpha
Ga0436389_024_66953_67345	33.08	5	8	*rpsK*	J	30S ribosomal protein S11
Ga0436389_024_67517_67819	40.00	6	20	*rpsM*	J	30S ribosomal protein S13
Ga0436389_024_68106_68657	54.10	7	20	*adk*	F	adenylate kinase
Ga0436389_024_70109_70564	23.18	2	11	*rplO*	J	50S ribosomal protein L15
Ga0436389_024_70573_71238	23.98	7	24	*rps5*	J	30S ribosomal protein S5
Ga0436389_024_71253_71618	54.55	6	18	*rplR*	J	50S ribosomal protein L18

Table B.20 continued from previous page

Protein Accession	Coverage (%)	UPN	PSM	Gene	COG	Description
Ga0436389_024_71651_72190	42.46	6	30	*rpl6*	J	50S ribosomal protein L6
Ga0436389_024_72207_72581	46.77	5	16	*rps8*	J	30S ribosomal protein S8
Ga0436389_024_72629_73168	51.96	9	35	*rpl5*	J	50S ribosomal protein L5
Ga0436389_024_73291_73647	36.44	5	24	*rplX*	J	50S ribosomal protein L24
Ga0436389_024_73649_74014	69.42	8	38	*rplN*	J	50S ribosomal protein L14
Ga0436389_024_74025_74276	45.78	4	10	*rpsQ*	J	30S ribosomal protein S17
Ga0436389_024_74294_74506	37.14	2	10	*rpmC*	J	50S ribosomal protein L29
Ga0436389_024_74509_74919	36.76	3	18	*rplP*	J	50S ribosomal protein L16
Ga0436389_024_74984_75712	48.35	14	48	*rps3*	J	30S ribosomal protein S3
Ga0436389_024_75732_76091	42.02	3	10	*rplV*	J	50S ribosomal protein L22
Ga0436389_024_76096_76371	45.05	5	15	*rpsS*	J	30S ribosomal protein S19
Ga0436389_024_76407_77270	36.24	11	40	*rpl2*	J	50S ribosomal protein L2
Ga0436389_024_77286_77564	66.30	6	19	*rplW*	J	50S ribosomal protein L23
Ga0436389_024_77581_78216	54.03	12	46	*rpl4*	J	50S ribosomal protein L4
Ga0436389_024_78228_78842	40.69	9	30	*rpl3*	J	50S ribosomal protein L3
Ga0436389_024_79288_79749	16.99	2	5	*ndhN*	C	NAD(P)H-quinone oxidoreductase subunit N
Ga0436389_024_9738_11801	32.17	15	33	*rne*	J	Rne/Rng family ribonuclease
Ga0436389_025_102808_104676	4.66	3	5	-	L	DUF2075 domain-containing protein
Ga0436389_025_110047_111048	19.82	5	8	*psr*	K	LCP family protein
Ga0436389_025_13183_13512	39.45	4	7	-	T	STAS domain-containing protein
Ga0436389_025_19236_20126	19.59	4	9	-	S	glycosyltransferase family 8 protein
Ga0436389_025_20_223	32.84	2	9	*psbA*	C	photosystem II q(b) protein
Ga0436389_025_32929_34473	22.96	10	18	*glgA*	F	glycogen synthase GlgA
Ga0436389_025_37123_38106	11.01	2	3	-	O	META domain-containing protein
Ga0436389_025_41503_49626	0.74	2	3	-	S	hypothetical protein
Ga0436389_025_49774_50304	6.25	1	2	-	J	peptide deformylase
Ga0436389_025_64356_65666	9.17	3	6	*cupA*	C	CO2 hydration protein
Ga0436389_025_67199_68260	55.81	17	50	*natF*	ET	amino acid ABC transporter substrate-binding protein
Ga0436389_025_70546_71295	31.33	7	13	*bgtA*	E	polar amino acid ABC transporter ATPase
Ga0436389_025_71305_72870	2.11	1	2	-	G	DUF1957 domain-containing protein
Ga0436389_025_72984_74609	16.64	9	13	*leuA*	E	2-isopropylmalate synthase

Table B.20 continued from previous page

Protein Accession	Coverage (%)	UPN	PSM	Gene	COG	Description
Ga0436389_025_80852_81733	5.46	1	2	*folD*	F	bifunctional methylenetetrahydrofolate
Ga0436389_025_82113_83015	19.00	5	8	*crtE*	H	polyprenyl synthetase family protein
Ga0436389_025_87990_89288	8.10	3	6	*opcA*	G	glucose-6-phosphate dehydrogenase assembly protein OpcA
Ga0436389_025_89356_90879	32.54	13	26	*zwf*	H	glucose-6-phosphate dehydrogenase
Ga0436389_025_91047_92216	59.64	18	101	*petH*	C	FAD-binding oxidoreductase
Ga0436389_025_95217_95657	57.53	9	37	-	S	cAMP phosphodiesterase
Ga0436389_025_95665_98376	8.97	8	12	*pepN*	E	aminopeptidase N
Ga0436389_025_98373_99626	6.47	3	5	-	S	family 10 glycosylhydrolase
Ga0436389_026_100307_100966	46.58	9	48	*tsf*	J	translation elongation factor Ts
Ga0436389_026_105759_107528	17.15	9	15	*sir*	C	sulfite reductase, ferredoxin dependent
Ga0436389_026_107642_109807	15.53	7	9	*glyS*	J	glycine–tRNA ligase subunit beta
Ga0436389_026_111252_112634	48.70	17	39	*chlP*	C	geranylgeranyl reductase
Ga0436389_026_113845_115653	22.43	10	23	*typA*	T	translational GTPase TypA
Ga0436389_026_116096_116584	29.01	4	9	-	M	LPS-assembly protein LptD
Ga0436389_026_116581_117309	13.64	3	5	-	S	LPS export ABC transporter ATP-binding protein
Ga0436389_026_119591_120274	27.31	5	21	*rpe*	G	ribulose-phosphate 3-epimerase
Ga0436389_026_11964_14078	14.63	8	16	*prlC*	E	M3 family metallopeptidase
Ga0436389_026_120479_121483	45.51	13	85	*glpX*	G	class II fructose-bisphosphatase
Ga0436389_026_122976_124271	33.18	13	35	*glgC*	H	glucose-1-phosphate adenylyltransferase
Ga0436389_026_124384_125802	25.21	9	29	*gnd*	H	NADP-dependent phosphogluconate dehydrogenase
Ga0436389_026_125799_126530	25.51	5	10	*pgl*	G	6-phosphogluconolactonase
Ga0436389_026_132651_134327	33.87	15	37	*ilvD*	H	dihydroxy-acid dehydratase
Ga0436389_026_135419_135991	35.79	6	10	-	S	pentapeptide repeat-containing protein
Ga0436389_026_139219_139575	21.19	1	2	-	S	hypothetical protein
Ga0436389_026_142026_143171	9.45	3	4	*pcxA*	U	proton extrusion protein PcxA
Ga0436389_026_143194_143910	17.65	4	9	-	S	TPM domain-containing protein
Ga0436389_026_144615_145748	8.22	2	3	-	M	LPS export ABC transporter periplasmic protein LptC
Ga0436389_026_145745_147295	10.47	3	4	*metG*	J	methionine–tRNA ligase
Ga0436389_026_149931_150245	33.65	3	6	-	S	hypothetical protein
Ga0436389_026_152298_152519	61.64	5	14	*rpsR*	J	30S ribosomal protein S18
Ga0436389_026_152609_152803	17.19	1	6	*rpmG*	J	50S ribosomal protein L33

Table B.20 continued from previous page

Protein Accession	Coverage (%)	UPN	PSM	Gene	COG	Description
Ga0436389_026_152916_155372	16.38	11	14	*pheT*	J	phenylalanine–tRNA ligase subunit beta
Ga0436389_026_156862_157356	42.07	6	16	*apcD*	C	allophycocyanin
Ga0436389_026_170827_171744	35.41	9	40	*ilvE*	E	branched-chain amino acid transaminase
Ga0436389_026_175590_176126	76.97	9	70	-	S	hypothetical protein
Ga0436389_026_179858_180349	23.31	3	12	-	S	flavin reductase
Ga0436389_026_180930_181877	53.65	14	45	-	C	hypothetical protein
Ga0436389_026_183265_183795	42.61	7	20	-	S	DUF4330 domain-containing protein
Ga0436389_026_183872_184561	28.82	4	5	-	S	DUF1995 family protein
Ga0436389_026_185765_186919	10.16	3	4	*dapF*	E	diaminopimelate epimerase
Ga0436389_026_187240_189861	10.08	7	9	*leuS*	J	leucine–tRNA ligase
Ga0436389_026_190725_191837	17.30	6	7	*adhC*	C	S-(Hydroxymethyl)glutathione dehydrogenase class III alcohol dehydrogenase
Ga0436389_026_191857_193458	18.95	7	10	*pgi*	G	glucose-6-phosphate isomerase
Ga0436389_026_195460_197208	6.19	3	5	-	S	M61 family metallopeptidase
Ga0436389_026_198203_199267	22.88	5	7	*argC*	E	N-acetyl-gamma-glutamyl-phosphate reductase
Ga0436389_026_199347_201056	3.16	2	2	*ribBA*	H	bifunctional 3,4-dihydroxy-2-butanone-4-phosphate synthase
Ga0436389_026_214652_215092	21.23	2	8	*rot*	O	peptidylprolyl isomerase
Ga0436389_026_21467_21841	38.71	3	18	-	S	VOC family protein
Ga0436389_026_215142_216113	25.39	6	12	*mtnP*	F	S-methyl-5'-thioadenosine phosphorylase
Ga0436389_026_217023_217424	25.56	2	4	-	S	DUF3110 domain-containing protein
Ga0436389_026_218461_220602	11.64	6	13	*dnaK1*	O	molecular chaperone DnaK
Ga0436389_026_21889_23265	14.85	6	11	*glmU*	M	UDP-N-acetylglucosamine diphosphorylase GlmU
Ga0436389_026_222690_223499	9.29	2	7	*pstB*	P	phosphate ABC transporter ATP-binding protein
Ga0436389_026_224151_225065	4.61	1	2	*suhB*	G	inositol monophosphatase
Ga0436389_026_225107_226294	20.51	7	8	*hisZ*	J	ATP phosphoribosyltransferase regulatory subunit
Ga0436389_026_226637_228556	33.96	17	33	*htpG*	O	molecular chaperone HtpG
Ga0436389_026_232726_232962	42.31	5	13	*rpmB*	J	50S ribosomal protein L28
Ga0436389_026_236904_238823	14.24	7	10	*dxs*	H	1-deoxy-D-xylulose-5-phosphate synthase
Ga0436389_026_238926_240527	4.69	3	7	*ilvA*	E	threonine ammonia-lyase, biosynthetic
Ga0436389_026_24196_25542	13.62	4	7	*aroA*	E	3-phosphoshikimate 1-carboxyvinyltransferase
Ga0436389_026_242104_242433	22.94	1	2	-	S	nucleoside triphosphate pyrophosphohydrolase family protein
Ga0436389_026_242511_244277	36.22	16	48	*pykF*	G	pyruvate kinase

Table B.20 continued from previous page

Protein Accession	Coverage (%)	UPN	PSM	Gene	COG	Description
Ga0436389_026_246109_248088	30.05	15	53	*ftsH1*	O	ATP-dependent zinc metalloprotease FtsH
Ga0436389_026_248155_248745	47.45	8	19	*clpP1*	O	ATP-dependent Clp endopeptidase proteolytic subunit ClpP
Ga0436389_026_248841_249536	36.80	6	16	*thf1*	S	photosystem II biogenesis protein Psp29
Ga0436389_026_252408_253691	12.65	4	5	*purK*	F	5-(carboxyamino)imidazole ribonucleotide synthase
Ga0436389_026_256610_257749	8.18	3	4	*aroB*	E	3-dehydroquinate synthase
Ga0436389_026_257888_258982	19.51	6	13	-	S	DUF4912 domain-containing protein
Ga0436389_026_265589_266674	22.16	7	17	-	G	fructose-bisphosphate aldolase class I
Ga0436389_026_284549_285352	8.61	2	4	-	T	GGDEF domain-containing protein
Ga0436389_026_285422_287809	23.27	17	36	*nrdJ*	F	ribonucleoside-triphosphate reductase, adenosylcobalamin-dependent
Ga0436389_026_28686_30254	7.09	3	3	*ubiD*	H	UbiD family decarboxylase
Ga0436389_026_288612_290264	8.73	4	5	*prfC*	J	peptide chain release factor 3
Ga0436389_026_294297_294674	23.20	2	11	-	S	hypothetical protein
Ga0436389_026_297303_298481	27.30	7	16	*aspC*	E	pyridoxal phosphate-dependent aminotransferase
Ga0436389_026_299273_300526	22.30	7	15	*ispG*	I	(E)-4-hydroxy-3-methylbut-2-enyl-diphosphate synthase
Ga0436389_026_3_773	12.45	1	6	*psbA*	C	photosystem II q(b) protein
Ga0436389_026_300596_301927	46.28	20	50	*prc*	M	S41 family peptidase
Ga0436389_026_302419_303681	14.29	4	7	-	E	4-Hydroxyphenylpyruvate dioxygenase
Ga0436389_026_30327_31082	11.55	3	7	*comB*	H	2-phosphosulfolactate phosphatase family protein
Ga0436389_026_303734_307321	2.51	4	7	*mfd*	L	transcription-repair coupling factor
Ga0436389_026_308331_309326	12.08	4	9	*gap1*	C	type I glyceraldehyde-3-phosphate dehydrogenase
Ga0436389_026_3098_3670	20.00	4	13	-	S	hypothetical protein
Ga0436389_026_31144_31962	38.24	7	17	-	S	carbon-nitrogen hydrolase family protein
Ga0436389_026_311703_312554	4.59	1	2	-	S	hypothetical protein
Ga0436389_026_316959_318353	12.07	5	8	*pmbA*	S	TldD/PmbA family protein
Ga0436389_026_318356_319816	17.70	8	17	*tldD*	S	TldD/PmbA family protein
Ga0436389_026_31969_33069	12.84	3	5	-	M	N-acetylmuramoyl-L-alanine amidase
Ga0436389_026_319856_320947	7.16	2	5	*acsF*	H	magnesium-protoporphyrin IX monomethyl ester
Ga0436389_026_322784_325090	19.79	13	26	*zam*	K	RNB domain-containing ribonuclease
Ga0436389_026_331902_332084	11.67	1	4	*rpmF*	J	50S ribosomal protein L32
Ga0436389_026_332191_334164	13.55	6	14	*ftsH4*	O	ATP-dependent zinc metalloprotease FtsH
Ga0436389_026_334102_335097	3.32	1	3	-	Q	cupin domain-containing protein

Table B.20 continued from previous page

Protein Accession	Coverage (%)	UPN	PSM	Gene	COG	Description
Ga0436389_026_335230_335505	47.25	3	7	-	H	MoaD/ThiS family protein
Ga0436389_026_335561_336841	30.99	11	22	-	E	threonine synthase
Ga0436389_026_339934_341007	36.69	14	74	*cbbA*	G	fructose-bisphosphate aldolase class II
Ga0436389_026_341101_342195	3.85	2	4	*mviM*	S	Gfo/Idh/MocA family oxidoreductase
Ga0436389_026_344854_345756	47.67	11	45	*prkB*	C	phosphoribulokinase
Ga0436389_026_345838_346920	31.11	8	19	*leuB*	CE	3-isopropylmalate dehydrogenase
Ga0436389_026_346950_347978	9.94	2	5	*lpxD*	I	UDP-3-O-(3-hydroxymyristoyl)glucosamine N-acyltransferase
Ga0436389_026_348863_351214	3.07	2	3	-	O	S8 family peptidase
Ga0436389_026_357189_357524	19.82	2	3	-	S	hypothetical protein
Ga0436389_026_357649_359166	5.54	3	6	-	S	DUF1254 domain-containing protein
Ga0436389_026_36540_38513	13.09	8	19	*acsA*	H	acetate–CoA ligase
Ga0436389_026_3667_4704	8.70	3	4	*trpS*	J	tryptophan–tRNA ligase
Ga0436389_026_373588_374766	11.73	4	7	-	S	F420-0:Gamma-glutamyl ligase
Ga0436389_026_374852_376231	8.06	3	5	-	T	PhoH family protein
Ga0436389_026_376999_378378	3.70	2	3	*fumC*	C	class II fumarate hydratase
Ga0436389_026_378378_378947	11.64	2	4	-	CO	thylakoid membrane photosystem I accumulation factor
Ga0436389_026_381291_382052	21.74	4	7	*hisA*	E	(5-phosphoribosylamino)methylideneamino]imidazole-4-carboxamide isomerase
Ga0436389_026_383156_384073	36.07	9	17	-	GM	NAD-dependent epimerase/dehydratase family protein
Ga0436389_026_387543_388010	4.52	1	3	-	S	CBS domain-containing protein
Ga0436389_026_39316_39780	9.09	1	4	-	O	peroxiredoxin
Ga0436389_026_400205_401221	3.55	1	2	-	P	ABC transporter substrate-binding protein
Ga0436389_026_406021_406554	33.33	4	8	-	P	ferritin
Ga0436389_026_411109_412116	32.24	8	16	-	P	phosphate ABC transporter substrate-binding protein PstS
Ga0436389_026_416356_417222	16.67	5	7	*dapB*	E	4-hydroxy-tetrahydrodipicolinate reductase
Ga0436389_026_419705_420244	20.11	3	6	*apt*	F	adenine phosphoribosyltransferase
Ga0436389_026_42286_43002	28.15	4	7	*msrA*	O	peptide-methionine (S)-S-oxide reductase MsrA
Ga0436389_026_425305_425562	62.35	2	8	-	S	DUF4912 domain-containing protein
Ga0436389_026_45716_46756	8.09	2	7	*trpD*	F	anthranilate phosphoribosyltransferase
Ga0436389_026_48822_50645	5.44	4	6	*rlmB*	J	23S rRNA (guanosine(2251)-2'-O)-methyltransferase RlmB
Ga0436389_026_50709_50960	27.71	2	7	-	S	DUF1816 domain-containing protein
Ga0436389_026_51005_52477	27.55	9	23	*gatA*	J	Asp-tRNA(Asn)/Glu-tRNA(Gln) amidotransferase subunit GatA

Table B.20 continued from previous page

Protein Accession	Coverage (%)	UPN	PSM	Gene	COG	Description
Ga0436389_026_57453_57722	17.98	2	5	*rpsO*	J	30S ribosomal protein S15
Ga0436389_026_6330_8177	9.92	6	8	*thrS*	J	threonine–tRNA ligase
Ga0436389_026_66694_67221	29.71	4	14	-	S	pentapeptide repeat-containing protein
Ga0436389_026_73288_74271	16.82	5	12	*pdhB*	C	alpha-ketoacid dehydrogenase subunit beta
Ga0436389_026_79101_79475	17.74	2	14	*psb28*	S	photosystem II reaction center protein Psb28
Ga0436389_026_82900_83418	51.74	9	26	*apcF*	C	allophycocyanin subunit beta
Ga0436389_026_83466_84662	20.10	7	21	*spt*	E	alanine–glyoxylate aminotransferase family protein
Ga0436389_026_9242_10216	7.41	3	3	*thrB*	F	homoserine kinase
Ga0436389_026_99579_100268	53.71	12	26	*rps2*	J	30S ribosomal protein S2
Ga0436389_027_11557_11769	28.57	3	15	-	S	DUF2862 domain-containing protein
Ga0436389_027_14982_16322	5.16	2	3	*sun*	J	16S rRNA (cytosine(967)-C(5))-methyltransferase
Ga0436389_027_19149_19844	21.21	4	8	-	K	response regulator transcription factor
Ga0436389_027_22226_22645	16.55	1	15	-	NU	prepilin-type N-terminal cleavage/methylation domain-containing protein
Ga0436389_027_31953_33761	7.48	5	7	*lepA*	M	translation elongation factor 4
Ga0436389_027_33903_34205	38.00	3	7	*nifU*	O	NifU-like domain
Ga0436389_027_39156_40481	28.80	10	17	*serS*	J	serine–tRNA ligase
Ga0436389_027_43269_43889	9.71	2	3	-	S	hypothetical protein
Ga0436389_027_43889_46591	12.33	8	18	-	S	hypothetical protein
Ga0436389_027_4812_6317	25.75	11	22	-	P	sulfatase-like hydrolase/transferase
Ga0436389_027_49480_51609	48.66	29	129	*pnp*	J	polyribonucleotide nucleotidyltransferase
Ga0436389_027_54441_55385	7.96	2	4	-	L	hypothetical protein
Ga0436389_027_8639_9517	20.21	4	6	*glyQ*	J	glycine–tRNA ligase subunit alpha
Ga0436389_029_10758_11861	5.99	2	5	*coaE*	H	oxygenase MpaB family protein
Ga0436389_029_16072_17667	18.46	9	17	*nirA*	C	ferredoxin–nitrite reductase
Ga0436389_029_19276_20589	45.54	20	89	*nrtA*	P	ABC transporter substrate-binding protein
Ga0436389_029_21478_23514	7.23	4	5	*nrtC*	P	ABC transporter substrate-binding protein
Ga0436389_029_23553_24386	14.80	3	5	*nrtD*	P	ATP-binding cassette domain-containing protein
Ga0436389_029_24959_27157	5.05	2	2	*narB*	C	nitrate reductase
Ga0436389_029_28040_30472	8.89	7	10	*xfp*	H	phosphoketolase family protein
Ga0436389_029_40651_42792	16.97	11	16	*ppk*	H	polyphosphate kinase 1
Ga0436389_029_46607_47680	9.24	4	12	*aroG*	E	3-deoxy-7-phosphoheptulonate synthase

Table B.20 continued from previous page

Protein Accession	Coverage (%)	UPN	PSM	Gene	COG	Description
Ga0436389_029_47829_50471	25.34	19	51	*acnB*	C	bifunctional aconitate hydratase 2/2-methylisocitrate dehydratase
Ga0436389_029_60466_61059	43.15	7	21	-	S	photosystem II protein PsbQ
Ga0436389_029_62289_64193	54.73	35	146	*dnaK2*	O	molecular chaperone DnaK
Ga0436389_029_65869_66231	50.83	5	27	*rps6*	J	30S ribosomal protein S6
Ga0436389_029_66995_68209	48.51	16	38	*argG*	F	argininosuccinate synthase
Ga0436389_029_68542_68814	27.78	3	3	-	S	DUF3134 domain-containing protein
Ga0436389_029_70514_71719	13.47	5	9	*purT*	F	formate-dependent phosphoribosylglycinamide formyltransferase
Ga0436389_029_72789_73283	45.12	5	14	-	P	TerB family tellurite resistance protein
Ga0436389_029_79625_81022	7.53	3	4	*gad*	E	glutamate decarboxylase
Ga0436389_029_89163_89789	41.83	7	14	-	S	alpha/beta hydrolase
Ga0436389_029_89839_90942	28.88	7	16	*thrC*	E	threonine synthase
Ga0436389_029_91292_92479	31.39	9	27	*dnaN*	L	DNA polymerase III subunit beta
Ga0436389_029_93403_95772	13.43	9	13	*purL*	F	phosphoribosylformylglycinamidine synthase subunit PurL
Ga0436389_029_958_1443	11.80	2	4	-	H	molybdenum cofactor biosynthesis protein MoaE
Ga0436389_029_95811_97331	9.29	5	8	*purF*	F	amidophosphoribosyltransferase
Ga0436389_029_98382_99137	3.59	1	2	-	S	phytochelatin synthase family protein
Ga0436389_030_12641_13579	8.65	2	6	-	M	LysM peptidoglycan-binding domain-containing protein
Ga0436389_030_17649_18806	33.51	10	26	-	M	efflux RND transporter periplasmic adaptor subunit
Ga0436389_030_22588_23661	20.73	6	15	-	M	efflux RND transporter periplasmic adaptor subunit
Ga0436389_030_26116_26676	52.69	8	63	-	S	alpha/beta hydrolase
Ga0436389_030_27016_27483	21.94	4	10	*bcp*	O	thioredoxin-dependent thiol peroxidase
Ga0436389_030_31004_33337	13.00	9	47	*hflX*	O	GTPase HflX
Ga0436389_030_34763_35467	17.09	5	7	*trkA*	P	TrkA family potassium uptake protein
Ga0436389_030_37202_37510	9.80	1	3	-	S	hypothetical protein
Ga0436389_030_37601_37846	66.67	4	13	-	S	hypothetical protein
Ga0436389_030_38087_39844	9.74	5	6	-	S	ABC-F family ATP-binding cassette domain-containing protein
Ga0436389_030_40722_41879	26.23	9	14	-	M	trypsin-like peptidase domain-containing protein
Ga0436389_030_42484_42870	46.09	5	13	-	S	hypothetical protein
Ga0436389_030_4887_6533	22.26	8	21	*pgm*	G	alpha-D-glucose phosphate-specific phosphoglucomutase
Ga0436389_030_9106_9573	32.90	4	7	-	S	hypothetical protein
Ga0436389_031_11820_12821	28.53	8	22	*hemB*	H	porphobilinogen synthase

Table B.20 continued from previous page

Protein Accession	Coverage (%)	UPN	PSM	Gene	COG	Description
Ga0436389_031_14039_14512	22.93	4	7	-	CO	hypothetical protein
Ga0436389_031_24793_25269	12.66	2	6	-	S	hypothetical protein
Ga0436389_031_25370_28978	4.58	5	8	*smc*	D	chromosome segregation protein SMC
Ga0436389_031_289_642	35.90	4	6	-	S	hypothetical protein
Ga0436389_031_29071_30279	25.37	7	12	-	S	PRC-barrel domain-containing protein
Ga0436389_031_30298_30714	21.74	2	4	*msrB*	C	peptide-methionine (R)-S-oxide reductase MsrB
Ga0436389_031_32334_33683	26.50	8	18	*accC*	I	acetyl-CoA carboxylase biotin carboxylase subunit
Ga0436389_031_34313_35287	9.57	2	3	*ycf66*	S	hypothetical protein
Ga0436389_031_36594_38615	7.58	4	8	*mcyH*	S	ABC transporter ATP-binding protein/permease
Ga0436389_031_38686_39030	38.60	2	2	*hit*	FG	histidine triad nucleotide-binding protein
Ga0436389_031_40957_41127	14.29	1	3	*rpsU*	J	30S ribosomal protein S21
Ga0436389_031_41189_42022	18.41	5	17	-	C	DUF3747 domain-containing protein
Ga0436389_031_46366_48477	16.79	10	17	-	IQ	bifunctional aldolase/short-chain dehydrogenase
Ga0436389_031_52307_53371	26.27	8	15	-	M	sulfurtransferase
Ga0436389_031_54371_55471	18.31	6	10	*mtnA*	J	S-methyl-5-thioribose-1-phosphate isomerase
Ga0436389_031_6303_6521	51.39	6	20	-	S	hypothetical protein
Ga0436389_031_645_1070	17.02	2	6	-	S	hypothetical protein
Ga0436389_031_6588_7580	6.06	2	4	*obg*	S	GTPase ObgE
Ga0436389_031_71970_73448	12.80	6	8	-	S	DUF1254 domain-containing protein
Ga0436389_031_74991_76298	8.51	3	3	*sufS*	H	SufS family cysteine desulfurase
Ga0436389_031_76295_77575	20.19	8	12	*sufD*	O	Fe-S cluster assembly protein SufD
Ga0436389_031_7645_8883	6.55	3	4	*malK*	P	ATP-binding cassette domain-containing protein
Ga0436389_031_77577_78380	22.85	5	10	*sufC*	O	Fe-S cluster assembly ATPase SufC
Ga0436389_031_78454_79896	13.75	8	17	*sufB*	O	Fe-S cluster assembly protein SufB
Ga0436389_031_79914_80279	38.02	3	9	*ftrC*	C	ferredoxin:thioredoxin reductase
Ga0436389_032_10755_11810	17.38	5	9	*fabH*	I	beta-ketoacyl-ACP synthase 3
Ga0436389_032_11838_12710	15.17	3	5	*fabD*	I	ACP S-malonyltransferase
Ga0436389_032_1569_2366	22.64	4	11	*rpaA*	K	response regulator transcription factor
Ga0436389_032_16414_16842	46.48	6	15	*rbp3*	S	RNA-binding protein
Ga0436389_032_17800_18822	17.65	4	6	-	P	aromatic ring-hydroxylating dioxygenase subunit alpha
Ga0436389_032_18824_19840	2.37	1	3	-	P	aromatic ring-hydroxylating dioxygenase subunit alpha

Table B.20 continued from previous page

Protein Accession	Coverage (%)	UPN	PSM	Gene	COG	Description
Ga0436389_032_26504_27268	4.33	1	2	-	D	MinD/ParA family protein
Ga0436389_032_28769_30382	5.96	2	2	-	MU	TolC family protein
Ga0436389_032_293_556	50.57	6	21	-	S	hypothetical protein
Ga0436389_032_31305_32726	9.94	5	6	*pds*	C	15-cis-phytoene desaturase
Ga0436389_032_32891_33238	42.61	4	14	*ndhM*	C	NAD(P)H-quinone oxidoreductase subunit M
Ga0436389_032_33968_34969	15.92	5	7	*rbcR*	K	LysR family transcriptional regulator
Ga0436389_032_35938_37950	3.28	3	5	*ndhF1*	CP	NAD(P)H-quinone oxidoreductase subunit 5
Ga0436389_032_41323_42405	24.44	9	16	*cugP*	H	sugar phosphate nucleotidyltransferase
Ga0436389_032_43693_44277	9.28	2	3	-	S	CYTH domain-containing protein
Ga0436389_032_46280_46882	39.00	7	23	*ndhI*	C	NAD(P)H-quinone oxidoreductase subunit I
Ga0436389_032_48164_49306	13.95	5	7	*gltA*	C	citrate synthase
Ga0436389_032_49450_49980	15.91	2	3	*sixA*	T	phophohistidine phosphatase SixA
Ga0436389_032_53894_55153	32.46	10	20	*trpB*	E	tryptophan synthase subunit beta
Ga0436389_032_56652_57287	36.02	4	10	*cysC*	F	adenylyl-sulfate kinase
Ga0436389_032_62224_63021	31.32	6	11	*ycf29*	K	response regulator
Ga0436389_032_6540_7061	17.34	2	3	*ycf3*	NU	photosystem I assembly protein Ycf3
Ga0436389_032_67162_68346	23.86	10	28	*ndhH*	C	NAD(P)H-quinone oxidoreductase subunit H
Ga0436389_032_71407_71673	15.91	1	4	-	S	TM2 domain-containing protein
Ga0436389_032_76480_77469	19.15	5	8	*gshB*	H	glutathione synthase
Ga0436389_032_79814_80881	21.69	7	12	*prfB*	J	peptide chain release factor 2
Ga0436389_032_82184_82780	13.13	2	4	*trpG*	EH	aminodeoxychorismate/anthranilate synthase component II
Ga0436389_032_82878_83618	46.34	8	19	-	S	MBL fold metallo-hydrolase
Ga0436389_032_8650_9396	55.65	10	27	*rpaB*	K	response regulator transcription factor
Ga0436389_032_9377_10729	8.89	4	5	*plsX*	I	phosphate acyltransferase PlsX
Ga0436389_033_1518_2282	69.29	17	130	*cpcG1*	G	phycobilisome rod-core linker polypeptide
Ga0436389_033_768_1400	17.62	2	2	*ycf58*	H	phycobiliprotein lyase
Ga0436389_034_100615_101928	30.89	12	18	*purA*	F	adenylosuccinate synthase
Ga0436389_034_102442_104241	15.86	8	12	*proS*	J	proline–tRNA ligase
Ga0436389_034_105439_106125	20.18	3	4	-	H	class I SAM-dependent methyltransferase
Ga0436389_034_106258_106620	49.17	5	33	-	C	(2Fe-2S)-binding protein
Ga0436389_034_106626_106994	24.59	3	5	*arsC*	P	Spx/MgsR family RNA polymerase-binding regulatory protein

Table B.20 continued from previous page

Protein Accession	Coverage (%)	UPN	PSM	Gene	COG	Description
Ga0436389_034_110575_111900	17.23	5	12	*gpmB*	G	histidine phosphatase family protein
Ga0436389_034_115884_117023	51.45	17	46	*tal*	H	transaldolase
Ga0436389_034_117107_117496	20.93	2	5	-	T	Crp/Fnr family transcriptional regulator
Ga0436389_034_118666_119205	49.16	8	27	*frr*	J	ribosome recycling factor
Ga0436389_034_119235_120056	28.57	5	15	*pyrH*	F	uridylate kinase
Ga0436389_034_120282_120986	7.69	2	2	*cobO*	H	cob(I)yrinic acid a,c-diamide adenosyltransferase
Ga0436389_034_122385_123560	15.35	5	10	*hemH*	H	ferrochelatase
Ga0436389_034_124251_126146	4.75	2	3	*ilvB*	H	biosynthetic-type acetolactate synthase large subunit
Ga0436389_034_126808_129276	4.62	3	5	*ppsA*	H	phosphoenolpyruvate synthase
Ga0436389_034_132099_132971	11.03	2	5	-	S	Tab2/Atab2 family RNA-binding protein
Ga0436389_034_132980_134293	26.09	9	26	*rps1b*	J	S1 RNA-binding domain-containing protein
Ga0436389_034_13470_16556	2.33	3	3	-	S	DUF499 domain-containing protein
Ga0436389_034_136032_137081	17.19	6	9	-	S	long-chain acyl-[acyl-carrier-protein] reductase
Ga0436389_034_137101_138090	17.02	4	5	*accA*	I	acetyl-CoA carboxylase carboxyltransferase subunit alpha
Ga0436389_034_138836_139579	23.48	6	9	*folE*	H	GTP cyclohydrolase I
Ga0436389_034_139586_140119	15.82	3	6	-	S	hypothetical protein
Ga0436389_034_144476_145480	27.25	8	21	*por*	IQ	protochlorophyllide reductase
Ga0436389_034_146075_146959	21.77	4	4	-	T	hypothetical protein
Ga0436389_034_147623_148528	20.93	6	8	*chlL*	P	ferredoxin:protochlorophyllide reductase (ATP-dependent) iron-sulfur ATP-binding protein
Ga0436389_034_156962_157243	17.20	2	4	-	E	transcriptional regulator
Ga0436389_034_157455_157766	57.28	4	106	*ccmK2*	CQ	BMC domain-containing protein
Ga0436389_034_157818_159230	41.70	21	146	*cbbL*	H	form I ribulose bisphosphate carboxylase large subunit
Ga0436389_034_159332_159673	53.10	5	24	*rbcS*	C	ribulose bisphosphate carboxylase small subunit
Ga0436389_034_159746_162112	61.42	33	135	*csoS2*	S	carboxysome shell protein
Ga0436389_034_162120_163658	12.50	6	11	*csoS3*	S	carboxysome shell carbonic anhydrase
Ga0436389_034_163661_163975	36.54	3	7	-	CQ	carboxysome peptide A
Ga0436389_034_19011_19817	4.48	1	2	-	S	DUF4393 domain-containing protein
Ga0436389_034_51799_54357	9.39	5	8	-	T	autotransporter outer membrane beta-barrel domain-containing protein
Ga0436389_034_55284_56603	19.36	7	10	*proA*	E	glutamate-5-semialdehyde dehydrogenase
Ga0436389_034_64984_65754	41.80	7	21	-	S	DUF3887 domain-containing protein
Ga0436389_034_65806_68100	23.43	15	29	*glgB*	G	1,4-alpha-glucan branching protein GlgB

Table B.20 continued from previous page

Protein Accession	Coverage (%)	UPN	PSM	Gene	COG	Description
Ga0436389_034_68197_69255	16.76	5	8	*hemE*	H	uroporphyrinogen decarboxylase
Ga0436389_034_695_1744	8.31	2	3	-	S	hypothetical protein
Ga0436389_034_70361_70702	24.78	2	28	*petJ*	C	c-type cytochrome
Ga0436389_034_71575_74205	8.56	6	20	*clpB*	O	ATP-dependent chaperone ClpB
Ga0436389_034_79200_79616	21.01	3	5	-	O	OsmC family protein
Ga0436389_034_81921_82508	47.69	8	20	-	C	inorganic diphosphatase
Ga0436389_034_82555_83199	30.37	6	17	-	S	L,D-transpeptidase
Ga0436389_034_85477_86430	31.55	7	20	*hemC*	H	hydroxymethylbilane synthase
Ga0436389_034_86538_86939	10.53	1	4	-	S	DUF1824 family protein
Ga0436389_034_90680_91993	21.51	9	20	*sigA*	K	RNA polymerase sigma factor RpoD
Ga0436389_034_95986_96882	18.12	4	7	*argB*	F	acetylglutamate kinase
Ga0436389_034_97526_97777	24.10	2	5	-	S	hypothetical protein
Ga0436389_034_97778_98308	44.89	5	11	*ycf41*	L	single-stranded DNA-binding protein
Ga0436389_035_10584_11219	12.80	3	9	-	S	DUF3318 domain-containing protein
Ga0436389_035_11212_11598	22.66	2	3	*rsfS*	J	ribosome silencing factor
Ga0436389_035_12160_13143	17.43	4	10	*ansA*	E	asparaginase
Ga0436389_035_16419_17534	9.43	3	7	*cbiX*	S	sirohydrochlorin chelatase
Ga0436389_035_17632_18282	27.31	4	6	-	S	chromosome partitioning protein ParB
Ga0436389_035_21873_22319	41.22	5	12	-	T	Universal stress protein
Ga0436389_035_25523_25975	52.67	4	9	-	S	hypothetical protein
Ga0436389_035_26926_27327	42.86	4	9	-	S	hypothetical protein
Ga0436389_035_30354_30695	59.29	4	54	*rbpD*	S	RNA-binding protein
Ga0436389_035_32118_33074	45.91	12	73	-	S	orange carotenoid-binding protein
Ga0436389_035_33266_34033	32.55	7	38	*cpcG1*	G	phycobilisome rod-core linker polypeptide
Ga0436389_035_3387_4148	32.02	7	31	*tpiA*	G	triose-phosphate isomerase
Ga0436389_035_4417_6156	3.28	2	4	*mdlB*	V	ABC transporter transmembrane domain-containing protein
Ga0436389_035_45633_46937	4.61	2	5	-	S	hypothetical protein
Ga0436389_035_47426_54061	15.37	22	46	-	MQU	tandem-95 repeat protein
Ga0436389_035_6196_6426	64.47	5	17	-	S	DUF6447 family protein
Ga0436389_035_6439_7116	43.11	7	16	-	S	DUF3386 domain-containing protein
Ga0436389_035_7139_10462	11.65	12	20	*carB*	F	carbamoyl-phosphate synthase large subunit

Table B.20 continued from previous page

Protein Accession	Coverage (%)	UPN	PSM	Gene	COG	Description
Ga0436389_037_14309_15898	7.94	3	3	*lytB*	D	SpoIID/LytB domain-containing protein
Ga0436389_037_15970_16170	31.82	2	3	*rpmI*	J	50S ribosomal protein L35
Ga0436389_037_16239_16586	14.78	2	4	*rplT*	J	50S ribosomal protein L20
Ga0436389_037_16624_17187	13.37	2	4	*ycf37*	S	tetratricopeptide repeat protein
Ga0436389_037_17206_17775	37.04	5	9	-	S	hypothetical protein
Ga0436389_037_18956_20152	39.20	12	37	*sqdB*	GM	NAD-dependent epimerase/dehydratase family protein
Ga0436389_037_1977_2906	19.09	5	8	*dapA*	E	4-hydroxy-tetrahydrodipicolinate synthase
Ga0436389_037_25191_26810	8.72	5	8	*guaA*	F	glutamine-hydrolyzing GMP synthase
Ga0436389_037_28018_29223	43.64	14	45	*dhsS*	E	alanine–glyoxylate aminotransferase family protein
Ga0436389_037_2903_3931	11.11	3	6	*asd*	E	aspartate-semialdehyde dehydrogenase
Ga0436389_037_29285_29752	37.42	4	16	-	S	hypothetical protein
Ga0436389_037_3_1904	22.08	14	23	*rnj*	J	RNase J family beta-CASP ribonuclease
Ga0436389_037_31804_32256	26.00	3	8	-	S	penicillin amidase
Ga0436389_037_32240_32641	9.77	1	2	-	S	hypothetical protein
Ga0436389_037_37116_37610	33.54	4	14	*accB*	I	acetyl-CoA carboxylase biotin carboxyl carrier protein
Ga0436389_037_37616_38176	22.58	4	30	*efp*	J	elongation factor P
Ga0436389_037_38288_39391	32.97	10	26	-	M	peptidylprolyl isomerase
Ga0436389_037_40445_41416	70.90	18	126	*gap2*	C	type I glyceraldehyde-3-phosphate dehydrogenase
Ga0436389_037_4250_5704	47.11	15	54	*tig*	D	trigger factor
Ga0436389_037_44026_44367	44.25	4	12	-	S	YbaB/EbfC family nucleoid-associated protein
Ga0436389_037_45677_46801	7.75	3	6	*dnaJ*	O	molecular chaperone DnaJ
Ga0436389_037_46801_47529	38.02	8	22	*grpE*	O	nucleotide exchange factor GrpE
Ga0436389_037_47667_49124	10.31	3	3	*pilB,*	NU	type II/IV secretion system protein
Ga0436389_037_49130_50209	61.00	17	40	*pilT*	NU	type IV pilus twitching motility protein PilT
Ga0436389_037_53784_54287	50.30	4	15	-	S	RNA-binding protein
Ga0436389_037_54429_55829	4.51	2	4	*argH*	E	argininosuccinate lyase
Ga0436389_037_5741_6487	26.61	4	13	*clpP2*	O	ATP-dependent Clp protease, proteolytic subunit ClpP
Ga0436389_037_58993_59619	11.06	2	4	*nusB*	K	transcription antitermination factor NusB
Ga0436389_037_60402_61052	20.37	4	16	-	S	HpsJ family protein
Ga0436389_037_62088_63002	28.62	8	28	-	S	Tetratricopeptide repeat
Ga0436389_037_63067_65535	9.12	7	9	-	L	DNA topoisomerase 4 subunit A

Table B.20 continued from previous page

Protein Accession	Coverage (%)	UPN	PSM	Gene	COG	Description
Ga0436389_037_6596_7969	13.79	6	13	*clpX*	O	ATP-dependent protease ATP-binding subunit ClpX
Ga0436389_038_14437_15219	8.85	2	4	*gloB*	S	hydroxyacylglutathione hydrolase
Ga0436389_038_18685_19071	48.44	4	8	*tdcF*	J	RidA family protein
Ga0436389_038_21532_23262	3.65	2	2	*me*	C	NAD-dependent malic enzyme
Ga0436389_038_24_332	33.33	1	73	*ccmK2*	CQ	BMC domain-containing protein
Ga0436389_038_338_805	11.61	1	2	-	P	bacterioferritin
Ga0436389_038_35516_39118	1.92	2	2	*nifJ*	C	pyruvate:ferredoxin (flavodoxin) oxidoreductase
Ga0436389_038_42686_43543	2.81	1	2	-	M	outer membrane beta-barrel protein
Ga0436389_038_47586_49181	6.78	3	5	*hoxF*	C	SLBB domain-containing protein
Ga0436389_038_5609_6577	7.45	2	3	*cbbX*	O	CbbX protein
Ga0436389_038_60913_61758	14.59	3	6	*trpA*	E	tryptophan synthase subunit alpha
Ga0436389_038_63285_63647	42.50	5	44	-	K	AbrB family transcriptional regulator
Ga0436389_038_8106_9233	46.93	13	33	-	O	trypsin-like peptidase domain-containing protein
Ga0436389_040_6902_7993	53.72	19	109	*idiA*	C	extracellular solute-binding protein
Ga0436389_041_15810_16178	42.62	5	31	-	K	AbrB family transcriptional regulator
Ga0436389_041_18408_18683	41.76	6	33	-	L	HU family DNA-binding protein
Ga0436389_041_2_2542	30.61	20	47	*infB*	J	translation initiation factor IF-2
Ga0436389_041_26392_27222	19.57	4	9	*proC*	E	pyrroline-5-carboxylate reductase
Ga0436389_041_29844_31208	19.16	8	12	*der*	S	ribosome biogenesis GTPase Der
Ga0436389_041_39208_39861	30.73	5	11	-	M	LysM peptidoglycan-binding domain-containing protein
Ga0436389_041_4021_4233	44.29	3	8	-	S	ribbon-helix-helix protein, CopG family
Ga0436389_042_304_510	66.18	1	18	-	M	iron uptake porin
Ga0436389_044_15264_16370	16.85	5	12	*ddl*	F	D-alanine–D-alanine ligase
Ga0436389_044_17931_19019	38.67	10	40	*ftsZ*	D	cell division protein FtsZ
Ga0436389_044_22522_23193	26.91	7	24	*clpP*	O	ATP-dependent Clp protease proteolytic subunit
Ga0436389_044_23310_23921	19.70	4	9	*clpP3*	O	ATP-dependent Clp protease proteolytic subunit
Ga0436389_044_23975_24970	57.40	15	63	*ilvC*	H	ketol-acid reductoisomerase
Ga0436389_044_3292_3540	65.85	4	9	-	S	hypothetical protein
Ga0436389_044_3813_5120	8.51	3	5	*murA*	M	UDP-N-acetylglucosamine 1-carboxyvinyltransferase
Ga0436389_044_40_924	22.11	5	6	*trpC*	E	indole-3-glycerol phosphate synthase TrpC
Ga0436389_044_5223_6404	24.17	7	11	*argD*	E	aspartate aminotransferase family protein

Table B.20 continued from previous page

Protein Accession	Coverage (%)	UPN	PSM	Gene	COG	Description
Ga0436389_044_7743_8291	37.91	4	5	-	S	pentapeptide repeat-containing protein
Ga0436389_044_926_2398	35.10	15	47	*lpdA*	C	dihydrolipoyl dehydrogenase
Ga0436389_045_2298_4007	38.31	18	51	*tolC*	MU	TolC family protein
Ga0436389_046_16533_17252	34.73	7	12	*gst1*	O	Glutathione S-transferase
Ga0436389_046_22209_22643	13.89	2	6	-	S	SRPBCC family protein
Ga0436389_046_28833_30602	31.41	11	102	-	S	carbohydrate porin
Ga0436389_046_30897_31103	66.18	1	14	-	M	porin
Ga0436389_046_3359_3736	60.80	6	52	*psbU*	U	photosystem II complex extrinsic protein PsbU
Ga0436389_046_5566_6492	13.31	4	11	-	CO	vitamin K epoxide reductase family protein
Ga0436389_046_6649_8010	5.74	3	5	*rimO*	H	30S ribosomal protein S12 methylthiotransferase RimO
Ga0436389_046_9521_11131	24.81	11	32	-	C	B12-binding domain-containing radical SAM protein
Ga0436389_047_10386_11198	23.70	5	12	-	O	META domain-containing protein
Ga0436389_047_105796_109314	0.77	1	2	-	L	helicase-related protein
Ga0436389_047_112270_112707	43.45	4	6	-	S	hypothetical protein
Ga0436389_047_117992_119026	34.88	8	19	-	P	phosphate ABC transporter substrate-binding protein PstS
Ga0436389_047_122388_123812	18.57	8	19	*rimK*	HJO	30S ribosomal protein S6–L-glutamate ligase
Ga0436389_047_129806_130741	41.16	11	47	*coaE*	H	phycobilisome rod-core linker polypeptide
Ga0436389_047_135731_136105	25.00	4	12	-	S	CAAD domain-containing protein
Ga0436389_047_136639_137808	39.07	11	28	*moeB*	HP	molybdopterin-synthase adenylyltransferase MoeB
Ga0436389_047_143128_146160	6.34	6	11	*ppc*	H	phosphoenolpyruvate carboxylase
Ga0436389_047_148951_149382	72.03	9	102	*psaD*	S	photosystem I reaction center subunit II
Ga0436389_047_152153_153259	20.92	5	8	*mrp*	F	Mrp/NBP35 family ATP-binding protein
Ga0436389_047_153453_154478	19.06	5	11	*hemF*	H	oxygen-dependent coproporphyrinogen oxidase
Ga0436389_047_155929_156594	34.84	6	11	*rnd*	J	ribonuclease D
Ga0436389_047_162129_163163	7.56	2	3	*purM*	F	phosphoribosylformylglycinamidine cyclo-ligase
Ga0436389_047_166282_166944	21.36	4	7	*cpcF*	C	HEAT repeat domain-containing protein
Ga0436389_047_1764_2963	26.32	7	10	*recA*	L	recombinase RecA
Ga0436389_047_24055_25611	5.60	2	3	-	CH	FAD-dependent monooxygenase
Ga0436389_047_3494_3745	19.28	2	4	-	S	DUF2839 domain-containing protein
Ga0436389_047_37485_38417	13.87	4	6	-	S	DUF262 domain-containing protein
Ga0436389_047_40433_41332	4.68	1	2	-	S	hypothetical protein

Table B.20 continued from previous page

Protein Accession	Coverage (%)	UPN	PSM	Gene	COG	Description
Ga0436389_047_44242_46971	5.72	5	9	-	L	hypothetical protein
Ga0436389_047_55922_56941	14.75	4	5	-	C	NAD(P)-dependent alcohol dehydrogenase
Ga0436389_047_66740_67207	63.87	7	23	-	S	hypothetical protein
Ga0436389_047_98548_99237	16.16	4	8	-	T	TerB family tellurite resistance protein
Ga0436389_047_99576_102875	4.64	5	5	-	S	DUF499 domain-containing protein
Ga0436389_048_51_539	67.90	5	229	*cpcA*	C	phycocyanin subunit alpha
Ga0436389_048_600_749	67.35	1	49	*cpcB*	C	phycocyanin beta subunit
Ga0436389_049_407_1258	56.89	17	95	*cpeC*	H	phycobilisome rod-core linker polypeptide
Ga0436389_050_11625_12095	15.38	2	5	-	C	2Fe-2S iron-sulfur cluster binding domain
Ga0436389_050_12238_13098	53.15	15	42	-	T	Universal stress protein
Ga0436389_050_18118_18513	24.43	3	6	*minE*	D	cell division topological specificity factor MinE
Ga0436389_050_18519_19343	33.94	8	21	*minD*	D	septum site-determining protein MinD
Ga0436389_050_19443_20441	24.10	6	10	-	I	cardiolipin synthase
Ga0436389_050_24975_25649	16.07	3	10	*petB*	C	cytochrome b6
Ga0436389_050_25707_26189	5.00	1	4	*petD*	C	cytochrome b6-f complex subunit IV
Ga0436389_050_490_747	81.18	6	48	-	S	CpcD/allophycocyanin linker domain-containing protein
Ga0436389_050_5813_7072	33.65	12	38	*metK*	H	methionine adenosyltransferase
Ga0436389_050_791_1636	40.21	13	83	*cpeC*	H	phycobilisome rod-core linker polypeptide
Ga0436389_050_7931_9007	37.99	12	65	*rps1a*	J	30S ribosomal protein S1
Ga0436389_050_9135_9548	17.52	2	2	*nrdR*	K	transcriptional regulator NrdR
Ga0436389_050_9873_11432	19.65	9	28	*psbB*	P	photosystem II chlorophyll-binding protein CP47
Ga0436389_051_14205_14576	27.64	3	7	-	K	AbrB family transcriptional regulator
Ga0436389_051_19732_19947	9.86	1	2	-	S	DUF3721 domain-containing protein
Ga0436389_051_24078_24575	26.06	3	5	-	S	DUF411 domain-containing protein
Ga0436389_051_45516_45962	61.49	7	22	*cynS*	H	cyanase
Ga0436389_051_70097_71521	14.77	5	9	*lysA*	E	diaminopimelate decarboxylase
Ga0436389_051_72247_74799	37.88	26	78	*clpC*	O	ATP-dependent Clp protease ATP-binding subunit
Ga0436389_053_3_782	12.31	1	8	*psbA*	C	photosystem II q(b) protein
Ga0436389_070_2_241	43.75	2	5	*glgP*	F	glycogen phosphorylase
Ga0436389_074_3_368	84.43	6	170	*cpcB*	C	phycocyanin subunit beta
Ga0436389_079_2_553	36.41	5	12	*glgP*	F	glycogen/starch/alpha-glucan phosphorylase

Table B.20 continued from previous page

Protein Accession	Coverage (%)	UPN	PSM	Gene	COG	Description
Ga0436389_083_3_254	26.19	1	6	*psbD*	C	photosystem II D2 protein (photosystem q(a) protein)
Ga0436389_084_1_297	43.43	1	55	-	M	porin
Ga0436389_096_1_843	23.13	4	15	*glgP*	F	glycogen/starch/alpha-glucan family phosphorylase

Table B.21: **Summary of intensities for quantified proteins**. Intensity values are normalised based on a pooled sample. ANOVA testing was carried out with associated p values FDR-adjusted (Q value) to determine differential expression between conditions. NoN: N-starvation; N: Asparagine; R: Arginine; P: Proline; E: Glutamate.)

Protein Accession	**NoN1**	**NoN2**	**NoN3**	**NO_3^-1**	**NO_3^-2**	**NO_3^-3**	**N1**	**N2**	**N3**	**R1**	**R2**	**R3**	**P1**	**P2**	**P3**	**E1**	**E2**	**E3**	**Q Value**
Ga0436389_001_101384_101671	25.7	108.2	91.3	28.1	171.2	152.1	74.4	84.9	156.9	39.4	39.9	137.4	138	124.6	115.9	79.8	127.5	134.3	0.765985296
Ga0436389_001_105813_106616	12.4	30.7	24.1	9	26	18.3	57.1	61.3	31.4	38.7	5.9	26.7	149.7	109.2	23.9	38.3	34.3	32.5	0.125280741
Ga0436389_001_107869_108948	2483.8	2749.9	2539.1	210	305.8	312.8	1093.3	869.6	906.9	422.7	317.5	770.7	1483.6	1293.9	707	853.7	575	771.4	1.84716E-05
Ga0436389_001_1594_2124	77.8	330.2	301	56.8	330.5	303.3	416.3	367.5	408.5	351.7	207.4	311.5	602.1	415.4	464.2	428.1	388	352.9	0.103883816
Ga0436389_001_17070_17834	2160.2	1832	1474.2	170.2	295.7	323	636.9	743.6	372.3	758.9	573.6	348.9	793.1	528.4	516.1	1103.1	445.5	371.1	0.001754392
Ga0436389_001_20756_21772	46.4	301.9	286.7	32.8	141.5	149.9	42.6	38.9	129.9	50.9	44.3	125.4	30.4	58.6	143.6	45.4	242.1	207.8	0.444105459
Ga0436389_001_265_831	193.3	100.4	100.7	294.5	129.9	112.1	202.3	228.4	155.9	123.4	245.2	143.8	166.8	179.6	126.8	101.3	134.5	173.1	0.826836516
Ga0436389_001_34324_36078	286.4	648.2	661.6	225.3	649.4	631	284.2	219.1	469	130.2	179.6	634.9	129.3	191.6	394.5	136.2	424	465.3	0.609161724
Ga0436389_001_36075_37373	205.6	460.7	447.1	133	341.3	323.4	169.8	148.9	413.4	91	85.1	379.2	174.4	208.1	348.3	113.4	310.8	380	0.794431521
Ga0436389_001_38108_38986	95.1	210	204.3	67.8	114.2	107.9	84.9	62.5	142.2	68.6	57	128.3	115.8	104.3	156.5	84.3	122.1	132.3	0.320822566
Ga0436389_001_39028_39897	412.7	610.8	573.9	228.7	355.3	345.1	358	285.8	388.2	186.8	216.7	456.3	365.8	384.3	485.3	301.5	359.9	319.3	0.111738351
Ga0436389_001_42666_43829	517.7	696.1	635.7	439.5	599.1	574.3	539.9	526.7	674.3	425.8	375.6	661.1	505.6	606.6	466.1	504	770.5	794.3	0.503465265
Ga0436389_001_43859_44758	76.7	87.4	63.3	71.8	44.8	44.5	147.7	107	118.6	132.2	90.6	66.6	221.2	232.7	87.9	158.2	83.3	103.8	0.092739695
Ga0436389_001_48625_50067	135	104.6	92.9	85.7	84.9	81.4	242.4	244.7	279.6	113.8	122.8	121	258.8	258	492.9	111	209.7	145.3	0.010905627
Ga0436389_001_50405_50593	60.4	64.7	49.3	29.6	92.5	74.6	39.7	53.4	50.4	6.6	54	33.2	44.4	46.4	29.7	34.1	35.1	33	0.312958974
Ga0436389_001_50769_51245	1073.5	1191.7	1215.8	1433.3	1871.8	1783.8	860.7	970.8	1033.4	528.9	898.5	1044.8	821.3	712.4	967.2	562.1	1188.3	911	0.010905627
Ga0436389_001_51713_52573	23.4	190.9	202.1	2.8	189.6	185	153	268.3	885.6	255.2	143.6	697.9	211.1	128.7	521.7	503.6	701.4	1208.9	0.181857903
Ga0436389_001_53396_54502	65.6	280.8	290	67.8	364.9	327.3	99.4	107.7	212.4	43.8	64.5	219	126.9	143.1	157.1	35.8	163.5	155.8	0.606188877
Ga0436389_001_54542_55054	1657.9	1446.6	1382.2	1881	1345.5	1311	912.3	1068.7	764.5	761.5	900.1	861.5	957.4	931.4	644	777.1	915.8	711.4	0.00392427
Ga0436389_001_55630_56544	746.3	1488.3	1395.4	543.9	1345.7	1382.5	469.3	530.1	902.2	632.1	660.3	1060.8	541.5	402.3	718.4	573.4	950.3	928.9	0.263945519
Ga0436389_001_57116_57607	650.5	443.2	424.1	578.7	443.8	453.8	729.4	769.3	302.5	904.5	855.2	463.8	1066.9	963.7	315.4	968.1	329.6	302.4	0.788072216
Ga0436389_001_6633_6902	72.4	83.5	88.3	71.6	83.4	96.4	111	101.1	191.4	169.8	123.4	104.6	73	61.5	169.3	238.6	153.2	127.9	0.247532176
Ga0436389_001_77097_77483	156.4	248.9	238.8	124.1	291.6	318.4	342.7	360.7	362.9	153.3	213.8	351.9	777	788.3	475.3	174.1	270.2	222.9	0.004820086
Ga0436389_001_834_1565	592.6	703.6	720.7	576.7	833.9	747.9	775.3	629.6	901.6	511.9	561.8	853.3	524.9	533.1	706.9	499.9	787.1	644	0.748339619
Ga0436389_001_85235_86230	57	230	232.8	118	182.2	191.5	47.9	59.6	253.8	50.5	39.8	193.7	39	66.7	272.7	79.2	229.7	210.8	0.913343434
Ga0436389_001_8994_10151	163.8	297.4	288.4	118.7	329.6	308.6	179.3	202.3	835.6	139.9	139	664.9	231.3	220.6	488.7	224.1	583.3	773.8	0.805585182
Ga0436389_001_97188_98384	104.6	129.3	134.4	125.4	201.1	189	159.8	179.6	139.5	100.7	106.5	116.1	166.4	159.3	131.3	63.3	120.9	100.8	0.045470018

Table B.21 continued from previous page

Protein Accession	NoN1	NoN2	NoN3	NO_3^-1	NO_3^-2	NO_3^-3	N1	N2	N3	R1	R2	R3	P1	P2	P3	E1	E2	E3	Q Value
Ga0436389_001_98381_99709	626.4	745.6	740.1	802.7	815.4	780.5	694.6	847.5	751	630.1	624.9	627.1	536.9	710.4	457	779.9	678.9	664.8	0.061989425
Ga0436389_002_8000_9232	313.6	599.3	588.4	186	464.4	458.5	377	313.7	557.9	312.4	265.4	395.5	293.3	261.3	732	307.5	457.9	321.2	0.857619793
Ga0436389_003_13722_14702	188.2	558.7	532.4	147	455.7	451.4	339.9	349.5	521.3	351.1	253.8	524.1	249.1	345.6	527.8	241.2	860.3	745.8	0.724896365
Ga0436389_003_14740_15174	43.4	141.9	152.8	43.9	166.1	146.3	115.1	79	212	97.1	67.7	138.6	122.6	93.9	173.2	148.3	321.6	272.1	0.237283117
Ga0436389_003_15567_16925	633	1156.4	1106	266.5	339.6	308.8	401.3	416.3	634.3	367.3	327.3	473.4	464.9	438.9	538.7	372.3	424.1	556.5	0.010479808
Ga0436389_003_22524_23453	225.3	198.6	179.4	21.9	12.7	15	140.4	125.3	174.7	103.9	54.9	114.1	204.5	161	40	88.1	54.2	145.2	0.025134825
Ga0436389_003_25618_26139	24.6	66.2	70.8	42.3	87.2	97.8	35	49.8	160	45.3	44.1	164.6	55.7	49.9	91.7	80	106	144.2	0.833821868
Ga0436389_003_33834_34154	5693.6	6062.6	5706.7	2004.9	3139.5	3388.2	2277.7	1921.6	1346.5	1060.8	1121.8	1229.5	3732.2	3725.6	1573	768	1180.8	1057.3	0.00025251
Ga0436389_003_35798_36316	7769.4	7343.9	6883.7	13711.2	12557.4	13164.1	8015.2	8840.3	8886	8246	11014.3	11363.2	4914.5	4953.7	5331.8	5645.2	8183.6	7454.3	0.000167728
Ga0436389_003_38378_38656	299.3	397	361.1	192.1	333.9	349.1	115.1	98.6	299	90.9	190.6	330	60.8	81.3	276.3	60.5	324.6	258.2	0.37176433
Ga0436389_003_40168_40587	222.9	259.7	249.1	94.7	246.5	205.9	167.7	178.9	85.7	254.1	128.8	64.7	276.5	185.1	132.2	283.2	65.8	64.1	0.710964196
Ga0436389_003_41071_43239	9710.7	8037.3	7669.8	7051.6	7012.9	6133.9	4073	3275.3	2961.8	3124.5	2050.7	2075.6	5669.3	4976.8	3308.2	3463.7	2956.8	3214.6	0.000160037
Ga0436389_003_45449_46516	194.9	100.4	96	116.4	65.4	54.9	223.5	192.5	33.3	203.6	148.8	38.6	240.1	291	30.9	202.4	25.8	22.2	0.799856853
Ga0436389_003_56891_57901	366.3	333.3	285.6	81.6	145.2	114.8	116.7	115.3	375.1	87.9	91.3	222.1	84	90.7	623.8	124.1	352.1	337.3	0.642408513
Ga0436389_003_63117_63968	324.6	105.5	108.3	119.1	60.2	68.4	180.8	185.6	146.7	138.1	98	71.4	274.2	191.2	36.9	144.2	84.8	108.8	0.626719426
Ga0436389_003_83842_84417	79.8	191	176.9	27.9	54	49.8	33.2	38.3	71.1	66.9	55	47.2	31.9	28.8	29.9	86.6	64.2	58.9	0.015030853
Ga0436389_004_29599_30195	53.3	40.8	36.9	67.7	51.6	39	102.6	98	155.5	256.5	249.9	144.1	168.6	107.6	47.6	468.2	81.5	137.1	0.246986337
Ga0436389_004_36659_37756	179.2	332.3	322.3	220.3	317.3	289.3	391.2	437.7	515.8	373.6	309.4	348.8	342.3	285.3	846.1	431.2	559.7	496.2	0.361692803
Ga0436389_004_4451_6448	241	300.1	290.3	36.6	98.8	87.3	119.6	111.8	185.4	153.8	83.1	135.4	187.1	170.9	159.1	150.7	235.3	272.7	0.004731823
Ga0436389_004_46165_47994	28.2	20.4	21.4	24.1	28.7	22.7	157.7	164.4	109.2	165.4	106.8	79.3	257.1	130.4	172.6	240.5	138.1	205	0.004139985
Ga0436389_004_47991_48395	3035.4	3151.3	2931.8	2585.2	4919.8	4635	1618.2	1645.9	1536.1	1161.5	2242.7	2340.9	872.7	1119.6	1466.3	917	1960	1257.1	0.006370734
Ga0436389_004_49277_51235	29	165.3	184.2	66.9	236.2	256.4	84	72.2	367.7	75.9	61.9	285.4	148.3	136.1	740.3	50.9	344.1	175.2	0.833821868
Ga0436389_004_62053_62373	79.5	200.7	196.6	62.4	205.6	204.2	43.3	42.7	77	27	35.7	89.9	29.3	26.1	52.4	36.5	62.7	30.6	0.053146209
Ga0436389_004_64037_64612	1316.1	1119.5	1065.9	1171.3	1257	1199.7	676.3	710.9	851.5	589.5	822	870.2	546	646.6	640	441.8	867.1	856.8	0.004731823
Ga0436389_004_64917_65240	304.6	621.7	507.1	164.5	429.5	460.8	184.4	227.4	377.5	123.1	134.6	357.6	625.4	362.1	303.4	193.6	240.9	328.1	0.313074519
Ga0436389_004_65878_66549	259.7	130.4	110.8	263.8	101.5	109.8	342.4	323.6	167.9	290.2	208.6	77	419.3	409.5	438.6	284.2	149.8	59.9	0.068567799
Ga0436389_004_71443_71565	166.5	249.7	250.6	409.1	525.2	403.6	196.1	304.3	334.7	158.1	162.7	305.6	219.7	245.1	207.6	94.2	162.7	151.5	0.006236576
Ga0436389_004_73565_74782	102.6	66.2	59.3	31.1	32.2	22.2	58.8	99.5	41.2	64.5	58.1	32.9	53.3	68.2	19.7	84.8	29.7	22.2	0.431711463
Ga0436389_004_7569_8897	2322.5	3390.6	3103.7	398.6	451.8	442.9	944.1	766.7	755.5	607	423.6	593.7	1676.7	1410.8	876.4	680.7	1031.7	1137.4	0.000137732
Ga0436389_005_16182_17213	41.3	46.7	36.7	34	25.5	24.6	41.7	59.7	21.7	31.8	35.9	29.1	107.4	81.3	31.6	31.7	24.3	7.5	0.141283578
Ga0436389_005_17805_19529	55.4	36.4	34.2	36.2	31.9	17.5	38.2	31.7	12.4	31.2	71.5	42.8	30.2	55.6	18.7	8.2	12.7	6.6	0.156578233

Table B.21 continued from previous page

Protein Accession	NoN1	NoN2	NoN3	NO$_3^-$1	NO$_3^-$2	NO$_3^-$3	N1	N2	N3	R1	R2	R3	P1	P2	P3	E1	E2	E3	Q Value
Ga0436389_005_2951_3700	484.3	549.2	536.8	182.1	160.2	158.8	435.1	447.9	493.6	534.7	870.3	873.6	468.9	519.2	418.1	394.8	553	582.4	0.00283184
Ga0436389_005_618_1670	28.9	62.1	69.1	12	153.2	142.5	103.9	113.5	199.7	87.3	134.3	162.5	151.1	208.5	360.3	101.7	204.1	159.3	0.160899603
Ga0436389_006_102407_103084	331.7	605.9	638.8	270.5	424.8	521.9	360.1	311.3	471.8	342	340.9	456.3	458.7	341.1	795.7	459.1	735.2	682.9	0.444105459
Ga0436389_006_109737_112418	397.2	457.4	471.9	329.5	364.2	321.5	442	517	602.5	469.6	412.1	441.1	659.8	462.3	578.8	678.5	645.5	674.1	0.00283184
Ga0436389_006_116837_119929	149.8	130.8	154.2	42.3	74.8	88.5	319.4	256.9	264.1	103.3	155.2	237.1	301.9	495.6	260	112.8	290.8	428	0.058987037
Ga0436389_006_125743_127530	114	122.1	112.8	116.1	106.9	115.8	1147.5	974.6	476.5	774.8	514.2	289.2	532.4	337.7	170.8	889.2	216.2	489.3	0.051981084
Ga0436389_006_127530_129344	93.1	230	221.2	73.3	169.7	177.5	41.7	39.7	157.8	25.2	63.4	183.1	27.8	36.1	201.4	15.2	162.2	146.3	0.714627317
Ga0436389_006_132039_134975	184.4	259.3	243.7	230.1	215.5	186.5	121.6	177.9	193.2	118	153.9	246.6	105.1	103.2	143.6	130.8	155.2	141	0.08595632
Ga0436389_006_139713_140174	301.4	923.2	886.3	454	1564.6	1380.4	836.3	826.5	1882.3	1281.4	801.3	1438.8	374.4	203.2	884.9	681.6	1346.6	1913.1	0.473658566
Ga0436389_006_143205_145160	48.3	87.1	85	38.4	108.7	107.7	505	189.6	267.4	121.4	43.1	179.7	507.5	409.7	120.9	214.1	188	197.2	0.095975137
Ga0436389_006_147182_147703	1245.1	1043	972	1691.8	933.3	906.5	1172.6	1057.4	764.6	1033.2	1452.3	1113.5	783.9	763.9	476.5	508.5	518.2	418.8	0.053982133
Ga0436389_006_149189_149611	84	260.8	275.3	140.6	493	460.1	152.5	183.3	219	121.8	164.3	339.4	204.3	153.1	188	132.2	260.8	182.1	0.478347226
Ga0436389_006_157336_160278	10.9	32.9	30.6	16	33.7	37.7	49.6	54.2	60.6	20.9	28	44	67.4	80	43.7	49.9	56.3	80.4	0.033343677
Ga0436389_006_169346_172165	382.4	424.4	370.5	238.1	381.2	314.5	1031.2	1122.5	371.5	1639.9	1065.9	335.5	1134.4	740.3	302.6	1476.6	342.5	281.9	0.576341431
Ga0436389_006_177279_178604	1759.1	2118.4	1996.2	104	188.2	144.5	446.2	401.1	766.3	832.6	532.4	587.9	560	345.7	586.1	898.3	521.3	840.2	1.84716E-05
Ga0436389_006_179061_179315	720.9	504.2	409	383	383.8	367.6	277.6	381.7	234.2	299.5	419	264.1	210.4	293.2	230.2	326.9	222.4	184.1	0.040298695
Ga0436389_006_182501_182869	2248.9	1861.1	1796.7	1908.2	2363.8	2405.5	1326.3	1370.1	1763.8	1099.8	1420.1	1620.9	1063.3	1284.1	1192.2	998.7	1701.8	1906.9	0.033343677
Ga0436389_006_211053_212597	185.6	230.6	206.1	11.4	25.2	21.8	97.9	94.9	47.3	52	27.8	19.3	77.9	115.8	97.4	64.3	92.8	68.1	4.78812E-05
Ga0436389_006_212614_213261	920	710.2	658	698.9	406.3	390.8	824.4	831.1	363.4	705.9	731.6	409.6	818	885	389.6	811.5	575.9	532.8	0.806884893
Ga0436389_006_214742_215242	181.9	163.2	161.3	69.4	87.4	89.7	99.2	114.3	71.7	54.2	82.8	74.7	88.8	131.7	84.7	69.5	55.4	68.2	0.001136206
Ga0436389_006_217480_217971	108.9	118.5	129.2	58.8	146	150	136.3	123.9	224.3	106	84.8	108.1	174.9	122.5	243	197	123.7	113.3	0.435942404
Ga0436389_006_2213_3172	1426.3	1129.8	1037.2	1629.1	1094.6	1054	1131	1200.4	833.8	991.1	1233.5	909.6	1472.1	1245.6	1109.9	872.4	811.6	579.3	0.183472688
Ga0436389_006_222551_223216	44.4	51.3	43.4	31.5	28.3	23	128.8	156.9	283.3	145.9	68.9	189.1	54.8	66.1	199.3	171.5	180.1	344.5	0.058414253
Ga0436389_006_224431_225378	407.2	867.6	917.1	386.8	1266.7	1287.3	793.3	859.5	1316.1	876.9	693.1	1097.6	1186.2	981.7	1137.5	1066.1	1054.9	1170.4	0.725555429
Ga0436389_006_22757_24919	70.6	70.3	72.5	45.3	46.1	41.8	115.9	79.4	60.9	116.1	57.7	43.3	94.3	135.4	45.5	92.1	51.5	38.7	0.584198868
Ga0436389_006_230698_232128	262.3	204.3	172.4	173.6	127.5	108.2	419.3	436.3	195.7	274	226.5	149.5	526.5	687.6	176.7	313.3	151.4	157.2	0.187137734
Ga0436389_006_232805_235027	824.2	1150.7	1119.7	197.8	478.8	487.8	1296.6	1087.8	815.7	751.7	531	514	2098.5	1806.9	2565.2	1251.6	793.7	519.2	0.000985048
Ga0436389_006_235033_236019	44.4	84.6	86.3	31.3	70.7	69.1	41	32.6	77.4	67.7	42.8	73.9	53.9	21.8	68.6	92.9	100.2	123.4	0.136302052
Ga0436389_006_238562_238903	248.6	102.1	93.5	184.2	112.9	105.7	127.3	130.2	59.4	73.7	150.1	47.6	94.4	135.8	104.9	107.3	59.2	34.3	0.603002482
Ga0436389_006_247274_249397	99.3	167.5	178	14.3	101.8	118.1	299.2	212.5	180.3	148.1	88	128.5	303.1	255.1	194.2	668.5	802	816.5	5.76564E-06
Ga0436389_006_249428_251854	834.7	1211.2	1128.7	435.4	416.7	376	282.8	265.3	464.1	449.8	321	320.5	255.2	284.1	568.8	341	422.5	488.5	0.001251338

Table B.21 continued from previous page

Protein Accession	NoN1	NoN2	NoN3	NO$_3^-$1	NO$_3^-$2	NO$_3^-$3	N1	N2	N3	R1	R2	R3	P1	P2	P3	E1	E2	E3	Q Value
Ga0436389_006_251960_252859	85	122.5	118.5	52.7	76.8	64.8	91.2	64.3	177.4	223.4	132.7	191.5	164.3	92.3	153.1	222.3	170.5	247.9	0.029633032
Ga0436389_006_252844_253824	144.8	94.8	95.4	76.1	52.5	47.4	333.3	159.7	84.5	170.5	96.6	29	376.4	348.6	259.2	213.5	138.6	109.9	0.030052607
Ga0436389_006_253818_255257	66.3	602.6	584	32.2	414	366.4	47.5	36.7	311.9	60.2	29.9	153.6	56.2	52.1	407.1	55	251.8	297.1	0.536635729
Ga0436389_006_255986_256747	983.2	203.5	199.1	152.9	30.5	33.4	417.6	277.8	109.9	224.5	215.9	85.8	547	549.6	179.3	243.1	89.9	55.8	0.38518359
Ga0436389_006_256732_258456	315.9	305.2	312.5	219.3	207	154.1	305.6	315.9	449.6	362	280.1	515.4	311.6	410.8	453	287	384.8	525.4	0.168289032
Ga0436389_006_258795_259097	250.2	325.8	309.4	115	180.6	178.8	152.5	156.2	301.6	178.4	183.7	236.8	204.7	185.9	693.6	225.9	283.8	185	0.58538959
Ga0436389_006_26936_27742	2011.1	1672.7	1643.5	1087.5	943.1	844.7	958.9	954.8	782.4	726.7	854.2	633.5	1820.7	784.3	801.2	916.1	731.6	533.2	0.026703127
Ga0436389_006_270480_271775	27095.9	21101.5	19900.4	17030.1	18901.4	16416.7	6584.6	5619.1	5086	5876.3	4794.4	4186.1	6392	6909.9	5989.5	6105.2	5847.7	5577.4	6.12887E-06
Ga0436389_006_274310_275056	2341.1	1803.5	1714.9	1059.9	1029	971.3	1314.8	974.8	582.4	787.1	574.9	415.5	901.1	1019.2	886	1274.1	636.2	562.6	0.007640158
Ga0436389_006_275088_275813	1104.7	2618.3	2454.3	638	1703.7	1545.7	654	910.1	1276.7	1033.7	751.1	1101.5	996.7	815.4	1199.5	1247.8	944.7	1208.5	0.168289032
Ga0436389_006_276270_277475	3533.2	3844.5	3947.7	1364	2172	1697.7	844.2	656.1	891.2	553.5	444.5	719.8	721.8	919	1026	755.8	902.8	980.2	7.40069E-07
Ga0436389_006_277642_278151	177.8	243.2	235.6	152.3	273.2	252	222.6	272.1	304.1	263.7	203.5	293.2	170.4	173.1	192	217.1	269.6	270.6	0.326122574
Ga0436389_006_281210_281758	55.6	78	80.1	15.2	32.2	32.6	57.5	86.8	24.3	64.5	60.3	45.5	26.8	36.2	18.7	76.6	55.9	64	0.058925315
Ga0436389_006_282399_284078	49	334.3	306.3	51.2	283.3	276	588.9	758.3	1284.7	1621.7	892.3	1169	658.9	301	378.7	1782.6	811.8	1856.2	0.013122001
Ga0436389_006_284265_284750	1751.7	1256.4	1079	422.9	221.6	227.6	490.3	477.2	280	545.3	524.9	234.9	563.2	570.6	370.6	408.7	330.6	275.6	0.001293886
Ga0436389_006_284840_286663	1038.6	1309	1251.9	1145.5	1257.9	1190.3	1506.5	1708.1	1481.2	1543.2	1167.2	1311.6	1906.8	1604.3	1532.3	1493.3	1360.2	1111.3	0.049687943
Ga0436389_006_286749_287864	207.2	354.7	344.9	186	334.1	292.6	185.6	165.6	457.5	144.4	182.2	418.4	259	243.2	382.7	152.5	309.3	493.7	0.990699199
Ga0436389_006_289696_290589	93	31.3	37.6	72.2	9.7	21.9	136	206.9	31.9	314.4	199.6	25.7	154.9	92.2	73.5	291.9	37.2	44.4	0.624233989
Ga0436389_006_294469_294783	1378.3	1568	1447.2	825.4	471.5	516.1	1194.6	841.6	615.7	672.5	535.9	649.7	1440.8	1571	418.9	662.1	530.6	592	0.059358619
Ga0436389_006_33675_35189	230.6	90.3	78.6	190.7	73.4	68.1	218.2	272.1	810.3	286	252.8	696.1	387.4	271	229.8	583.9	406.2	982.6	0.15638334
Ga0436389_006_3692_4618	203.4	240.6	247.5	152.2	176.8	158.7	288.7	271.9	223.2	239.5	205.5	176.9	345.7	328.3	284.3	360.3	217.2	211.2	0.047000664
Ga0436389_006_37954_39603	99.6	122.3	121.1	46.2	66	65.8	116.3	138.5	64.9	119.6	115.7	46.9	224.3	154.9	41.9	180.2	83.7	58.8	0.694713171
Ga0436389_006_47155_47538	52.8	87.3	83	30.1	70.8	58.1	57.1	66.8	64.9	46	54.4	70.3	51.3	46.1	168.9	62.7	56.5	34.6	0.776232307
Ga0436389_006_47679_48884	5917.7	5258.1	5174.5	6836.7	5956.8	5420.8	3840.6	4253.9	3318.5	3669.9	4720.2	3956.2	3394	3625.4	3019.2	3366.1	4084.1	3268.9	0.001280805
Ga0436389_006_48929_49336	347.6	667.4	623.1	218.7	116	124.4	206.4	210.6	399.8	88.8	67	236.3	349.3	275.2	378.7	139.2	330.5	410.5	0.035792517
Ga0436389_006_51533_52327	648.6	511.9	499.7	197.9	237.6	231	193.5	193.7	267.2	245.5	209.5	209.5	173.2	172.6	302.2	244.4	319.6	237.9	0.000585037
Ga0436389_006_59218_61458	2959.7	1841.9	1854.3	2833.8	1775.7	1678.5	1911.4	2060.3	1178	2075.4	2261.9	1260.9	1813.3	1697.9	1516.6	2141.7	1375.5	999.2	0.703617122
Ga0436389_006_62459_63187	71.9	284.4	262.9	138.4	319.3	306.6	220	302	424.2	484.6	453.9	381.1	228.9	168.5	311.6	594.9	391.7	577.2	0.038960755
Ga0436389_006_63184_64002	345.3	180.1	138.8	505.5	162.2	153.6	268.6	291.8	121.9	525.2	464.5	161.7	237.3	244.3	141.3	379.6	115.1	99.1	0.748339619
Ga0436389_006_66146_66901	116.3	53	63.1	81	70.6	72.9	170	168.7	32.2	171.6	143.1	38.3	372	364.1	279.7	258.4	83.2	34.2	0.025134825
Ga0436389_006_67191_67577	3523.5	2493.6	2090.8	4639.4	4875	4690.2	2902.8	3066.1	2421	2150.7	3674.7	2875.6	1997.5	2237.7	2080.7	1951	2488.6	1739.4	0.00283184

Table B.21 continued from previous page

Protein Accession	NoN1	NoN2	NoN3	NO$_3^-$1	NO$_3^-$2	NO$_3^-$3	N1	N2	N3	R1	R2	R3	P1	P2	P3	E1	E2	E3	Q Value
Ga0436389_006_67639_68166	1283.1	1428.4	1350.2	1623.5	2595.9	2377.6	1508	1498	1279	1313.6	1461	1387.4	1257.6	1310.5	890.3	1350.6	1284.7	961.8	0.017711586
Ga0436389_006_69080_69787	1861.6	1709.2	1715.4	2292	2841.4	2641.1	3570.8	3345.1	5671.5	2684.7	2454.3	3841.6	2955.3	2923.6	3478.1	3501.5	4734.5	6554.1	0.04376215
Ga0436389_006_69877_70302	1083.6	1434.4	1431.2	1523.8	2678.1	2534.9	2703.4	2832.1	1559.9	2385.9	1698.1	1417.4	1914.7	1425.4	1085.5	2253.1	1529.8	1422.3	0.281752357
Ga0436389_006_70462_71172	982.9	1001.4	954.7	867.8	1326.2	1269.1	1089.9	1165.4	919.7	1105.3	880.3	817.6	969.5	950.4	754.8	1237.5	942.5	915.9	0.562107381
Ga0436389_006_71515_74388	63	20.3	20.6	45.5	14.8	12.3	133.4	88	34.3	56.4	61	49	112.4	143.7	36.1	68.6	83.6	84.1	0.210351732
Ga0436389_006_74880_76175	1042.6	832.9	779.4	717.4	578.5	481.3	576.2	524.6	481.9	923.8	665.3	475.8	600.9	618.7	267.1	386.7	386.6	467.8	0.071433399
Ga0436389_006_76194_77891	195.7	106.5	110.1	209.5	105.1	98.7	206.5	219.8	136.1	142.1	168.3	138.7	235.6	262	107.5	148.8	98.7	135.3	0.585880628
Ga0436389_006_79714_80949	221.8	174.4	176.2	165	137	135.6	295.7	322.3	289.7	256.4	194.4	238.3	288.4	310	317.3	499.6	255.4	286.9	0.035079555
Ga0436389_006_84420_85865	1354.8	1008.4	984.8	1128	1013	998.4	1094.7	1191.5	1143.8	967.6	1113.3	949.7	1181.2	1165.2	809.9	1244.7	974.2	930.1	0.919505908
Ga0436389_006_89006_89674	183.7	348.4	321	133.3	265.2	291.8	103.2	99.5	255.5	136.3	131.2	223.7	151.2	104.2	136.2	273.7	245.3	335.8	0.139167839
Ga0436389_006_89745_90203	1227.8	562.2	520.4	774.6	411.3	383.2	737	704.9	264.7	453.5	628.6	345.9	567.6	1038.7	185.4	312.1	232.6	159.9	0.536635729
Ga0436389_006_92519_94498	153.5	148.7	139.8	84.6	118	123.7	190.7	178	100.7	145.9	106.6	94.8	200.3	289.8	122.7	208	93.4	90.2	0.429926088
Ga0436389_006_98226_98930	134.7	130.1	123.4	72.4	100.7	90.5	459.1	557.8	137.9	888.8	522.2	141.4	451.2	289.5	283.2	1026.2	149	140.7	0.530107706
Ga0436389_007_103059_104093	167.2	69.5	60.1	108.3	37	45.1	131.1	179.3	59.1	166.7	176.4	42.2	337.5	268.4	152.6	192.4	76.1	56.4	0.169659813
Ga0436389_007_106454_107632	50.1	11.4	3.5	35	8.3	6.3	457.2	369.8	97.2	403.3	231.3	11.4	212.2	115.3	339.6	409	84.4	11.6	0.308037375
Ga0436389_007_11941_12342	491.3	293.4	270.9	623.3	415.4	396.3	798.3	767.5	662.4	546.3	461.5	879.6	1040.2	1123.2	298.7	567.2	482.2	801.7	0.361692803
Ga0436389_007_126216_127262	361.4	410.1	397.7	249.6	271.2	271.6	258.5	257.9	403.1	241.6	340	301.6	189.1	196.7	863.8	191.1	323.9	208.1	0.83253458
Ga0436389_007_128358_129482	547.4	920.2	882.2	325.7	815.8	819.8	471.7	415.7	845.9	328.7	428.1	654.5	970.6	713.6	790.2	309.1	779	750.9	0.565518388
Ga0436389_007_132187_132612	405.1	735.1	696.6	630.2	975	1020.9	1032.9	961	1391.1	792.9	711	950	1055.6	1091.3	1953.7	823.5	1289.1	962.5	0.165891133
Ga0436389_007_132659_134155	30.4	232.6	230.7	38.3	283.3	315.2	82.7	56.1	289.9	54	57.2	294	94	81.3	419	103	334.7	265.8	0.960821659
Ga0436389_007_13269_14798	19.5	231	196.8	9.4	183	195.1	61.7	71.1	256.4	114.2	60.4	246.3	174.5	135.8	300.5	171.2	183.8	222.7	0.907724681
Ga0436389_007_136249_137358	384.8	275.3	258.1	557.2	368.4	306.4	783.2	725.9	462.3	580.8	611	472.7	541.5	613.7	404.5	550.2	494.9	610.2	0.075256474
Ga0436389_007_143700_144581	525.9	123.5	113.2	464.9	64.6	67.4	399.1	411	96.6	498.6	641.4	120.4	233.4	352.4	141.4	319.7	103.6	74	0.794431521
Ga0436389_007_150004_151392	443.1	403.5	412.8	1297.8	623.9	795.9	756	799.3	586.6	708	732.4	638.2	810.5	456.5	560.6	830.6	689.2	551.6	0.21032359
Ga0436389_007_151376_151486	34.3	58.7	52.7	87.5	135.7	148.1	50.2	41.3	65.5	55.6	36.6	77	46	30.8	25.2	44.9	103.8	67.3	0.02020216
Ga0436389_007_15768_16574	289.9	359	319.3	361.6	468.6	453.9	287.7	317.8	300.8	239.3	341.5	332	255	203.8	250.1	284.1	300.7	219.4	0.014632382
Ga0436389_007_16624_17823	22	43.7	37.3	15.5	25.7	19.7	57.9	64.1	43.8	95.2	70.8	39	78.5	63.5	29.3	165.2	86.4	155.1	0.010266945
Ga0436389_007_20391_21530	179.5	175.2	136.5	286.1	201.5	222.5	188.5	231.7	493.8	145.9	284.8	473.5	202	144.7	226.2	154.1	443.6	455.9	0.558388933
Ga0436389_007_21606_22052	40.8	112.8	106.8	33.5	126.1	124.1	87.7	125.2	152.1	235.2	125.7	157.2	127.8	56.3	118.9	280.3	190.3	252.3	0.036138173
Ga0436389_007_22652_23269	173.3	832.4	765	56.4	452.8	349.3	143.1	164.7	350.6	410.8	273.5	346.9	162.6	115.8	310.3	533.5	300.4	333	0.351469031
Ga0436389_007_23747_24559	37	263.2	262.9	47.3	297.3	282.8	75.1	154.3	386.4	196.8	145.1	345.1	102.5	61.8	490.1	210.8	376.7	330.4	0.960821659

Table B.21 continued from previous page

Protein Accession	NoN1	NoN2	NoN3	NO_3^-1	NO_3^-2	NO_3^-3	N1	N2	N3	R1	R2	R3	P1	P2	P3	E1	E2	E3	Q Value
Ga0436389_007_34915_35502	982.2	572	618.2	871.9	654.4	738.3	1009.5	945.2	476	564.3	598.1	443.5	1084.5	1689	496.2	534	561.1	431.5	0.38518359
Ga0436389_007_36310_38259	157.3	787.6	788.7	125.8	631.6	607.2	123.3	119.6	393.8	92.2	106.7	394.1	170	149	594.8	88.3	457.8	249.6	0.547929064
Ga0436389_007_38319_38762	2610.2	1564.2	1498.6	1540.9	1931.5	1769.4	2011	2073.3	851.8	1691.2	1991.6	1120.3	1382	1444.1	864.2	1804.3	888	658.9	0.576341431
Ga0436389_007_38809_40152	1310.5	1428.8	1404.4	1282.6	1607.9	1532	1731.1	1803.2	984.5	1346.1	1283.1	970.6	6496.8	4009.1	687.9	1577	1100.6	1215.4	0.296850136
Ga0436389_007_41358_41957	8377.7	4106	4041.4	3354.9	1938.5	1934.8	2483	2189.1	1705	2105.1	2267.3	1491.1	2180	1750.5	2253.9	2332.5	1575.4	1231	0.044531423
Ga0436389_007_42027_42698	1991.6	1235.6	1198.1	1569.8	1859.3	1696.8	1764.6	1762.3	1107.6	1033.5	1595.3	1210.9	1406.4	1296.3	863.6	1326.1	1298.8	970.9	0.441497853
Ga0436389_007_42736_43077	209.8	133.5	118.9	107.9	100.4	92.9	189.3	213.4	116.6	203.3	175.2	67.9	259.7	193.4	148.6	197.2	113.8	110.4	0.478705645
Ga0436389_007_45787_46485	25.8	91.4	76.3	20.9	58.5	61.4	67	80.5	147.2	83.9	55	84.2	71.8	49.8	194	161.6	172.4	184.1	0.092484161
Ga0436389_007_47651_48175	188.8	95.9	109.7	534.1	295.4	290.5	178.7	193.2	77.3	170.5	260.7	155.3	95.4	104.6	46.4	149.9	113	63.9	0.023400743
Ga0436389_007_49464_52028	4759.5	5129.8	5187.3	1315.3	2107.1	2382.3	1860.1	1557.8	1650.9	1349.9	1383.8	1456.7	1644.8	1999.4	1536.1	1382.2	1695	1593.5	2.40887E-06
Ga0436389_007_53434_54396	364.8	462	406.7	243.3	337.2	306.7	445.8	593.7	487.2	546.9	492.3	391.4	729.8	497.2	1314.5	799.9	459.1	371.9	0.189108417
Ga0436389_007_57631_59523	2352.6	2035.2	1892.8	1630.3	1608.5	1603.3	2552	2637	1937.7	2266.1	2684.7	2045	1891.3	1819.5	1369.9	2017.3	1397	1439.4	0.052604149
Ga0436389_007_60363_60617	58.5	92.9	62.7	73.7	111.5	135.1	74.2	102.8	114.8	64.4	79.7	89.6	61.3	68.6	134.7	53.8	76.6	49.2	0.432931315
Ga0436389_007_6310_6993	2.5	3.7	2.3	1.6	3.9	5.1	538.1	444.5	506.6	884.8	449.5	413.3	409.4	135.8	19.4	1062.8	164.3	637.7	0.047829133
Ga0436389_007_63822_66329	174.3	268.6	278.3	142.6	227.6	236	348.6	249.3	269.6	211.3	186.7	251.6	505.5	473.1	350.8	241	304.9	279.2	0.011410112
Ga0436389_007_7050_8879	73.5	99.2	102	87.6	111.1	120.2	101.7	94	48.3	66.8	52.9	44.4	104.8	104.9	71.5	68.6	50.1	48	0.066603213
Ga0436389_007_73340_74347	53.1	69.9	72.8	160.1	49.4	48.3	71.6	76.8	79.3	65.4	75.3	55.8	128.5	77.1	114.7	77.5	153.3	159	0.361692803
Ga0436389_007_76955_78070	43.5	247.4	222.7	42.9	218.2	172.6	85	75	398.2	67.8	79.7	249.4	101.3	75.1	386.2	147.5	228	202.6	0.990699199
Ga0436389_007_85979_87067	716	703.4	699.9	494.9	551.1	563.4	656.6	746.8	639	649.8	513	615.1	654.8	603.7	392.7	663.1	488.1	576.4	0.200619907
Ga0436389_007_87566_89926	2261	1633.2	1629	2395	1681.7	1667.9	3356.9	2403	1645.8	4382.8	3524.7	1961	2561.4	2453.1	1901.8	2597.9	1390	1237.6	0.30873986
Ga0436389_007_97943_99694	294.7	152.7	140.4	194	99.8	114.5	375.9	338.6	230.7	333.4	281.5	127.5	966.7	203.3	110.8	372.2	196.8	170.8	0.743515551
Ga0436389_008_100726_101124	436.1	305.7	287.9	596.8	510.3	516.7	280.5	338.9	291.5	168.9	261.9	369.9	314.6	329.3	184.2	128.4	272.2	229.8	0.01336335
Ga0436389_008_101180_101446	40.4	199	212.3	56.2	254.7	282.6	54.2	76.1	427.2	43.6	31.9	279.9	73.9	57	238	79.4	407.3	477.1	0.740020191
Ga0436389_008_103159_103935	19.7	105.6	100.9	8.1	81.3	80.4	67.5	48	367.5	36.6	8.6	339.2	73.7	104.6	330.8	80.2	304.7	487.2	0.609161724
Ga0436389_008_107232_107804	961	1235	1170.7	869	1815.8	1869.4	772.7	646	867.6	518.9	670.9	977.3	418.5	502.9	868.7	416.7	890.2	640.9	0.054683339
Ga0436389_008_108347_108646	5685.2	8444.7	7582.2	1826.8	3254.8	3289.3	1471	1151.8	2066.7	1317.1	1519.3	2498.4	744.7	785.1	2112.8	1371.9	2166.5	1667.7	0.000289927
Ga0436389_008_10887_14177	2993.6	4083.5	3925.1	3974.3	5875.5	5398.9	3361.8	3677.8	4427.8	2478	2891.2	4570.7	4879.5	4165.5	4949.1	2663.4	4952.4	4194.1	0.325930411
Ga0436389_008_109754_111340	1973.7	1553.5	1413.1	1471	1375.3	1393.6	2016.3	1885.4	1841.8	1728	1909.4	1579	3140.7	2424.1	1955.7	2091.2	1559.2	1612.7	0.056531658
Ga0436389_008_113114_114532	239.5	291.2	268	642.3	263.4	264.7	276.4	266.4	223.3	364.7	502.7	345.8	196.1	445.6	226.3	216.2	268.2	278.8	0.545815122
Ga0436389_008_116021_116248	57.8	247.1	246.1	50.3	457.3	409.7	82.6	114.2	207.9	106.6	123.4	247.1	60.8	49	157.4	265.9	179.4	135	0.512813407
Ga0436389_008_118044_118871	485.1	95.4	102.8	353.3	64.6	61.7	642.4	1389.2	99.4	916.2	830.5	75.6	699.8	338.9	66.7	1111.8	91.8	106.5	0.74018071

Table B.21 continued from previous page

Protein Accession	NoN1	NoN2	NoN3	NO_3^-1	NO_3^-2	NO_3^-3	N1	N2	N3	R1	R2	R3	P1	P2	P3	E1	E2	E3	Q Value
Ga0436389_008_118912_119130	318.2	107.5	109	266	164.6	178.9	115.5	192.8	129.3	75.5	154.5	84.2	48.5	80.5	115.8	101.6	118.1	67.3	0.308461832
Ga0436389_008_121728_122231	55.1	117	98.3	57.7	74.5	80.3	121.5	99.6	141.8	117.5	110	120.6	171.3	171.5	160.1	243.5	331.9	229.5	0.000538839
Ga0436389_008_124196_125833	10880.1	8569.2	8007.6	11923.6	12525.6	11803.1	11837.5	12639	9731.5	8013.2	10135.4	7818.9	7414	7147.4	6174.3	6867.4	9137.7	6870.1	0.007606261
Ga0436389_008_125887_126198	7451.5	7291.6	5695.8	2082.7	3865.6	3787.4	2313.2	2268.7	2145.4	1578.3	2004.5	2020.9	1068.4	1166.3	1151.4	1654	1967.4	1551.5	4.05456E-05
Ga0436389_008_126447_127913	6342.4	6383.2	6165.4	7242.2	7728.4	7401.9	6287	8325.2	9506.9	7506.9	6893.5	7473.9	6802.5	5923	8035.1	7699.2	7551.7	8466.7	0.328242734
Ga0436389_008_137011_137478	1126.6	1501.6	1331.9	1146.8	2083.8	2042.1	1370.6	1221.8	1467	793.9	933.3	1522.6	1037	882.7	1895	835.2	1574.2	1421.7	0.609161724
Ga0436389_008_14236_16140	2183.7	2204.8	2135.3	2479.1	2649	2435.8	4800.9	4994.4	5790.2	4813.8	3414.7	4314.3	6307.8	5007.4	4876.9	5014.7	4061	6201.8	0.00154649
Ga0436389_008_144427_145575	498.9	347	326	226.7	141.7	128.9	426.2	442.2	365.5	368.9	306.5	174.5	348.9	306.5	202.1	378.2	198.9	296	0.078411999
Ga0436389_008_146464_147237	42.7	117.5	107.8	31.2	56.1	51.8	162.8	177.6	121.7	211	158.7	53.9	133.3	92.8	140.7	224.6	461.9	413.6	0.007926666
Ga0436389_008_159598_160773	179.4	290.1	314.9	163.8	205.3	190.3	160.2	167.3	151.8	113.4	162.9	167.5	119	135.7	230.7	198.9	204	187.4	0.136704212
Ga0436389_008_16214_20329	3780.1	5258.7	5366.2	5457.1	7650.5	5681.3	4828.6	5076.6	6042.8	3652.2	3946.2	5383.1	4092	4222.8	5649.3	3854.7	6037.9	4941.9	0.400407826
Ga0436389_008_163251_164024	3910.1	2772.1	2502.8	1736.8	1746.8	1763.7	2045.5	2023.4	2154.8	1807.4	2054.1	1934.9	2001.1	2018.5	2081.1	1601.5	1979.9	2173.5	0.025479647
Ga0436389_008_164060_164410	401.5	379.4	394.5	342.3	305.9	303.8	502.1	557.1	419.3	381.5	371.5	271	592.5	392.4	519.8	752.1	369.6	274.8	0.453487363
Ga0436389_008_167769_168452	53	44.6	39.6	51.2	33.6	30.4	128.5	110.9	38.4	56.3	61.6	41.9	203.5	183.1	34.1	82.9	26.5	37.3	0.212501218
Ga0436389_008_170222_173014	723.3	1130.4	1080.5	784.2	1299.4	1187.4	1012.2	1137.2	1241.5	697.3	752.7	1302	1767.6	1502.8	939.5	538.2	1118.5	980.6	0.487767376
Ga0436389_008_179502_179948	405.7	308.2	298.2	133.1	216.9	257.1	171	172.1	203.1	162.4	145.3	173.9	159.1	160.9	230.1	272.5	217.1	263.3	0.014632382
Ga0436389_008_186580_187602	56.3	42.9	44.7	50.9	43	33.6	73.5	70	33.8	55.1	40.4	40.3	83.8	90	50	72.8	39.5	27.6	0.383579104
Ga0436389_008_189786_191453	3158.1	1994.1	1952.9	3463.3	2765.5	2530.8	6914.2	7168.5	3790.8	6399.5	5009.9	3167.8	8307.9	6507.2	3214	6968	3347.5	3862.5	0.202147531
Ga0436389_008_191842_192594	640	554.5	536	460.7	496.5	488.2	563.5	565.4	818.5	473.3	451.8	731.8	488.3	462.1	550.1	478.8	600.1	887.1	0.594063618
Ga0436389_008_193810_194652	245.6	339.3	328.6	297.7	262	242.8	223.6	232.5	306.3	120.7	157.9	293.2	389.1	348.8	240.6	151.7	273.9	234.6	0.285099289
Ga0436389_008_194738_195439	113.3	44	46.6	77	34.3	21.9	126.3	137.8	138.4	119.7	114.5	82	137.5	111	51.3	113.2	129.6	128.1	0.058987037
Ga0436389_008_200112_200912	72	59.2	59.7	54.2	54.8	46.1	116.7	119.5	131.7	160.1	119.3	102.4	162.2	141	131.6	246.4	106.3	130.8	0.03864506
Ga0436389_008_204171_205112	1221.8	1137.8	1019.9	1171.9	1623.6	1526	769.2	878	732	579.3	923.2	996.6	666	791.3	445.4	481.4	699.5	618	0.003400378
Ga0436389_008_20510_21568	1.5	39.4	38.4	2.7	101.2	89	312.6	287.6	51.6	96.1	65.1	41.4	1448.2	1064.6	49.3	101.2	44.5	33	0.103285174
Ga0436389_008_205169_205729	598.2	303	303.9	734.8	353.3	377.7	359.8	390.4	325.2	291.7	417.3	249	191.6	212.1	199.3	290.6	235	188.9	0.210351732
Ga0436389_008_209555_210064	2003.6	1934.1	1961.5	4226.4	2843.1	2513	1678.5	2457.1	2406.8	1958.2	2582.8	2657.1	1417	1344.1	1581.5	1341.4	1789.8	1404.5	0.025134825
Ga0436389_008_212849_213967	144.1	150.3	125.7	126.9	123.5	97.5	351.3	334.6	205.3	244.2	220.2	164.6	460	314.5	209.7	283.6	261.6	239.8	0.03561983
Ga0436389_008_215246_216337	307.2	244.9	242.4	265.3	200.3	207.1	234.1	317	231.9	341.9	326.3	219.9	275.5	229.7	220.9	401.2	175.2	221.9	0.877551687
Ga0436389_008_221859_222092	80.8	78	41.3	89	91.5	132.3	99	81.8	67.5	30.5	47.4	50.2	41.8	37.5	48.7	14.8	133.1	143.1	0.292496601
Ga0436389_008_229434_230771	127.9	92.5	84.4	123	58.8	74.4	121.4	144	132.9	233.6	173.8	93.5	135.3	130.5	135.1	256.3	137	143.3	0.260240059
Ga0436389_008_233101_234186	536.1	1165.6	1222.7	525.6	1315.5	1220	598.6	559.2	1031.3	549.3	500.2	965.4	786.4	721.5	1192.1	1023.3	1918.4	1573.5	0.243497196

Table B.21 continued from previous page

Protein Accession	NoN1	NoN2	NoN3	NO_3^-1	NO_3^-2	NO_3^-3	N1	N2	N3	R1	R2	R3	P1	P2	P3	E1	E2	E3	Q Value
Ga0436389_008_234183_235682	461.9	299.1	294.7	263.4	190.6	208.5	553.1	615.9	231.1	487.6	287.9	212.6	539.5	516.7	220.8	563.5	247.5	263.9	0.61319273
Ga0436389_008_235682_236746	220.7	136.7	135.8	141.7	98	93.5	168.7	197.6	368.1	196.1	212	353.8	304.5	228.6	94.5	262.5	238.1	513.1	0.288927401
Ga0436389_008_242405_242728	3994.1	2853	2478.1	3455.9	3059.2	2912.5	1831.6	2390.2	1831.3	1631.6	2587.2	1861.3	1356.5	1566.5	1087.8	1720.6	1637.4	1800.9	0.007811318
Ga0436389_008_242971_244134	4885.2	4025.3	4251.3	2386	1938.5	1644.4	2708.8	2992.8	3496.4	2410.9	2528	2807.3	2580.5	2547	2327.7	2646.3	2344.6	3243.7	0.001077545
Ga0436389_008_246097_248760	406.6	744.7	640.6	319.7	851.9	805	314.9	281.1	903.4	455.1	345	935.9	573.2	519.9	1139.9	247	915.8	946.4	0.960821659
Ga0436389_008_253221_253568	174.3	190	122.7	116.3	157.6	197.9	112.6	113.4	212.1	97	65.7	146.2	127.5	120.8	126.4	98.9	127	180.5	0.62709948
Ga0436389_008_254467_255165	131	617.6	611.8	119.8	452	477	135.6	155.1	986.8	280.4	198.2	703.8	111.7	108	556.5	379.4	543	849.7	0.885895212
Ga0436389_008_3_554	581.5	424.3	429.1	674.6	590.1	581.3	1598.3	772.2	524.1	1885.7	1515.5	926.6	874.2	744.7	555.8	820.9	550.9	425.9	0.090741437
Ga0436389_008_41003_41635	48.3	38.1	37.8	44.4	43.5	44	120.9	150.4	54	71.7	66.2	54.1	107.5	155.8	156.7	184.7	56.4	18.2	0.195171329
Ga0436389_008_46935_48230	256	397.3	383.8	214	446.3	399.6	179.8	207.3	368.2	130.5	222.7	367.4	223.2	246.1	318.2	109	359	263.9	0.707823824
Ga0436389_008_49509_49847	5639.9	3107.1	2886.5	1586.3	1550.2	1558.9	2642.3	1595.8	989.4	1600.2	1484.9	875.3	2234.8	2073.9	1052.2	1725.4	1481.2	1351.8	0.054038613
Ga0436389_008_5382_6095	913.9	723.7	746.9	351.9	402.1	420.5	504.7	450.5	587.9	553	529.4	372.2	289.5	250.4	245.5	736.2	387.6	479.3	0.006236576
Ga0436389_008_53909_54109	413.3	759.9	796	464.5	697.3	882.6	545.3	659	549.2	533	448.5	711.6	781.2	682.6	388.3	453.3	507	355.7	0.669359899
Ga0436389_008_54156_54644	6019.3	5252.4	5113.4	6488.8	6385.7	6582.5	5486.4	5723.6	6247.5	5026.7	6695.7	7161.2	3503.3	3692.5	4140.6	4209.8	5062.5	5345	0.008543143
Ga0436389_008_54688_55173	4000.5	4726.4	4543.7	4462.8	5680.1	6050.3	3596.8	3916.1	4242.3	4249.5	4878.4	5218.3	3688.8	3093.6	3575.4	3048.5	3854	3664.1	0.014911347
Ga0436389_008_55558_58584	5733.9	5210.1	5006.1	6830.1	5206.8	5906.2	5002.1	5352.5	4359.6	5411.7	5629.6	5453.3	4189	4141.9	4183.1	4728.5	4192.8	3284.1	0.020802167
Ga0436389_008_58734_59969	56	169.5	172.1	51.5	175.4	183.8	117.3	116.4	193.3	165	117.9	142.7	118.1	120.3	251.7	91.5	232.3	199.6	0.965454609
Ga0436389_008_61457_61672	455.4	421.8	359.9	567.1	637.2	475.5	585.3	689.3	298.6	329.8	470.8	352.1	342.8	359.5	174.8	330.7	294	234	0.097310933
Ga0436389_008_61773_62255	512	723	622.6	581.7	1082	996.2	539	603.2	638.1	565.4	640.8	820.9	830.4	872.9	496.8	308.3	1026	974.7	0.749001815
Ga0436389_008_62252_62779	1431.6	1223.3	1043.6	1563.4	1825.4	1680.1	1492	1476.2	1038.1	1613.1	1660.4	1118.7	1768.4	1646.9	1455	1279.5	906.7	737.2	0.068567799
Ga0436389_008_62779_63327	1140.7	1099.8	999.7	1679.7	1413.2	1300.8	867.3	984.5	917.8	900.5	1089.8	943	752.5	600.1	605.8	1070.5	851.8	743.6	0.001541807
Ga0436389_008_63390_64907	7728.8	7216.1	6867.4	8160.4	9101.5	8358.5	6039	6366.1	6042.8	5383.3	6276.7	6850.2	4300.9	4380.3	3995.6	5646.1	6037.2	5852.9	5.3563E-05
Ga0436389_008_64921_65871	936.4	1082.3	966.9	1098.1	1504.2	1291.3	1018	1207.2	900.3	1031.9	900.5	916.2	1128.1	1026.2	421.7	987.8	694.7	551.2	0.211491823
Ga0436389_008_71122_71487	43.7	84	88.6	27.8	86.6	89.4	440.4	617.6	84.6	1190.8	653.7	81.1	627	238.9	52.5	1293.7	50	75.5	0.615998298
Ga0436389_008_72755_73261	25	184.3	181.2	29.8	95.8	94.1	256.8	280	554.5	1428	783.7	880.9	442.4	131.5	155.2	1583.9	224.1	632.6	0.066603213
Ga0436389_008_73445_74017	8565.9	7263.2	7245.8	8430.3	6133	5813.6	4596.5	4485.2	3632.6	3695.8	4984.7	4279.9	3530.1	3573.5	3056.9	2978	3562.4	3063.1	0.000736838
Ga0436389_008_78195_79499	75.6	76.5	74.6	55.4	90	71.4	214.3	202	103.3	135.4	93.6	68	425.8	439.6	83	144.1	88.7	72.5	0.10412797
Ga0436389_008_8027_9430	220.7	195	193.8	142.7	205.2	192	208.1	175.1	164.7	84.3	109	166	318	208.3	125.3	120.1	249.2	244.8	0.522051737
Ga0436389_008_83561_85066	747.9	518.6	466.2	1032.9	524.6	460.8	1090.2	695.1	439.2	620.5	636.7	453.7	493.1	491.2	494.4	576.5	376.6	293.5	0.609161724
Ga0436389_008_87560_88366	268.2	165.8	176.3	337.7	204.4	180	246	340.5	468.2	268.2	215.6	326.9	281.5	205.3	263.2	369.4	321.2	581.7	0.178013453
Ga0436389_008_88386_88907	302.2	172.4	154.3	214.6	159.7	155.7	236.6	255.8	241.2	218.7	173.4	159.4	193.9	148.7	260.9	272.3	200.1	232.5	0.593279349

Table B.21 continued from previous page

Protein Accession	NoN1	NoN2	NoN3	NO$_3^-$1	NO$_3^-$2	NO$_3^-$3	N1	N2	N3	R1	R2	R3	P1	P2	P3	E1	E2	E3	Q Value
Ga0436389_008_88912_89784	84.4	77.9	92.2	93.5	85.8	101.6	134.1	177.8	165.4	157	117.4	155.1	181.5	160.7	87.1	199.1	129.7	164.6	0.058414253
Ga0436389_008_89784_92189	2055.7	2503.6	2346.4	1699.1	2264.8	2201.3	1399	1465.8	2530	1302.2	1551.6	1905.1	1121.3	1182.8	1993.3	1639.1	1685.6	1737.2	0.260291619
Ga0436389_008_904_2343	1168.9	749.3	720.2	1038.3	745.6	701.3	2230.6	962.4	689.8	799.2	737.5	567.5	849.6	640.7	821.3	1052.2	852.6	779.9	0.612302307
Ga0436389_008_92233_92991	28.9	125.9	106.6	35.4	107.5	92.4	51	87.9	236.3	65.8	53.6	203.7	159.4	91.9	76.7	79.2	122.1	261.4	0.880099815
Ga0436389_008_92988_93992	29.7	118.6	120.8	23.5	91	101.5	87.5	80.7	70.2	56.7	34.2	76.8	120.4	112.8	116.2	45.6	80.6	59.6	0.371975359
Ga0436389_008_94299_95657	134.4	402.8	384.3	63.2	327.7	320.8	262.9	226.4	567	171	117.2	329.7	194.6	220.9	303.2	216.7	394.8	448	0.764183164
Ga0436389_008_9533_9844	587.2	238.2	227.4	853.7	412.6	364.4	485.9	497	233.8	491.5	559.6	370.6	324.6	370.1	222.9	494.1	265.5	146.6	0.626719426
Ga0436389_008_97722_99221	418	378.5	378.6	250.6	344.1	350.2	629.4	555.7	803.1	408.6	317.6	649.3	549.3	746.7	821.1	536.4	622.6	842.3	0.035792517
Ga0436389_008_99359_99718	63.9	259.9	251.5	72.4	231.1	243.2	66.6	56.6	254.8	45.3	33.1	236.8	32.3	49.8	265	70.7	254	259.1	0.884558714
Ga0436389_015_2539_2748	345.4	321.2	278.7	177.9	400.2	397.8	192.5	221.8	143.6	222.9	163.8	151.3	168	190.2	160.3	182	148.7	65.8	0.044344903
Ga0436389_015_5500_5874	146.9	14.9	13.1	76.7	19.5	14.1	124	90.3	141	86.5	24.1	170.7	136.6	153.3	45	113.4	96.1	235.2	0.45710066
Ga0436389_016_16760_17248	302.6	457.1	491.8	178.6	168.9	169.2	240.7	252.2	260.5	273.3	250.8	305.9	287.2	214.7	208.4	265.6	193.7	181.3	0.008889556
Ga0436389_016_17245_18153	521.6	286.4	303.3	321.6	248.7	250.5	713.3	586	289.8	462	463.7	244.3	798.8	831.7	328.5	539.8	385.9	330.3	0.311545106
Ga0436389_016_18150_18749	174	90.3	95.5	114	58	70	210.2	232.9	112.8	361.2	286.1	91.8	291.6	215.9	130.5	386.2	96.5	108.5	0.536635729
Ga0436389_016_19892_21154	28.5	30.2	30.2	48	35.8	35.2	32.2	36.6	32.7	21.7	44.6	24.9	46.8	31.5	26.6	43.8	33.6	25.1	0.810100052
Ga0436389_016_25417_26532	58.5	64.3	59.4	54.4	44.6	49.6	86.8	98.1	93.2	107.3	120.2	85.8	95.1	84.2	85.6	204.4	72.4	113.8	0.122012313
Ga0436389_016_9559_11370	187.6	275	285.2	182	218.6	199.1	258.6	247.6	400.8	316.5	174.3	344.4	225.2	196.9	403	236.3	415.5	520.8	0.443914053
Ga0436389_017_1370_2116	31.2	82.3	64.6	42.1	75.4	78.1	234.1	167.7	188.7	87	44.4	195.4	336.4	475.6	195	39.4	168	249.4	0.042972213
Ga0436389_017_13730_14719	534.1	498.7	462.1	412.5	429.2	341.8	425.6	421.4	447.7	298.6	376.7	454.1	314.9	350.5	501.2	714.7	310.3	328.2	0.828926114
Ga0436389_017_15679_16260	262.3	572.3	554.8	120.9	775.3	770.1	502	392.3	894.6	477.2	299.2	814.1	838.3	730.1	715.1	538.2	599.8	768.9	0.80722407
Ga0436389_017_17466_19424	175	560	544.4	257.9	613.3	599.1	441.7	457.2	868.2	266.2	261.9	695.9	370.9	330.1	981.9	360.6	902.5	946.9	0.795772104
Ga0436389_017_19761_20240	586.9	388	374.6	661	364.8	322.1	614.9	561.7	268.3	422.3	501.7	370.8	347.1	382.4	263.6	334.8	197.6	130.8	0.341654089
Ga0436389_017_9594_12428	1469	1033	964.7	1572.2	1401.5	1279.1	2266.2	2249.6	1646.2	2306.2	1743.6	1094.7	2654.8	2278.4	1882.7	3112.9	1460.6	1189.8	0.351875919
Ga0436389_018_1_504	293.9	348.9	368.3	39.3	118.3	89.2	152.6	166.9	188.1	143.1	147.8	121.7	66.8	79.6	43.3	190.5	71.5	169	0.000354976
Ga0436389_019_10490_11986	324.6	511.4	515.7	228.5	368.8	385.5	533.6	384.4	690.6	252.5	257.6	560.8	357.2	336.4	679.8	1194.9	736.4	859.6	0.038351239
Ga0436389_019_13170_13655	94.9	122.5	127.9	134.4	102.2	113.7	380.2	134.6	85.9	143.2	61	69.9	106.7	132.6	92.4	87.9	83.1	135.8	0.611163452
Ga0436389_019_13652_13888	261.8	109.9	125	141.6	67.7	73.9	92.8	114.6	95.9	100.4	123.3	59.8	68.9	70.6	144.1	109.9	146.3	132.6	0.493714508
Ga0436389_019_16179_17207	245.1	627.3	611.7	212.3	405.2	366.4	181.4	174.8	434.7	122.7	145.8	355.6	176.9	159.4	584.7	128.8	374.3	228.8	0.556639881
Ga0436389_019_17204_19351	92.7	88.7	77	72.5	89.1	74.8	144.6	138.9	135	55.1	75.1	103.5	159	191.5	59.7	93.6	105	129.9	0.194499817
Ga0436389_019_2836_4215	918.9	1142.7	1149	1915.5	1575.1	1494.6	1920	1689.7	1704.5	1390.1	1374.5	1331	4066.6	2426.9	1905.6	1514.4	1883.7	1611	0.054350005
Ga0436389_019_5093_5488	591.5	807.4	771.2	198.7	397.9	400.3	259.9	194.4	375.2	216.9	214.9	306.5	270.6	251.7	436.9	202	421	353.5	0.006886424

Table B.21 continued from previous page

Protein Accession	NoN1	NoN2	NoN3	NO$_3^-$1	NO$_3^-$2	NO$_3^-$3	N1	N2	N3	R1	R2	R3	P1	P2	P3	E1	E2	E3	Q Value
Ga0436389_019_5648_6694	2112.5	1584.2	1517	1965.5	1520.2	1439.2	1694.3	2065.1	1524.4	2034	2038.8	1407.3	1604.6	1777.3	1620.7	2480.4	1396.8	1148.7	0.993376593
Ga0436389_020_13321_14094	21.1	260.2	252.2	20.3	198.4	191.7	97	127.7	459.3	229.1	116.8	341.7	109.3	61.6	372.3	257.6	490.3	594.8	0.382330624
Ga0436389_020_17589_18644	189	259.8	268.7	96.9	204.8	191.4	204.1	202.4	165.8	113.9	148.8	203.6	162.5	133.8	199.9	474.3	154.8	120.7	0.774858409
Ga0436389_020_23707_25512	79.3	266.1	249.4	52.8	210.9	180.8	76.6	62.9	209	52.1	53.8	144.8	47.8	42.1	176.8	207.6	148.3	138	0.567426481
Ga0436389_020_25561_26364	712.8	818.1	838.5	418	447.1	460.2	461.8	455.3	563.4	520.9	508.9	561.4	410	316.3	779.6	618.6	447.5	421.9	0.096385027
Ga0436389_020_27372_28523	117	40.3	34.3	101.4	23.2	16.9	211.5	269.6	77	253.6	192	66.8	208.1	150.6	91.8	426.8	63.8	106.4	0.536635729
Ga0436389_020_28999_29982	46.1	68.8	67	39.9	59.1	63	153.7	141.9	56.3	42.5	33.5	67.7	372.8	311.4	63.2	37.2	83	83.7	0.095975137
Ga0436389_020_34055_36073	6447	4030	3941	7043	5346.1	5051.3	6226.3	6884.6	4311.8	5443.8	5771	3963.9	4869.6	4638.1	2633.7	5835.8	3283.1	3375.5	0.547639566
Ga0436389_020_36149_37396	95.4	549	517.2	97.3	587.2	521.8	111.1	96.3	667.3	85.7	112.1	569.1	158.4	133.5	306.2	85.8	510.4	587.8	0.922915439
Ga0436389_020_37796_38041	2932.3	1908.4	1704.3	4229.1	2751.8	2822	2178.8	2458.9	1227.5	2557.7	3392.4	1868.5	1721.4	1783.8	975.8	2510.8	1883.7	1923.5	0.164547025
Ga0436389_020_38119_40017	502.5	276.8	265.7	316	205.8	220.4	631	566.2	736.9	503.5	379.5	495.6	611.4	544	638.7	1406.2	635.8	975.8	0.016250918
Ga0436389_020_41281_42729	237.7	236.1	255.6	159.5	179.9	162.2	909	849.2	419.9	446.5	336.1	324.3	1149.9	944.1	1126.5	573.8	394.7	286.5	0.000558886
Ga0436389_020_44896_45663	68	147.1	170.6	84	115.6	113.1	181.2	179.5	896.7	206.2	148	507.3	115.3	76.5	166.9	127.4	487.6	945.9	0.420409164
Ga0436389_020_45810_46403	491.7	514	588.4	72.5	61.2	45.5	58.9	60.1	54.3	65.3	89	67.4	41.4	47	55.9	34.4	46.1	32.4	1.68921E-08
Ga0436389_020_48116_48628	714.2	319.8	271	206.5	131	123.2	370.8	438.9	200.6	497.4	403.9	151.8	518.4	501.6	211.2	726	215.5	177.9	0.703617122
Ga0436389_020_4902_6290	179.2	364	355.2	286.8	564.1	500.8	361.1	409.3	545.6	383.8	295.4	416.2	352.4	325.7	508.7	462.5	564.6	411.8	0.486202764
Ga0436389_021_1_360	226	769	696.6	75.2	251.2	229.6	142.5	110.5	251.3	75.7	66.3	229.8	153	138.2	248.5	83.5	234.9	195	0.067980435
Ga0436389_022_100421_100738	77.1	62.5	50.6	77.9	48.6	46	64.2	50.9	39.6	18.4	33.9	22.1	31.8	21.4	19.7	162.9	176.9	181.3	3.80939E-06
Ga0436389_022_103393_104133	130.9	136.8	129.2	87.6	94.7	84.2	140.5	135.3	204.8	117.1	66.8	186.7	234.4	158	196.6	168.6	126.8	226.3	0.139167839
Ga0436389_022_104201_104953	442.1	34	34	146.8	15.4	15.9	188.6	174.3	12.7	148	129.4	17.8	78.9	76.8	65.5	122.6	15.4	12.6	0.858546293
Ga0436389_022_123591_124127	251.7	169.9	159.7	344	524.7	572.7	231.2	280.7	280.2	293.5	368.9	387.4	228	204.8	788.7	403.3	236.9	172.7	0.444105459
Ga0436389_022_124148_124717	267.9	317.4	304.9	446.9	576.2	534.6	409.4	404.2	310.3	466.5	422.2	348	242.5	280.1	219.5	527.4	316.7	234.6	0.052604149
Ga0436389_022_132852_134576	215.7	292.3	307.6	241.9	302.5	345.2	404.6	383.9	341.3	535.4	426.6	323.3	665.1	554.8	331	813	546.7	609.7	0.029633032
Ga0436389_022_135558_136343	1280.4	1727.6	1738.7	1350.1	1813.7	2040.5	1261.5	1246.4	1394.2	1149.2	1019	1696.6	2134.7	1662.6	1290.3	1186.8	1281.6	1336.2	0.361834806
Ga0436389_022_136489_136995	840.3	465	477.7	1514.4	890.5	1072.7	980.7	839.1	616.5	1119.3	1304.8	690.9	847.9	803.2	630.2	1134.9	671.9	633.8	0.281432556
Ga0436389_022_139886_141661	504.4	549	510.1	558.4	384.4	437.4	526.7	563.2	514.6	500.3	501.1	488.2	609.4	517.5	527.9	547.2	501	561.1	0.383579104
Ga0436389_022_1512_2375	52.5	101.5	110.7	49.4	106.1	96.6	67.6	71.5	185.1	71.4	41.3	131.2	68.1	81.1	169.2	74.8	147.8	206.1	0.789529004
Ga0436389_022_157260_160280	27.6	17.4	17.8	51.6	12.7	12.2	328.8	248	22.8	71.8	52.8	27.5	658.2	841.5	40.4	35.1	26.8	22.9	0.100904754
Ga0436389_022_16464_17726	152.9	616.7	662.5	112.1	400	361	373	354.7	864.8	220.5	180.3	422.8	609.6	544.7	1291.5	281.9	1867.9	1541.9	0.232085881
Ga0436389_022_177048_177530	320.8	769.5	718	205.5	463.5	404	209.5	117.6	129.8	58.2	41.3	91.5	227.1	375.4	396.8	63.6	109.9	73	0.008889556
Ga0436389_022_187481_188332	98.5	85.9	84.6	72	103.5	96	133.4	158.9	119.6	136.2	88.1	111.1	115.3	111.3	114.2	157.3	112.5	159.9	0.052412269

Table B.21 continued from previous page

Protein Accession	NoN1	NoN2	NoN3	NO$_3^-$1	NO$_3^-$2	NO$_3^-$3	N1	N2	N3	R1	R2	R3	P1	P2	P3	E1	E2	E3	Q Value
Ga0436389_022_205356_205733	345.3	403.6	380.1	169.4	165.6	166.5	70.4	67.8	144.3	93.3	71.5	142.2	87.7	82.3	119.9	72.3	115.4	146	4.05456E-05
Ga0436389_022_220013_220420	427.7	773.3	743.5	244.2	475.4	536.2	277.9	270.7	370.9	275.3	270.9	409.4	277.4	276.7	297	296.5	291.7	307.5	0.03984082
Ga0436389_022_223953_226880	240.5	473.7	463	256.3	305.8	300.6	511.9	359.5	594.5	251.1	322.7	445.4	523.8	559.9	739.4	395.9	552.1	498.4	0.071433399
Ga0436389_022_22543_23955	1137.4	1335.6	1268.4	976.9	1102.9	1047.1	1183.5	1171.8	1140.8	1176	1133.9	1145.8	1587.8	1295.2	1339.5	1475.8	1101.9	1144.7	0.108217112
Ga0436389_022_227772_229340	313.3	389.1	380.3	490.6	482	500.4	623.9	583.2	631.2	519.2	419.2	490.6	935	792.6	789.5	679.2	413.3	355.8	0.002919058
Ga0436389_022_230442_230735	56.9	101.9	91.7	34.2	105.3	105.4	172.5	161	130.8	95.7	86.8	111.8	257.6	206.4	255.9	117.6	119.1	102.3	0.000795241
Ga0436389_022_235409_235606	600.7	395.7	320.2	403.9	354.5	338.8	600.1	473.8	328.1	358.5	347.9	175.4	997.2	764.8	463.6	347.6	213.2	164.4	0.058925315
Ga0436389_022_23956_24897	428	280.9	286.8	474.1	262	265.3	431.9	439.3	218.9	377.3	430.8	221.8	572.4	453.6	229.7	434.4	261.3	213	0.933159492
Ga0436389_022_243328_244350	139.1	212.4	196.4	283.7	124.5	151.6	180.8	214.5	161.2	248.4	287.1	219.3	136.8	169.3	106.2	237.3	178.8	196.7	0.285676368
Ga0436389_022_257979_258380	200.9	168.8	131.9	184.9	200.5	193.5	141.9	154.4	135.6	171.4	174.9	195.9	123.2	160.1	210	219.6	157.8	147.5	0.593279349
Ga0436389_022_258493_259305	3975.7	3627.4	3440	5370.3	5836.5	5828.6	3247.3	3896.8	3214.1	2692.4	4308.2	4045.5	2667.3	2717.4	2110.7	2253.9	3091	2054.3	0.000875663
Ga0436389_022_259490_260659	357.2	386.7	411.6	594	759.2	557.1	564.9	620.5	593.4	355.1	446.1	531.1	278.7	259.3	511.8	379.9	533.1	511.1	0.04376215
Ga0436389_022_260726_262570	597.5	881.8	871.1	772	1213	1134	657.3	732.9	1033.3	555	600.2	1046.5	661.4	613.6	891.7	700.6	998.6	952	0.577593457
Ga0436389_022_263199_264287	24.1	514	500.4	24.2	453.9	413	26	40.8	382.9	39.1	37.7	335.3	64.4	47.1	496.8	86.9	471.2	440.5	0.84461268
Ga0436389_022_28222_29262	684.2	505.7	528.3	668	616.7	563.7	658.3	674.6	386.3	357.5	479.3	341.4	609.9	588.3	300.3	549.3	423.4	337.6	0.371975359
Ga0436389_022_29262_29672	88.8	98.3	82.9	120.3	98.6	97.2	66.1	90.4	130.2	78.2	94.8	186.6	46.8	49	84.6	61.6	129.5	189.7	0.547929064
Ga0436389_022_30192_30947	290.5	1187.7	1117.4	373.5	957	842.3	205.4	238.6	807.4	131.1	185.8	864.1	227.9	232.9	848.2	129.3	777.7	637	0.706399868
Ga0436389_022_30940_31545	525.2	703.7	676.8	297.5	509.4	457.4	369.4	361.1	399.7	434.3	418	371.1	333.2	291.8	273.8	438.5	301	233.5	0.013879759
Ga0436389_022_38803_39825	29.6	43.9	37.5	22.9	30.8	34.5	113.5	75.4	26.6	51.7	41	29.5	182.6	234.8	33.3	48.3	40.9	27.1	0.146138709
Ga0436389_022_43048_43608	290	246.8	247.7	146.7	153.9	149	200.5	149.8	89.6	189.2	189.4	105.9	548.7	125.1	104.9	221.7	110.4	66.3	0.675417179
Ga0436389_022_48384_49166	457.6	366.5	377.7	550.8	323.4	316	668.5	691.4	371.8	657.5	576.1	405.2	925.4	711.4	284.5	675.2	321.4	396.5	0.66577182
Ga0436389_022_50756_51397	267.4	387.7	395.8	155.8	246.3	229.3	148.3	106	210	46.5	106.3	239.1	44.7	62.8	123.4	37.9	187.9	203.9	0.03470805
Ga0436389_022_59508_60977	463.8	477.5	492.5	230.9	260.4	272.8	493.2	501.8	623.7	442.9	346.1	406.3	597.5	411.1	475	725.8	454.2	676.6	0.013822851
Ga0436389_022_62594_63559	122.7	207.2	187.6	135.1	266.4	177.5	146.1	135.3	248.2	122.3	99.1	193.7	76.3	62	211.7	156.9	194.2	216.7	0.689643438
Ga0436389_022_63597_64190	199.3	303.4	267.5	206.9	265	247.6	155.3	204.9	239.8	206.7	191.1	240.8	148.4	121.2	166.4	189.5	224.3	177.2	0.081110466
Ga0436389_022_6410_7840	1421.1	1478.5	1415.9	944	1002.6	959.9	1036.1	959.3	1547.1	837	881.2	1308.4	1104	1022.8	880.1	806.8	989.2	1364.1	0.247136259
Ga0436389_022_67292_68284	213.8	97.3	87.3	162.7	64.4	60.9	233.1	225.2	101.9	166	202.5	78.9	205.1	252.2	205.3	200.5	87.7	73.8	0.371975359
Ga0436389_022_77174_78718	728.8	812.2	798.7	564.7	699.5	657.3	1054.6	1032.5	927.2	910.2	698.6	707.7	1542.1	1559.3	1343	922.2	890	695.4	0.000116372
Ga0436389_022_80105_81313	69.9	99.1	96.3	276.5	207	181.6	197.7	250.5	447.1	292.9	157	213.5	143.4	75.9	726.9	343	252.6	259	0.68995434
Ga0436389_022_81418_82803	186.3	321.9	312.8	238.7	207.2	181	515.8	527.4	730.5	314	181.1	117.9	1776.6	1339.4	1676.4	300.1	880.7	370.4	0.000137732
Ga0436389_022_8440_9153	142.7	348	374.1	493.6	756.8	737.2	248	291	556	193.9	161.9	422.4	213	267.9	601.7	252.8	536.3	248.6	0.20382254

Table B.21 continued from previous page

Protein Accession	NoN1	NoN2	NoN3	NO$_3^-$1	NO$_3^-$2	NO$_3^-$3	N1	N2	N3	R1	R2	R3	P1	P2	P3	E1	E2	E3	Q Value
Ga0436389_022_86912_88222	1337.9	1484	1467.7	1263.1	1604.4	1492.3	1221.9	1369.2	1430.4	779.5	851.9	1011.8	1223.1	1170.8	1138.7	921	1303	1035.2	0.008332703
Ga0436389_022_91266_92708	13.8	393.2	366.1	13	441.1	374.6	59.6	71.4	439.7	48.1	27.7	340.7	129.8	78.5	516.3	44.6	601	546.3	0.880099815
Ga0436389_022_92705_93337	22.9	87.3	82.5	31.9	123.1	103.5	58.3	54.3	99.4	42.8	36.8	75.6	119.1	77.3	81.2	41.2	107.1	87.1	0.775560037
Ga0436389_022_93334_96531	15.8	269.8	235.5	67.2	151.3	131.7	148.5	426.1	431	188.9	177.2	299.7	72	76.1	274.8	124.1	930.5	1034.9	0.162677093
Ga0436389_022_96927_99914	87.8	55.9	66.2	59.4	31.7	30.2	89.4	87.1	29.8	111.7	99.1	38.1	87.1	78	33.8	86.2	56.1	24.6	0.710964196
Ga0436389_023_3_1748	2364.6	1492.4	1409.4	180.1	159	147.9	508.2	484.6	444	302.8	255.3	339.7	595.5	477.1	370	593.7	380.9	495.1	0.000538839
Ga0436389_023_5570_6466	79.2	787.5	784.2	14.4	184.2	175.6	34.1	47.6	364.1	28.4	20.3	304.2	32.8	29	382.1	60.6	466.9	470.1	0.37176433
Ga0436389_024_13006_13851	78.6	73.6	60.7	87.8	86	74.8	84.2	93.5	211.3	87.5	91.4	187.8	93.1	129.8	50.3	84.7	127.9	249.1	0.572054844
Ga0436389_024_15652_15972	1575.4	1159.8	1171.8	2123.9	1772.1	1856.6	1417.5	1573.2	1086.1	1293.8	1611.3	1167	1525.2	1428.8	1365.5	1206	1246.3	883.7	0.036515336
Ga0436389_024_16103_17332	13312.6	9731.4	9594.1	11317.8	9900.9	9579.8	9742.2	10858.9	7740.6	9510.3	10431.6	7795.7	9039.8	8881.3	5457.6	10580.2	7534.9	7987.7	0.468015301
Ga0436389_024_17375_19450	4787.1	5119.4	4995.5	6817.3	6322.8	5683.8	4278.4	4916.2	4646.7	4271.4	5688.3	5441.4	3833.5	3847.7	3279.5	4936.1	5742.8	4897.1	0.005942455
Ga0436389_024_19563_20033	1172.4	1078.2	1036.6	2202.3	1798.7	1559.6	1095.2	1341.7	1290.1	889.9	1176.6	1330.8	943.6	1001.4	927.5	1218.9	1032.5	958.9	0.006236576
Ga0436389_024_20096_20470	396.5	389.9	357.3	576.2	539.8	487.6	1120.3	1184.8	497.6	1427.9	1009.8	863.8	879	754.2	395.1	861.4	482.5	389.9	0.087772189
Ga0436389_024_2079_4040	541.5	1138.5	1098.6	31.9	102.1	81.6	140.9	106.5	291.3	90.8	141.5	311.9	114.2	181	537	64.9	173.1	103.7	0.005942455
Ga0436389_024_23133_27662	940.9	771	739.7	907.8	793.7	751.2	1028.1	1303.5	869.7	1129.2	1285.9	632.8	1174	1116.9	1061	1274.7	995.8	900.1	0.444105459
Ga0436389_024_32741_35002	602	1128.1	1126.6	1486.1	1142.4	1260.2	1761	1526.5	1417.1	1144.8	1016.1	1165.3	859.5	512.7	837.2	1338.2	2030.7	1488.9	0.019454238
Ga0436389_024_35131_37245	577.6	524.6	527	1692.7	494.7	608.5	705	630.9	644.7	523.7	688.9	640.1	946.8	698.5	492.5	695.8	761.7	739.8	0.775560037
Ga0436389_024_37846_43668	8.6	103.2	100	3.5	47.2	38.2	439.4	420	415.5	110.6	50.2	173.1	1567.1	1200.5	958.1	109.7	320.2	190.7	5.55925E-05
Ga0436389_024_4925_5278	51.6	72	87.9	60.1	100.1	44.2	154.6	120.1	86.3	75.7	48.6	56.4	113	143	87.4	88.6	65.2	75.4	0.096385027
Ga0436389_024_5314_6540	278.3	219.8	223.9	269.3	198.7	163.5	410.5	368.9	326.3	297.8	306.3	221.6	807.6	499.2	391.9	297.4	336.7	360.2	0.035079555
Ga0436389_024_63385_63681	642.9	1119.9	1025.6	611.5	2474.6	2302.5	849.3	835	759.2	747.1	740.3	1157.8	392.4	443.9	628.8	1099.4	1065.8	625.9	0.165891133
Ga0436389_024_63690_64097	228.1	217.9	288	312.1	270.9	242.1	390.1	380.2	262.5	250.5	281.7	211.7	931.7	790.1	287.2	312.2	341.1	306.9	0.079331393
Ga0436389_024_64094_64546	300.7	233.9	231.6	369.8	372.7	376.2	803.3	845.8	302.3	935.3	662.7	306.8	825.7	592.4	241.9	1280.2	295.4	273.9	0.703617122
Ga0436389_024_65535_65885	594.4	281.3	283.5	686.7	387.9	414.6	760.5	821.9	682.8	511.7	747	477.2	951.2	954.5	928	510	715.5	649	0.014911347
Ga0436389_024_65932_66870	2345.9	3235.7	3010.3	2514.3	4495.6	4375.3	1988.1	2053.8	2953.6	1668.5	2172	2617.6	1535.3	1457	4058.7	1687.9	2946.2	1673.8	0.376124969
Ga0436389_024_66953_67345	144.5	178.7	162.5	184.8	245	253.1	82.4	103.3	262.4	93.2	122.2	278.6	60.8	53.9	191.9	111.5	270.5	277.2	0.560457344
Ga0436389_024_67517_67819	996.1	694.7	675.7	1228.8	896.4	1102.2	1207.7	1190	823.9	754.6	872.8	757.7	1241.2	1567.8	1085.9	693.5	1110.9	842.4	0.107104498
Ga0436389_024_68106_68657	1001.9	450.9	412.7	701.8	341.1	344.4	1235.2	1381.1	571.6	914.2	912.9	382.2	1906.2	1489.2	657.3	639.7	616.4	506.9	0.200619907
Ga0436389_024_70109_70564	461.7	222.3	232.4	885.6	425.1	433.2	423.1	437.1	236.9	315.2	462.5	294.9	165.8	302.9	171.6	275.7	296.3	245.9	0.189991068
Ga0436389_024_70573_71238	1031.4	810.7	761.6	1517.7	1360.8	1216.4	1136.7	1088.5	964.2	1030.1	1159.9	969	1007.8	1150.8	624	1077.6	783.7	704.7	0.081099435
Ga0436389_024_71253_71618	550.4	382.5	396	531.5	561.3	571.5	1049.6	991.2	1379.6	921.1	721.3	986.4	1022.9	869.4	671.5	882.1	945	1426.1	0.014478567

Table B.21 continued from previous page

Protein Accession	NoN1	NoN2	NoN3	NO_3^-1	NO_3^-2	NO_3^-3	N1	N2	N3	R1	R2	R3	P1	P2	P3	E1	E2	E3	Q Value
Ga0436389_024_71651_72190	813.4	802.8	746	1103.3	1301.6	1218.7	1752.8	1659.2	1998.6	962.1	955.6	1314.8	2981.9	3047.1	2281.4	857	1855.5	1819.7	0.001184822
Ga0436389_024_72207_72581	611.1	726.3	639	915.8	1162.8	1121.7	485.2	533.4	567.7	420.6	691	821.1	350.1	429.3	443.7	349.9	677.4	438	0.004776308
Ga0436389_024_72629_73168	1467.7	1526.1	1409	2350	2621.9	2280.2	1409.7	1504.6	2071	1594.3	1972.5	2291.4	951.2	1188.3	2294.5	1095.6	2381.9	1964.9	0.30826916
Ga0436389_024_73291_73647	817.6	843.6	797.6	907.1	1617.5	1625.1	788.4	744.2	1117.3	567.3	793.3	1063.1	638.6	882.6	755.4	576.9	1510.9	1356.7	0.27289012
Ga0436389_024_73649_74014	1455.8	1389	1378.7	2491.9	2273.7	2343.6	2873	3050.8	2299.5	1760.7	1644.5	2173.8	2486.2	2104.6	1378.5	1277.4	1842.7	1925.5	0.025134825
Ga0436389_024_74025_74276	201	382.7	392.9	163.8	578.9	589	547.5	427	598.7	136.1	183	378.4	758.3	983.4	1036.4	156.8	591.8	454.9	0.02169711
Ga0436389_024_74294_74506	284.8	216.1	204.3	367.4	543.7	478.6	413.5	391.6	199.1	505.8	488.9	255.6	216.2	219	161.2	547.3	210.1	117.5	0.281752357
Ga0436389_024_74509_74919	576.6	584.7	520.8	899.8	823.2	831.8	275	405.3	670.8	247.4	490.9	641.6	240.5	214.8	318.2	333.9	522.1	584.7	0.019132506
Ga0436389_024_74984_75712	1735.3	2168	2053.4	2311.1	3391	3210	1808	2080.4	2579.8	1972.4	2291.9	2578.8	1911.1	1773	2128	2949	2865.4	2450.7	0.055817155
Ga0436389_024_75732_76091	260.9	171.2	188.3	490.4	246.9	230	460.5	412.7	157.7	279.9	240.8	148.9	794.5	730.7	239.3	330	182.1	101.4	0.200619907
Ga0436389_024_76096_76371	740.4	711.7	685.8	898.4	1230.9	1213.1	563.1	518.6	769.9	442.1	611.8	868.1	324.3	357.8	527.6	550.8	919.8	799.6	0.018780364
Ga0436389_024_76407_77270	1352.3	1134.1	1043.7	1913.1	1691.5	1463.5	1699.7	1804.6	1545.1	1157	1709.4	1476.5	2504.3	2384.5	2469.5	1488.1	1980.9	1402.1	0.00283184
Ga0436389_024_77286_77564	465.8	1196.9	1144.7	499.2	1252.3	1234	437.1	448.3	906.7	453.5	457.8	1376.5	443	404.6	1512	450.7	1034.8	900.3	0.93362638
Ga0436389_024_77581_78216	1691.7	1105	1067.1	2677.3	1808.2	1878.3	1629.1	1883.2	1532.1	1808.4	2051.8	1547.4	1110.7	1232.8	1339.2	1762.7	1588.4	1264	0.078411999
Ga0436389_024_78228_78842	714.9	727.4	739.1	982.5	1152.3	1077.3	937.8	927.3	859	793	828.3	854.6	1068.2	990.8	1054.5	803.5	1093.3	834.2	0.010905627
Ga0436389_024_79288_79749	87.7	226.5	230.8	82.4	129.3	122.4	75.7	80.3	214.6	99.7	91.1	250.6	55.9	73.8	92.9	94.8	359.7	393.2	0.287145187
Ga0436389_024_9738_11801	950.1	722.2	740	1039.2	701.8	688.8	1665.2	960.4	871.8	931	947.8	1014.5	787.4	712.2	699.6	810.3	1105.8	1194.4	0.351873051
Ga0436389_025_102808_104676	4.7	15.9	17.6	1.2	9.9	7.3	78.7	104.4	92.9	174.4	71	19.4	61.8	37.1	219.5	233.1	81.4	41.5	0.371975359
Ga0436389_025_110047_111048	262.2	333	340.9	184.9	316.2	300.2	186	263	478.3	203.9	221.1	317.3	167.2	161.7	437.9	263.3	399.2	344.6	0.905781503
Ga0436389_025_13183_13512	230.4	448.3	387.5	187.6	427.1	449.2	165	225	288.7	146.8	214.2	414.6	177.3	168.1	214.2	216.1	238.3	187.5	0.365950509
Ga0436389_025_19236_20126	456.1	199.7	187.2	312.2	121.1	115.5	459	547.1	134.5	624.9	523.3	146.7	538.3	396	119.3	715.8	105.1	95.3	0.880099815
Ga0436389_025_20_223	130.3	28.2	32.6	363.8	34.9	41.8	114.5	122.5	16	113.2	179.5	16.4	149	114.3	22.2	136.7	14.8	32.1	0.931326766
Ga0436389_025_32929_34473	245.1	967.3	962	188	952.5	882.3	348.4	352	1020.6	209.6	224.3	1005.4	415.8	313.1	757.4	326.4	947.6	868.1	0.958790764
Ga0436389_025_37123_38106	30.6	57.3	50.1	33	38.2	51.8	138.3	142.3	41.4	155.8	105.9	35.5	95.3	56.6	32.5	120.7	25.5	19.1	0.50228876
Ga0436389_025_41503_49626	75.2	38.9	34.5	15.1	20.1	17.8	84.4	73.6	74.6	68.3	65.5	54.1	176.1	106.6	108.6	353.1	63.2	77.5	0.312958974
Ga0436389_025_49774_50304	59.3	56.5	52.6	24.9	26.1	23.3	55.2	44.1	42.2	55.2	38.6	23.4	34.7	45.4	45.1	69.3	110.7	102.1	0.00283184
Ga0436389_025_64356_65666	152.5	203.1	158.8	80.2	178	145.4	361.4	273.6	149.4	135.9	98	119.6	205.9	189.2	108.9	105.4	94.4	86.4	0.096312398
Ga0436389_025_67199_68260	3319.7	2506.6	2443.8	2730.7	2500.6	2185	2297.3	2558.8	1872.7	2335.8	2484.7	1943	2169.5	2236.1	1611.9	2182.9	1689.9	1699.9	0.173110619
Ga0436389_025_70546_71295	241.3	316.2	281	194.2	269.9	242.7	187.7	178.6	724.9	153.1	154.8	395.7	176.5	160.1	374.4	195.6	385.1	836.9	0.748339619
Ga0436389_025_71305_72870	13.9	85.8	97	11.2	91.5	83	379.8	270.2	175	16.5	12.2	79.1	98.3	99	228.2	17.1	94.4	29	0.02042856
Ga0436389_025_72984_74609	131.3	367.8	323.5	120.2	378.6	376.1	546.1	419	580.7	475.8	301.6	417	201.6	179.3	582.4	352.7	419.5	452.4	0.444744984

Table B.21 continued from previous page

Protein Accession	NoN1	NoN2	NoN3	NO$_3^-$1	NO$_3^-$2	NO$_3^-$3	N1	N2	N3	R1	R2	R3	P1	P2	P3	E1	E2	E3	Q Value
Ga0436389_025_80852_81733	60.2	97.1	123.5	48	68.9	77.5	57.7	57.2	116.7	50.3	38.3	108.3	71	65.2	38	48.5	107.5	147.3	0.68629898
Ga0436389_025_82113_83015	84.5	139.6	130.8	102.6	191.3	181.3	309.7	277.9	209.5	184.3	152.5	244.1	564.1	459.9	193.8	245.9	154	150.1	0.068567799
Ga0436389_025_87990_89288	103.4	205.1	198	52.5	88.8	77.4	204.3	189.6	132.7	187.4	142.7	120.8	296.7	329.6	134.7	195.5	102.1	102.5	0.119196104
Ga0436389_025_89356_90879	1021.4	1086.8	1048.1	444.1	682.9	683.3	1688.6	1091.9	754.7	852.8	835.8	732.7	1416.6	1312.7	1017.8	904.4	914	817.2	0.087206177
Ga0436389_025_91047_92216	6256.9	5092.4	4986.4	4713.5	4571.7	3983.5	4291.4	4498.4	2850.3	3573	3834.5	2954.6	3832.4	3573.7	2588.5	3209.8	2916.9	2363.7	0.017384523
Ga0436389_025_95217_95657	2868.9	2616.1	2580.8	1798.9	2238.3	2240.3	1562.4	1475.1	1222	1977.3	2119.5	1372.4	1402.5	1689.7	1545.7	1467.4	1348.4	1072	0.001541807
Ga0436389_025_95665_98376	45.2	278.3	294.1	25.5	202.4	194.8	23.1	11.8	610.5	32.5	16.2	543.4	13.3	24.8	407.3	39.1	488.5	814.9	0.833821868
Ga0436389_025_98373_99626	121	422.3	396.7	71.1	218	213.8	74.8	83.8	106.3	98.1	77.4	116.2	44.4	67.6	89.4	58.1	95.7	63.2	0.054350005
Ga0436389_026_100307_100966	2943.7	3002.4	2757.5	1820.3	2597.1	2390	1989.9	2063.5	1792.2	2026.2	2147.3	2029.8	1584.4	1820	1691.8	1818.7	1832.7	1672.2	0.001136206
Ga0436389_026_105759_107528	260	613.8	606.2	171.1	414.8	390	151.2	160.1	453.2	124.3	125.3	351.6	127.5	119.5	373.3	151.2	782.9	766.1	0.35556729
Ga0436389_026_107642_109807	62.1	235	217.7	86	273.9	237.7	43.2	53	259.7	32.3	30.9	213.2	12.8	19.2	291.7	8.1	185.5	169.8	0.897521283
Ga0436389_026_111252_112634	420.5	1084.7	1102.7	1171.8	1979.5	1931.7	600.6	672.4	1678.5	473.3	515	1616.7	587.9	466.4	1179.3	517.5	1620.6	1510.8	0.472367616
Ga0436389_026_113845_115653	336	659.4	696.3	897	1470.5	1175.7	919.1	942.5	1497.5	881.5	633.6	1021.2	1058	711.3	971.5	1253.4	1751.2	2090.6	0.031800823
Ga0436389_026_116096_116584	460.7	398	381	402.2	648.3	623.1	374.6	251.6	605.5	270.2	281.4	346.6	162.4	258.7	725.4	253.5	333.4	236.9	0.50228876
Ga0436389_026_116581_117309	143.6	101.7	104.8	327.1	119	107.7	464.7	517.8	304.7	216.1	140.5	87.6	610.5	612.2	551	216.7	325.5	169.6	0.001077545
Ga0436389_026_119591_120274	923.2	866	771.6	746.8	954.9	825.9	732.1	739.7	633.3	843.4	833	706.4	699.4	670.2	412.1	531.6	657.1	690.9	0.077810335
Ga0436389_026_11964_14078	165.3	316.4	291.3	87.5	230.9	234	341	262.6	305.3	157	155.6	301.5	909.4	543.3	505.4	122.7	321.2	342.4	0.023400743
Ga0436389_026_120479_121483	5887.9	5437.1	5393	6320	6191.2	5378.7	3970.9	4364	3657	2720.3	4450	4134.5	2618.6	2765.8	2920.2	2582.9	3381.4	2446.1	0.000389029
Ga0436389_026_122976_124271	1434	1964.6	1996.7	1112.7	1781.1	1623.6	878.8	958.6	1595.1	979.2	1159.6	1812	596.3	526.5	1326.7	750.6	1655.4	1606.4	0.293676867
Ga0436389_026_124384_125802	450.9	898.3	796.5	255.4	1013.3	876.3	310.9	273.6	877.9	255.7	356.8	1214.2	306	344.3	580.9	280.8	953	1227.8	0.84461268
Ga0436389_026_125799_126530	109.5	644.1	645	87.5	568.6	449.6	202.9	194.4	714.9	353.1	296.2	703.3	292.9	195.8	431.2	655.5	528.6	766.6	0.68995434
Ga0436389_026_132651_134327	1999.9	1613	1572	1598.1	1138.4	1109.9	1404.8	1377.3	918.1	1130.8	1461.3	1237.5	1262.2	1155.7	744.1	1015.6	785.2	613.4	0.045470018
Ga0436389_026_135419_135991	224.6	311.4	311.6	176.7	255.6	246.4	167	150.5	272.2	135.1	209.5	338.6	94.8	71.5	335.7	289.5	359.2	326.6	0.435986526
Ga0436389_026_139219_139575	45.2	57.9	48.6	53.7	71.9	79.4	52.8	66	26.2	40.2	48	33	263.7	131.4	21.7	32.9	24.5	26.5	0.311545106
Ga0436389_026_142026_143171	60.7	44.2	34.5	36.8	41.1	25.6	231.4	286.3	214.4	550.5	332.8	87	275.8	157.8	44.4	528	102.5	198.1	0.232085881
Ga0436389_026_143194_143910	153.4	148.9	136.1	218.8	173.6	137.5	310.4	275.3	325.6	191.1	209.6	185.7	220.8	261	530.7	193.9	340.9	256.3	0.148860042
Ga0436389_026_144615_145748	75.3	94.6	87.2	81.6	102.6	95	145.2	95.5	125	102.8	80.2	151.7	167.6	191.1	102.4	70.6	124.9	147.8	0.331409887
Ga0436389_026_145745_147295	42	150.8	161.4	24.1	119.2	119.1	84.6	66.8	208.2	51.4	53.3	147.2	157.4	203.4	307.5	66	206.8	179.6	0.349746955
Ga0436389_026_149931_150245	307.4	190.7	181.5	279.5	231.7	258.7	244.2	258.9	115.1	161.6	217	141.8	361.2	358.1	164.7	208.8	267.8	249.2	0.527500481
Ga0436389_026_152298_152519	622.5	226.2	220.7	961.7	392.5	375.5	631.3	696.2	270	565.6	658.1	202.3	2858.3	1232	287.3	429.9	262.4	122.9	0.357816641
Ga0436389_026_152609_152803	183.9	98.4	89.8	165.2	75.2	88.1	230.1	177.9	72.9	158.5	162.4	55.7	624.1	386.3	136.3	146.1	89.4	59.4	0.162677093

Table B.21 continued from previous page

Protein Accession	NoN1	NoN2	NoN3	NO$_3^{-}$1	NO$_3^{-}$2	NO$_3^{-}$3	N1	N2	N3	R1	R2	R3	P1	P2	P3	E1	E2	E3	Q Value
Ga0436389_026_152916_155372	157.1	573.5	546.5	143.7	567.1	556.6	233.8	214.6	761.5	301.1	204.1	818.7	283.1	253.3	762.5	418.3	724.3	789	0.921611604
Ga0436389_026_156862_157356	542.9	395.4	377.8	740.6	424	437.4	730.5	747.7	423.8	545.7	657.6	472.4	660.3	683.6	262.2	561.5	312.8	314.3	0.675417179
Ga0436389_026_170827_171744	2546.7	1923.4	1898.2	906.7	676	627	1363.1	1226.3	900.2	1065.6	883.1	782.6	1544.6	1273.1	870.3	1271.9	827.8	827	0.004514614
Ga0436389_026_175590_176126	2995.7	2301	2197.4	3191.5	3488.1	3997.1	2669.1	2830	2011.7	2267.1	3244.4	2470.4	1965.3	2139.6	1971.4	2662.4	2482	2259.4	0.034625813
Ga0436389_026_179858_180349	1118.5	533.8	504.5	557.7	371.3	385.9	715.3	631.2	442.4	380.9	555.5	358.9	471.3	711.5	349.3	375.6	454.3	411.7	0.476947397
Ga0436389_026_180930_181877	1050.8	1447.3	1333.1	736.8	1313	1263.5	644.5	651.1	1160.2	736.5	582.5	782.5	778.4	781.6	1658.5	1048.3	1370.3	1097.6	0.365681504
Ga0436389_026_183265_183795	636.1	1223.8	1317.1	491.3	1219.4	1186.3	443.4	417.4	1260	409.5	418.5	978.9	308.7	448.5	1432.4	280.2	1234.6	1092.8	0.880099815
Ga0436389_026_183872_184561	166	114.7	106.4	177.7	131.5	118.3	184	197.7	224.5	202.4	202.8	139.2	201.9	146.8	455.4	220.6	145.5	104.1	0.46266519
Ga0436389_026_185765_186919	92.8	47.8	49.7	108	32	35.9	76.7	82	96.4	82.3	84.8	62.1	123.7	176.4	157.3	147	153.9	149.1	0.006236576
Ga0436389_026_187240_189861	187.4	466.4	463.2	94.7	312.9	298.2	534.4	472.6	927.9	442.1	291.2	795.7	255.8	131.3	305.6	621.5	478.7	1108.2	0.153386574
Ga0436389_026_190725_191837	166.7	156.2	143.6	141.1	131.7	118.8	279.1	320.5	376.6	746.4	475.8	397.1	550.7	334.5	284	1205.4	285.5	513.6	0.153386574
Ga0436389_026_191857_193458	182	231.8	213.1	128.3	210	204.6	193.7	154.7	476.3	135.8	165.4	664.9	246.4	223.8	139.7	166.7	290.8	628.2	0.833821868
Ga0436389_026_195460_197208	75.3	124.5	117.4	71.8	92	79.1	149.9	96.4	80.1	113.2	90.7	71.9	92	104.8	89.4	112.9	58.4	79.4	0.753132186
Ga0436389_026_198203_199267	136.7	193.9	184.6	114.5	148	147.1	190.8	216.4	160.8	176	149.7	174.4	271	206.4	204	228.5	192.4	123.3	0.200619907
Ga0436389_026_199347_201056	18.5	23.3	22.8	11.5	23.1	23	89.7	118.1	32.1	235.9	113.3	29.1	122.8	58.2	69.9	312.2	54.4	40.8	0.522051737
Ga0436389_026_214652_215092	443.4	231.3	268.2	366.2	304.8	279.8	422.9	350.3	206.8	272.2	381.8	195.2	226.7	297.2	165.4	265.5	229.8	113.2	0.58538959
Ga0436389_026_21467_21841	2144.4	1776.2	1726.8	239.3	287	317.8	344.2	243.9	260.8	285.9	303.4	265.6	222.3	258.4	304.5	248.2	215.6	162.3	2.09994E-07
Ga0436389_026_215142_216113	299.3	635.3	584.5	242.7	380.6	350.6	217.5	227.8	529.6	143.5	207.3	473.5	166.2	219.3	397.6	138.9	468.1	411.9	0.612302307
Ga0436389_026_217023_217424	1.6	107.6	86.7	0.8	91.5	97.7	174.6	153.8	72.4	202.4	90.4	46.9	110.9	58.4	94.4	235.8	77.5	78.9	0.727791178
Ga0436389_026_218461_220602	96	502.3	475.9	122.6	418.7	388.2	131.4	129.1	355.3	131.4	125.3	360.6	198.6	170.4	215.9	267.8	350.1	322.1	0.703617122
Ga0436389_026_21889_23265	273.5	100.5	91.9	233.8	89	55.6	341.5	319	103.5	554.5	396.5	131	487.5	393	119.3	447.3	125.1	139.7	0.584198868
Ga0436389_026_222690_223499	112.1	88.9	87.4	97.5	89	97.3	353.8	424	295.4	1139	683.2	271.5	1106.1	421.8	260.9	2031.7	231.2	739.1	0.312877733
Ga0436389_026_224151_225065	59.3	22.1	23.7	56.7	19.1	28.1	61.3	83.8	16	38.3	51.6	16.2	104.3	106.7	18.1	35.8	7.9	3.3	0.408256007
Ga0436389_026_225107_226294	57.4	325.1	336.2	36.2	283.8	265.2	23.2	57.5	660.2	47.8	33.6	380.5	31.8	36.8	551.3	37.7	613.7	760.9	0.842741845
Ga0436389_026_226637_228556	669.9	1061.3	1004.2	714.6	1187.7	1166.8	1048.4	1215.5	1227.8	847.7	862.8	981.7	893.4	1157.3	722.4	751.9	1048.5	918.5	0.600924728
Ga0436389_026_232726_232962	336.1	430.5	396.8	371	650.2	652.1	461.3	456.5	464	404.6	408	404.9	480.6	376.1	564.3	632.1	392.6	269.1	0.624931824
Ga0436389_026_236904_238823	175.7	256.2	256.8	283.7	393.8	364.5	343.4	346.5	420.9	263.3	280.5	323.2	651.6	480.4	536.8	334.6	753.4	674.1	0.029037399
Ga0436389_026_238926_240527	60.6	90.9	73.3	86.2	63.2	73	129.5	237.5	577.5	355.9	200.3	382.6	157.3	91.5	508.2	411	297.1	534.7	0.16721543
Ga0436389_026_24196_25542	49.8	238.6	232.4	36.6	220.3	218.7	26.3	39	202	27.5	28.1	217	33.6	22.6	171.1	18	176.1	159	0.84461268
Ga0436389_026_242104_242433	90.2	157.1	129.6	28.1	64.4	57	30.8	24.8	22.1	22.5	23.5	44.3	9.4	19.3	15.4	40.7	21.1	20.7	0.000980459
Ga0436389_026_242511_244277	1524.6	1800.5	1843.3	1786.4	2136.5	1908.9	1616.1	1735.6	1913.8	1738.8	1736.1	1857.9	2303.1	1725	2173.9	1704.4	1839.9	1612.8	0.304231937

Table B.21 continued from previous page

Protein Accession	NoN1	NoN2	NoN3	NO_3^-1	NO_3^-2	NO_3^-3	N1	N2	N3	R1	R2	R3	P1	P2	P3	E1	E2	E3	Q Value
Ga0436389_026_246109_248088	874.6	835.9	787.4	1155.6	930.1	854.8	5646.5	7941.2	5995.7	6917.1	4385.6	3871.9	3981.6	2592.1	1186.2	5520.3	2663.4	5238.8	0.003551791
Ga0436389_026_248155_248745	1206.8	775.2	762.2	838	728.1	704.8	1012.5	912	405.1	636.6	782	441.9	805.4	887.3	396.1	651.1	366.8	236.6	0.37650427
Ga0436389_026_248841_249536	544.6	314.8	300.4	847.8	449.6	442	723.6	876.3	431.3	478.1	588	448.9	792.3	722.7	244.6	504.1	389.7	424.7	0.626719426
Ga0436389_026_252408_253691	135.6	111.9	115.2	230.1	112.6	122.7	318.7	397.3	78.6	317.6	265	95.9	259.4	219.9	79.5	412.9	78.8	53.1	0.84461268
Ga0436389_026_256610_257749	53.8	246.7	222.8	48.8	205.4	203.6	88.2	75.1	220.4	72.2	60.4	283.3	74.4	69.6	176.2	86.1	238.8	203.5	0.952429044
Ga0436389_026_257888_258982	248.3	370.7	322.5	219.7	505	455.3	362.2	368	532.7	338.3	217.7	320.4	382.3	350	456.9	324.4	549.9	662	0.392497552
Ga0436389_026_265589_266674	187.7	487.6	473.1	155.2	333.9	344.3	1700	1855.1	4047.4	3311.8	1730.1	3115	1268.9	386.4	805.4	3074.9	1586.1	5097	0.034876657
Ga0436389_026_284549_285352	79.5	120.8	125.6	52	92.6	104.3	63.6	62.4	105	89.4	66.3	83.2	68.9	66.5	62.7	104.8	53.7	97.5	0.50228876
Ga0436389_026_285422_287809	410.7	1180.9	1209.7	607.4	1677.2	1477.4	431.1	462.6	1941.6	469.1	313.7	1357.7	541.1	481.2	1579.1	511.7	1929.3	2347.4	0.775560037
Ga0436389_026_28686_30254	93.7	37	29.2	67.2	27.6	14.8	81.9	84.9	90.3	93.8	69	93.1	71.5	66.2	23.9	70	55.8	86.3	0.229003841
Ga0436389_026_288612_290264	166.1	16.5	15.4	171.3	15	11.6	345.2	426.4	182.8	503.6	347.6	76.6	331.3	258.3	196.4	683.5	125.4	160.8	0.389702517
Ga0436389_026_294297_294674	699.2	506.3	452.5	881.2	661.2	826.9	360.1	450.2	259.9	422.1	781.5	429.1	216	403.8	283.8	261.9	318	208.6	0.012752538
Ga0436389_026_297303_298481	348.1	225.2	212.3	293.5	168.7	174.4	751	336.1	215.9	342.6	347.3	192.5	458.4	387.5	214.3	526.8	165.6	136.3	0.749670117
Ga0436389_026_299273_300526	517.6	326.1	321	671.3	301.9	300	978	940.9	587.4	761.5	703.6	585	534.6	477.1	244.6	605.8	334.7	522	0.081110466
Ga0436389_026_3_773	81.1	52.7	60.5	381.6	190.7	260.8	106.9	245.8	176.8	432.2	300.8	228.1	231.7	103.8	36.3	686.6	205.9	264.9	0.172544266
Ga0436389_026_300596_301927	2853	3042	2932.5	1928.1	2872.2	2825.6	2279.1	2316.9	1900.8	1526.6	1343.6	1842.6	2932.7	2861.7	2029.1	1931.7	1822.6	1744.6	0.015030853
Ga0436389_026_302419_303681	61.5	159.3	143	52.6	206.6	178.8	65	55.6	229.6	27.9	37	196.1	110.9	115.1	132.8	34.3	242.7	417.3	0.748339619
Ga0436389_026_30327_31082	252	105.3	101.6	145.5	117.2	109.2	396.6	366.1	112.6	513.6	372.4	96.9	594.5	665.1	132.6	810.4	136.5	163.3	0.597327575
Ga0436389_026_303734_307321	41.7	183.9	173.2	10.8	148.5	141.2	330.7	242	1023.7	100	50.2	462.6	971	500.8	1816.7	195	929.6	900.1	0.119196104
Ga0436389_026_308331_309326	554.1	453.5	468.9	607.1	568	564.5	464.9	470.7	471.6	389	452.9	460.7	484.9	621.5	489	426.1	443.3	383.7	0.03212425
Ga0436389_026_3098_3670	618.5	332.6	295.7	397.6	345.2	314.4	2221.9	656.2	499.6	345.8	600.9	365.2	1152.3	1078.1	576.3	311.7	429.4	368	0.297811484
Ga0436389_026_31144_31962	349.8	339.4	333.8	281.3	327.7	293.1	263.1	367.1	505.9	541.3	435.2	416	409.5	324.7	487.1	647.8	440.8	469.7	0.139167839
Ga0436389_026_311703_312554	62.1	67.8	73.7	74.8	81.7	73.5	99.9	94.8	88.2	168.6	129.5	75.1	104.9	55.2	32.4	271.1	52.7	64.8	0.703617122
Ga0436389_026_316959_318353	191.1	104.3	100.7	231.4	119.3	124.8	360.6	376.1	136.8	443.2	427	150	377.1	253.6	124.4	667.9	116	147.8	0.702942521
Ga0436389_026_318356_319816	181.5	492.5	443	228.8	552.2	509.6	323	312.3	471.2	303.3	238.6	503.4	307.6	315.1	522.1	282.9	536.5	482.4	0.973155451
Ga0436389_026_31969_33069	37.6	229.8	231.1	24.7	199.1	232.3	61.9	48.4	207.7	53	31.3	205	45.3	36.9	198.1	60.8	237.3	230.3	0.880099815
Ga0436389_026_319856_320947	51.4	57.1	67.7	78.9	84.1	77.2	163	206.3	287.9	206	140.7	215.3	194	104.5	167.8	288.3	205.5	328.9	0.004663153
Ga0436389_026_322784_325090	520.4	737.3	699.6	504.6	704.2	645.2	583.4	652.2	1074.7	663.8	544.8	960.2	754.9	533	916.7	517.9	910.5	1260.4	0.710964196
Ga0436389_026_331902_332084	123	111	102	138.6	170.5	177.5	295.8	475.3	281.3	738.9	647.2	214.4	744.7	270	96.3	848.2	153.6	271.5	0.458701554
Ga0436389_026_332191_334164	55	241.4	219	47.6	287.1	257.1	104.8	65.6	354.9	37.7	42.9	259.3	157.6	117.1	386.4	55.1	317.8	336.4	0.933159492
Ga0436389_026_334102_335097	43	42.4	46.6	37.1	34.2	30.4	74.3	68.1	75.7	43.7	59.6	39	63.1	42	58.1	94.3	53.6	53.7	0.048418863

Table B.21 continued from previous page

Protein Accession	NoN1	NoN2	NoN3	NO$_3^-$1	NO$_3^-$2	NO$_3^-$3	N1	N2	N3	R1	R2	R3	P1	P2	P3	E1	E2	E3	Q Value
Ga0436389_026_335230_335505	202	65.5	54.3	178.3	69.6	55	382.1	375.8	106.8	240.9	289.3	68.4	574.1	613.5	228.1	186.6	117	52.1	0.0772504
Ga0436389_026_335561_336841	739.1	637.8	637.2	804.5	764.7	685.9	800.9	1013.5	625.9	2246	1648	708.9	861.6	809.5	418.3	1075.8	674	486.2	0.195171329
Ga0436389_026_339934_341007	2595.9	2898.5	2912.9	3884.1	5201.3	4901.2	3422.3	3816.9	3843.8	3536.4	3765.9	3932.1	2359.9	2047.8	1720.7	4188.2	3366.6	3768.3	0.000874633
Ga0436389_026_341101_342195	171.1	47	48.8	118.7	50.7	33.1	615.4	583.5	418	756.5	505.9	440.6	637.3	348.8	162.6	852.7	183	482.8	0.058949344
Ga0436389_026_344854_345756	1363.6	1398.9	1314.8	1460	1728.5	1648.9	1037.5	1336.2	1892.5	1216.7	1348.1	1998.2	888	793.8	1000.3	1185	1188.2	1597.5	0.192856887
Ga0436389_026_345838_346920	544.2	608.9	585	487.9	549.8	563.4	530.6	508.2	565.5	482.1	503.5	612.7	692.9	712	518.3	506.7	435.2	408.2	0.117754186
Ga0436389_026_346950_347978	73.4	148.1	159.3	68.6	154.6	137.1	79.8	101.7	167.6	139.9	114	103.5	92.3	79.1	140.7	68.5	150.9	122.7	0.991494579
Ga0436389_026_348863_351214	130.7	36.9	39.1	71.7	30.2	26.7	324.7	350.9	28.6	216.1	244.6	16.4	297.1	224.6	80	154	17.7	12	0.351909003
Ga0436389_026_357189_357524	131.3	63.6	75.9	29.3	14.5	13.7	301.9	216.3	46.3	105.8	82	30.7	166.8	108.8	50.1	335.2	28.7	44.1	0.546994445
Ga0436389_026_357649_359166	69.6	333.1	308.3	56	477.9	451.2	312	202.5	295.5	104.9	80.4	324.9	429.2	428.8	227.9	139.3	258.5	262.1	0.703617122
Ga0436389_026_36540_38513	619.4	725.8	694	565.9	687.9	590.6	699.6	725.5	562.8	851.4	616.9	526.9	768.8	564.2	474.4	743.3	616.1	518.8	0.960821659
Ga0436389_026_3667_4704	39.7	208.3	220.2	31.7	226.1	200.6	104.4	118.5	554.9	161.7	95.8	472.2	109.4	48	155.9	219.5	329.6	745.9	0.521018186
Ga0436389_026_373588_374766	73.8	165.6	153.1	97.8	151.5	168.5	148	156.3	603.5	178.8	145.1	593.9	240.1	199.7	139.6	290.1	265.5	629.7	0.537320858
Ga0436389_026_374852_376231	120.2	240.2	234.9	99.8	221.2	204.6	126.2	129.9	236.3	99.7	123.9	227.8	97.8	85.8	228.5	66.7	172.2	194	0.923951325
Ga0436389_026_376999_378378	106.4	52.1	43.1	111.7	27.2	31	464.1	321.1	89.7	160.8	97.5	24.3	188.5	175	21	149	42.5	28.8	0.193729723
Ga0436389_026_378378_378947	51	72.9	73.6	33.3	65.7	50	88.6	86	80.7	41.5	36.1	48.4	125.5	128.4	118.8	49.8	133.3	99.8	0.013792556
Ga0436389_026_381291_382052	60.8	97	97.8	67.9	143.6	153	58.4	58.4	114.1	65.7	64.2	118.4	48.2	50.7	68.6	57.6	94	114	0.417949327
Ga0436389_026_383156_384073	519	801.6	782.1	422.8	582.3	564.2	398.9	378.4	654.2	603.8	600.7	658	462.2	320.3	485.3	652.2	601.6	728.5	0.120158392
Ga0436389_026_387543_388010	47.5	38.8	34	32.2	24.8	27.2	59.3	81.5	18.2	137.5	100.8	21.2	74.4	35.7	26.8	262.3	24.2	4.8	0.834312757
Ga0436389_026_39316_39780	277.5	237.3	232.2	176.8	136.4	127.5	166.2	168	116.1	159	149.6	112.7	84.2	103.1	119.9	95.9	89.6	57.4	0.000840212
Ga0436389_026_400205_401221	43.9	89.6	86.9	22.9	71.6	73.8	11.4	29.6	22.3	32.1	25.8	41.2	38.7	25.3	25.5	29.8	20.7	8	0.051322758
Ga0436389_026_406021_406554	1019.9	821.8	767.7	99.8	456	399.6	444.8	419.5	313.3	249.1	359.5	259.2	468.5	499.6	558.1	382	259.1	136.5	0.00291824
Ga0436389_026_411109_412116	981.4	802	738.1	167.1	132.7	129.5	437.8	403.4	215.3	514.2	268.9	139.7	716.9	406.9	290.6	780.6	205.8	294.7	0.048914251
Ga0436389_026_416356_417222	148.2	215.8	206.3	97.8	195.2	163.8	359.3	302.7	320	187.6	182.4	254.2	228.2	292.7	256	186.6	271	313.6	0.025479647
Ga0436389_026_419705_420244	78.3	244.4	239.4	52.4	228	212.7	107	89.4	183.7	128.7	97.7	190.4	95.9	50.7	133.9	167.8	184.1	175.7	0.649261156
Ga0436389_026_42286_43002	182.7	78.6	77.7	304.7	100.4	86.8	236.8	238.3	64.2	234.1	240	67.9	154.1	197.2	44.9	191.4	53.3	34.8	0.857619793
Ga0436389_026_425305_425562	130.8	379.8	285.1	59.4	216.9	213.1	106.7	76.1	109.6	81.8	80.2	102.3	85.2	103.5	146.7	45.2	108.9	111.2	0.112652408
Ga0436389_026_45716_46756	278.5	189.9	201	67	73.2	83.8	97.4	102.4	94.3	60.4	65.6	85.8	58.1	89.9	130.5	127.3	97.8	75.3	0.002498765
Ga0436389_026_48822_50645	38.4	171.5	175.1	31.9	129.1	153.1	47.4	53.5	204.5	125.7	61.1	213.2	76.3	44.9	285.4	251	458.6	503.9	0.063555366
Ga0436389_026_50709_50960	390.9	238.2	240.1	263.2	218.5	237	264	222.6	122.9	189.9	223.6	127.8	159.2	191.1	100.8	141.2	135.4	58.1	0.087763652
Ga0436389_026_51005_52477	483.2	526	498.4	541.5	484.6	440.4	831.5	919.1	404.7	684.6	683.5	491.4	679.9	574.2	301.9	997	501.2	367.8	0.789135603

Table B.21 continued from previous page

Protein Accession	NoN1	NoN2	NoN3	NO$_3^-$1	NO$_3^-$2	NO$_3^-$3	N1	N2	N3	R1	R2	R3	P1	P2	P3	E1	E2	E3	Q Value
Ga0436389_026_57453_57722	87.1	215.5	200.9	89.3	226.5	218.8	85.9	85.8	304.9	101.2	69.3	197	99.9	89.4	356.3	158.8	434	432.4	0.468015301
Ga0436389_026_6330_8177	316.6	110.5	119.6	251.1	97.8	81.4	278.8	304.9	114.3	302.4	269.2	75.1	316.6	278.3	120.9	454.2	287.4	277.5	0.536635729
Ga0436389_026_66694_67221	407	387.4	328.1	180.1	279.6	271.6	420.9	258.2	366.1	259	248.9	331.3	207.7	270.7	265.9	232.3	574.6	568.9	0.204552299
Ga0436389_026_73288_74271	399.4	671	655.5	524.9	885	769	595.7	587.2	658	467.5	576.3	706.5	402.1	322.1	511.8	404.7	618.7	418.2	0.232174223
Ga0436389_026_79101_79475	601.2	554.3	487.2	335	624.4	587	489.1	471.6	363.9	490.6	418	447.5	550.9	447.9	346.3	551.2	353.6	346	0.703617122
Ga0436389_026_82900_83418	595.4	693.2	673.1	664	694.3	714.9	772.7	807.4	736.5	721.4	670.8	908.6	1026.8	836.6	542.2	709.6	1241.7	1276.1	0.336295418
Ga0436389_026_83466_84662	781.8	549.8	515.1	827.7	640.5	617.6	804.6	852.9	516.5	625.3	677.4	403.4	623.4	692.1	548.7	890.3	818.5	654.8	0.55477985
Ga0436389_026_9242_10216	16.2	228.6	222.1	4.8	179.9	127.4	199.4	210.3	164.8	296.8	103.6	160.3	427.9	537	162.4	138.1	277.4	195.1	0.27289012
Ga0436389_026_99579_100268	509.7	526.3	572.2	865.8	714.6	613.5	873.9	946.7	674.2	1002.1	1018.1	666.6	854.2	820.1	768.9	1133.4	838	890.7	0.079331393
Ga0436389_027_11557_11769	1136.5	581.8	521.5	696.6	743.7	785.4	618.4	655.4	440.7	571.3	770.4	422.7	449.9	500	621.8	534.6	399.4	314.8	0.366916709
Ga0436389_027_14982_16322	49.2	137.1	117.2	111.7	132.9	96	157.7	144.7	224.3	191.3	110.2	219.8	102.5	147.8	204.1	134.1	274	374.1	0.207875349
Ga0436389_027_19149_19844	108.6	180.4	200.9	106.6	181.4	168.7	162.5	136.2	198	103	67.2	143.9	218.4	186.9	190.8	110.5	132.8	160.6	0.194499817
Ga0436389_027_22226_22645	402.2	427	423	525.3	336	233.8	311.2	311.4	270.4	249.8	549.1	516.5	507.1	433.3	383.7	223.7	390.1	394.7	0.595324896
Ga0436389_027_31953_33761	76.1	206.7	210.2	60	201.3	200.8	418.7	457.7	263.1	412.1	233.6	197.1	314	222.8	346.5	502	246.2	212.3	0.239109818
Ga0436389_027_33903_34205	69.9	175.1	168.4	41.4	236.3	246	137.4	108.2	271.9	93	53.1	85.1	157.3	144.3	480.3	117.5	220.3	113.3	0.596078339
Ga0436389_027_39156_40481	554.8	294.7	299.5	604.4	322.9	268.5	1550.7	1924.4	1022.7	2743.2	1777.7	890.2	1550.8	823.4	295	3359.6	537.4	1232.2	0.272574478
Ga0436389_027_43269_43889	61.7	113.7	105.2	74	80.9	72.3	139.5	92	267.3	68.9	71.1	105.7	125.8	102.6	491.9	84.2	226.5	117.2	0.535372012
Ga0436389_027_43889_46591	548.7	474.8	470.6	402.9	247.5	244.5	684.2	648.5	471.8	506.8	522	251.2	830.6	713.4	364.6	718.3	694.6	565.5	0.119196104
Ga0436389_027_4812_6317	913.4	282.4	285.7	828	291.7	303.1	866.4	1004	437.4	883.4	856.3	315.4	709.2	681.2	432.5	874.7	365.5	399.2	0.863823735
Ga0436389_027_49480_51609	7577.9	7498.1	7198.8	7629.8	6821.9	6392.1	4336.6	4730	5079.3	4174.1	5495.3	5348.9	3504.8	4270.4	4525.3	3906.9	5178.2	4348.1	0.000538839
Ga0436389_027_54441_55385	100.3	148.2	142.4	55	157.8	154.7	389.8	329.8	131	152.3	115.6	166.6	408.4	506.2	74.9	176.4	150.6	153.8	0.302158158
Ga0436389_027_8639_9517	58.5	71.3	62.3	57.9	65	56.7	81	124	58.4	76.4	101.6	68.1	46.7	48.9	49.8	74.9	67.9	39.7	0.243379562
Ga0436389_029_10758_11861	72.9	47.7	51.3	62.4	42.6	48.7	125.8	136.4	232.6	131.9	102.7	247.2	269.1	286.8	88.5	230.2	425.9	701.5	0.041697393
Ga0436389_029_16072_17667	86.6	1020.2	986.3	73.9	1082.5	970.2	139.7	186.6	428.5	417.4	332.7	396.9	335.2	165	604.3	436.8	713.5	693.7	0.597327575
Ga0436389_029_19276_20589	8038.9	6352.1	6213.6	5756.8	7538.1	6693.3	3009.1	2473.9	1801.6	2722.6	1704.1	1122.5	3504.7	3695.6	2463.2	2935.2	1572.5	1465.3	0.000222036
Ga0436389_029_21478_23514	167.2	98.9	113.7	149.9	166.9	174.5	143	173.5	190.4	178.7	139.3	117.8	235.8	226.3	284.8	209.6	137.9	163.7	0.025629089
Ga0436389_029_23553_24386	56.8	152.3	139.6	41.9	198.9	166.8	64.7	59.5	42.2	98.6	52.4	31.3	44.6	22.7	61.7	176.1	48.7	53.7	0.408256007
Ga0436389_029_24959_27157	28	33.6	39.8	35.8	41.5	35.3	31.6	31.5	29.4	31.2	26.7	23.9	13.2	28.3	34.1	40.4	29.7	27.7	0.372138571
Ga0436389_029_28040_30472	55	739.2	702.2	19	400.1	352.1	100.8	107.9	560.7	50.1	65.5	372.2	270.5	212.2	650.1	95.4	425.6	379.1	0.775560037
Ga0436389_029_40651_42792	96.4	669.5	615.2	83.9	561.2	536.4	50.9	71.3	766.4	161.1	130.7	667.9	80	64.4	651.7	75.1	571.5	632.9	0.976365372
Ga0436389_029_46607_47680	90.6	158.9	164.6	81.9	152.6	134.4	787.9	746.1	835.5	1123.7	610.1	523.2	1019.9	865.3	702.8	1015.6	528.7	770.2	0.003736291

Table B.21 continued from previous page

Protein Accession	NoN1	NoN2	NoN3	NO_3^-1	NO_3^-2	NO_3^-3	N1	N2	N3	R1	R2	R3	P1	P2	P3	E1	E2	E3	Q Value
Ga0436389_029_47829_50471	2298.6	1714.6	1618.4	1036.7	1182.3	1004.7	2214.6	2242.9	2170.7	2004	1239.5	1298.7	1773.1	1436.3	1149.8	3461	1360	1999.9	0.191553874
Ga0436389_029_60466_61059	348.1	753.4	708.3	410.4	1136.1	933.8	643.3	883.5	1147.7	976.3	668.3	1023.9	607	646.2	1202.7	1074.1	910.8	897.8	0.7608511
Ga0436389_029_62289_64193	6443.6	5837.6	5456.5	3963.1	5615.7	5679.7	7590.3	6347	6322.4	4074.2	4661.5	5203.3	6263.2	6034.9	4495.2	3922.4	4969.4	5406.2	0.11313324
Ga0436389_029_65869_66231	556.3	659.1	690.3	737.8	1374.3	1202.4	899.2	952.9	694.7	845.3	769.9	743.6	1038	1199.8	675.8	1071.1	606.1	541.4	0.333382375
Ga0436389_029_66995_68209	1787.2	942.7	904.3	2048.3	913.4	831	1927.2	2139.2	1485.3	2032	2041.4	1464.2	1758.3	1344.7	494.3	1662.7	1052.9	1731.9	0.540117653
Ga0436389_029_68542_68814	53.3	47.6	42.9	97.8	77.2	68.1	44.1	86.3	178.4	63.3	71.9	87.1	31.7	25.8	476.9	74.6	164	94.1	0.849587163
Ga0436389_029_70514_71719	137.9	137.3	149.7	98.1	110.6	120.2	187.3	185.7	228.6	187	145.6	201.6	250.4	292.2	130.7	208.6	164.6	270.5	0.12022744
Ga0436389_029_72789_73283	444.2	332.3	298.9	348	352.7	341.7	373	454.1	323.8	684.1	677.4	495.1	261.8	269.7	311.3	719.1	250.9	314.7	0.15484998
Ga0436389_029_79625_81022	68.8	208	198.6	42.5	139.1	135.3	131.7	160.5	141.6	85.6	50.8	119	611	414.3	157.2	90.6	131	128	0.081110466
Ga0436389_029_89163_89789	364.7	549.5	507	249.9	445.4	403.6	277	231.5	618.5	222.4	258.7	491.5	384.5	280.1	641	226.2	558.3	564.9	0.900587056
Ga0436389_029_89839_90942	629.8	644.4	637	555.6	543.7	530.5	415.8	505.6	390.1	331.7	424.9	385.7	293.4	318.9	378.2	288.4	346.6	230.6	0.000137732
Ga0436389_029_91292_92479	1211.7	1116.6	1116	1153.9	906.5	848.4	1288.5	1206.8	826.1	965.5	1117.1	894.7	1563.3	1621.2	863.7	1001.5	899.8	882.1	0.395817504
Ga0436389_029_93403_95772	152.6	592.7	572.8	167.1	668.4	593.9	200.2	171.7	550.5	120.7	157.3	529.2	494.7	360.4	492.6	105.6	503.6	363.1	0.84461268
Ga0436389_029_958_1443	21.2	33.9	24.5	16.5	39	36.2	319.8	284.6	51.4	46.1	38.6	36.1	1085.2	1914.7	57.5	92.7	49.4	24.7	0.121779507
Ga0436389_029_95811_97331	367	300.4	282.4	503	397.4	406.1	398.8	350.7	439	303.1	289.7	366.1	218.4	289.2	396.1	255.4	354.4	332.1	0.159393564
Ga0436389_029_98382_99137	49.6	65.5	53.9	58.9	32.9	29.6	89.5	101.6	60	219.4	147.2	75.9	105.7	84.6	59.2	107.2	50.2	37.3	0.118155327
Ga0436389_030_12641_13579	30.2	105.2	120	38.3	121	129.1	68.6	78.2	144.3	76.2	56.3	158.4	51.2	53.8	281.6	89.6	208.4	193.2	0.842921471
Ga0436389_030_17649_18806	801.2	747.4	759.9	836.9	942.2	972.8	816.2	836.8	675.7	903.2	854.8	726.6	585.5	678.3	766.2	863	699.5	651.6	0.148860042
Ga0436389_030_22588_23661	644.6	466.1	460.6	383.9	347.7	407.8	797	650	483.1	592.4	484.7	359.3	1127.3	1232.8	654.9	796	424.2	464.2	0.051827955
Ga0436389_030_26116_26676	2809.7	2584.7	2594.9	1022.1	1857.1	1267.1	1987.7	1851.4	1595.5	1560	1321.5	1505.5	2018.7	2087.1	1506.2	1105	1386.9	1508.2	0.003508896
Ga0436389_030_27016_27483	716.1	841.7	849.3	567.7	724.8	707.5	594.5	659	579.8	429.1	581.6	645.5	836.4	818.6	767.8	413.4	698.5	407.2	0.039326564
Ga0436389_030_31004_33337	4640.5	2946.7	2864.3	2864.3	3002	3095.4	2532	2741.4	2582.2	1960.6	3161.7	2264.7	1597.7	2249.9	1800.5	1601.7	2165.3	1983.8	0.058925315
Ga0436389_030_34763_35467	87.4	239.3	240.2	36.3	146.2	160.3	177.3	208.3	228	953.3	609	167.5	371.3	207.6	266.4	803.7	201	195.1	0.344403191
Ga0436389_030_37202_37510	90.2	82	76.8	63.8	73.8	67.2	102.9	159.7	130.1	119.6	91.7	118.8	81.4	77.7	41.2	161.8	67.4	135.9	0.095975137
Ga0436389_030_37601_37846	196.3	455.9	285.8	192.3	721.5	804.8	154.4	165.4	411.4	160.5	192.3	568.4	113.6	156.1	292.3	139.5	360.9	437.3	0.45710066
Ga0436389_030_38087_39844	17.4	67.9	69	22.9	73.1	69	153.8	155.5	259.4	62	60.8	205	677.5	403	240.1	129.3	187.8	330.1	0.035258266
Ga0436389_030_40722_41879	203	416.5	437.8	83.9	279.8	308.7	341.2	328.3	560.1	728.5	504	453.2	447.3	248.9	567.8	1194.8	489.6	572.6	0.19619854
Ga0436389_030_42484_42870	369.9	607.1	620.5	310.2	613.8	576.9	496.8	534	495.5	418.1	438	387.9	387.7	357.8	593	527	348.8	226.5	0.680745895
Ga0436389_030_4887_6533	577.3	598.8	573.9	510.6	519	477	607.8	632	677.9	561.9	561.9	605.8	858	690.7	1377	714.9	668.9	510.7	0.095240545
Ga0436389_030_9106_9573	172.2	72.6	63.5	164.1	68.8	77.9	315.1	251	161	315.4	229.2	106.8	363.7	256.5	57.5	267.2	79.4	143.2	0.504794396
Ga0436389_031_11820_12821	738	926.7	879.8	547.2	580.7	537	446.3	535.1	648.5	353.6	332	596.7	655.3	506	850.4	401.5	635.3	583.9	0.057236091

Table B.21 continued from previous page

Protein Accession	NoN1	NoN2	NoN3	NO$_3^-$1	NO$_3^-$2	NO$_3^-$3	N1	N2	N3	R1	R2	R3	P1	P2	P3	E1	E2	E3	Q Value
Ga0436389_031_14039_14512	104.4	181.7	197.7	87.2	227.5	241.1	152.6	134.9	147.7	143.7	148.9	148.6	131	109.8	156.6	186.1	250	253.2	0.29131425
Ga0436389_031_24793_25269	509.9	128	102.8	522.9	97.1	94.1	503	608.8	203.2	329	351.8	180.7	413.3	393.8	213.6	323.9	123.2	209	0.76821363
Ga0436389_031_25370_28978	123.4	259.2	244.1	91.4	157.7	160.4	170.6	139.5	322	237.2	235.1	240.6	366.5	188.7	299.5	686.1	241.6	435.3	0.148860042
Ga0436389_031_289_642	124.3	371.6	337.7	98.7	472.6	407.7	134.4	189.8	398.4	235.3	212.2	392.3	215.6	147	550.9	430.7	363.5	168.9	0.990699199
Ga0436389_031_29071_30279	213.9	531	536.2	204.5	541.4	540.2	260.1	270.6	632.9	310.8	240.6	731.5	259	245.4	606.5	372.2	821.5	755.5	0.77254526
Ga0436389_031_30298_30714	103.4	46.8	44.7	173.5	58.7	64.8	197.5	199.5	62.3	156.3	128.9	48.8	235.5	217.1	22.5	135.3	41.3	42.9	0.648711355
Ga0436389_031_32334_33683	164	273.9	275.6	167.4	271.1	246.8	317.8	324.6	444.2	300.3	249.1	360.1	524.2	476.9	491.2	344.5	532.5	528.5	0.008424705
Ga0436389_031_34313_35287	68	124	101.6	80.3	168.5	171.2	111.9	107.3	173.7	73	87.9	147.6	156.2	132.8	237.5	74.4	208.8	135	0.580478053
Ga0436389_031_36594_38615	59.4	122.5	127.5	73.4	151.8	184.1	153.5	162.3	123.8	108.9	102.5	159.9	398.3	268.7	97.8	182.3	312.7	213.6	0.247532176
Ga0436389_031_38686_39030	15.6	38.3	33.7	14	22.7	25.5	26.6	41.1	41.1	55.6	41.3	30.1	53.5	34.4	28.2	62.1	30.1	37	0.393178772
Ga0436389_031_40957_41127	138.1	34.5	39.8	218.2	20.2	19.9	112.9	120.1	121	78.5	119	63.9	54.1	83.3	172.9	64.4	64.2	116.8	0.960821659
Ga0436389_031_41189_42022	1231.7	848.5	843.6	731.8	581.5	629.7	398.5	500	464.2	422.6	514.2	556	543.9	723.3	1011.6	598	661.5	445	0.039326564
Ga0436389_031_46366_48477	571.8	329.2	313.7	472.4	289.4	303.3	1000.6	1131.5	495.3	670.8	664.4	329.2	694.2	863.8	615.2	879.8	401.4	292.6	0.210351732
Ga0436389_031_52307_53371	977	386.4	387.7	799.4	637.9	587.9	881.4	934.8	510.8	496.5	520.8	372.8	407.5	654.4	670.6	484.3	451.1	422.3	0.472830503
Ga0436389_031_54371_55471	213.2	169	151.1	196.5	140.3	126	261.5	410.1	744.3	431.1	394.8	487.2	247.3	197.5	696.4	487.8	373.7	458.8	0.194499817
Ga0436389_031_6303_6521	822.3	1252.3	791.8	1064.7	966	1224.8	1121.8	921.5	1421.2	629.4	1233.5	1256.8	460.8	490.1	604.4	481.8	749.4	1127.4	0.178013453
Ga0436389_031_645_1070	170.4	27	23.8	148.6	25.8	27.5	292.7	285.1	69.3	265.6	211.7	74.9	266.6	251.1	46.2	280.4	49.4	70.2	0.585880628
Ga0436389_031_6588_7580	101.6	36.9	32	73.2	26	34.6	199.5	139.5	25.9	222.2	134.1	28.2	568.7	345.8	22.7	184.4	25	9	0.352563451
Ga0436389_031_71970_73448	82.5	107.2	106.3	9.1	135	136.8	157.7	137.8	240.1	141.4	100.9	311.5	84.8	63.4	285.4	979	172.1	292.8	0.35933369
Ga0436389_031_74991_76298	29.7	76.6	77.3	23.6	45.3	48.8	60.3	75.1	107.3	33.1	40.4	46	88.2	89.8	95.9	15.4	115.5	68.3	0.283569116
Ga0436389_031_76295_77575	495	307.4	286.2	598	301.7	301.1	749.5	703.1	383.1	471.6	579.5	322	932.5	831.2	312.2	784	559	496.4	0.444105459
Ga0436389_031_7645_8883	64.9	117.5	124	49.3	139.7	126.5	112.3	71.1	200.9	43.8	50.6	116.4	36	35.4	341.3	65.7	162.3	121	0.960821659
Ga0436389_031_77577_78380	286.1	402.6	399.9	234	420.1	362.3	370.2	297.4	462.2	186.1	202.6	314.4	282	260.4	442.3	335.6	660.2	553.7	0.200619907
Ga0436389_031_78454_79896	396.5	279.4	282.6	596.3	378.4	309.5	1008.3	807.9	531.6	751.2	567.7	416.2	674.7	647.7	823.9	687.2	519.5	484.9	0.068027819
Ga0436389_031_79914_80279	156.8	242.8	202.7	201.4	328.8	300.3	101.5	133.1	344.2	117.6	112.4	283.4	85.1	95.7	626.6	112.8	278.2	169.6	0.947865961
Ga0436389_032_10755_11810	33.1	236	222.6	32.1	186.2	164.5	52	67	312.7	64.3	34.6	269.4	710.3	66.4	320.5	84.4	465.3	449.4	0.580478053
Ga0436389_032_11838_12710	136	66.7	71.8	124.1	59.3	64.8	300.1	262.1	108.2	135.6	144.8	81	395.6	390.4	91	123.6	170.5	178	0.168713497
Ga0436389_032_1569_2366	82.1	255.1	218.4	125.2	185.9	176.4	78.1	113.9	215.2	136.1	129.1	262.5	90	100.6	188.9	143.9	173.7	190.3	0.883259142
Ga0436389_032_16414_16842	260	249	227.1	526.3	1010.4	1043.1	528.6	506.9	545.6	340.6	480.4	485.5	286.3	452.5	644.2	347.5	685.1	558.6	0.04637225
Ga0436389_032_17800_18822	81.5	125.1	116.8	85.1	91.8	100.2	52.3	83.4	161.6	117.7	102.1	107.6	53	57.8	235.6	97.7	213.1	218	0.662091974
Ga0436389_032_18824_19840	134.9	154.5	154	169.2	145.2	183.9	174.3	145.9	128.9	197	271.1	215.5	291.6	344.3	248	191.3	196	186.3	0.00283184

Table B.21 continued from previous page

Protein Accession	NoN1	NoN2	NoN3	NO_3^-1	NO_3^-2	NO_3^-3	N1	N2	N3	R1	R2	R3	P1	P2	P3	E1	E2	E3	Q Value
Ga0436389_032_26504_27268	125.7	67.1	64.9	68.3	32.5	40.5	144.3	103.8	79.2	90.3	93.8	62	97.3	99.4	34.2	104.8	41.4	61.7	0.442723876
Ga0436389_032_28769_30382	36.4	27.1	24.9	29.9	23.7	20.7	44.3	41.2	94.2	50.5	30	61.7	66.6	36.4	438.4	56.1	149.9	104.1	0.526501556
Ga0436389_032_293_556	1269.3	1544.3	1573.7	475.7	727.7	726.2	459.4	374.4	654.6	252.6	387	595.8	254.2	444.9	966.5	307.6	595.1	455.4	0.003697608
Ga0436389_032_31305_32726	226.4	223.5	244.5	286.9	324.9	314	292.6	300.9	466	233.4	296	409.9	157.6	162.1	375.1	186	363.9	497.5	0.626719426
Ga0436389_032_32891_33238	309.6	462.8	364.5	184.4	404.6	381.8	240.7	262.3	365.5	297.8	274	330.8	239.5	239.1	344.6	487.4	296.3	257.1	0.789381707
Ga0436389_032_33968_34969	194.5	194.1	224.7	157.7	103.8	96.1	631.8	469.4	142.3	218.6	187.8	131.5	266.7	249.1	154.4	328.2	165.4	171.2	0.217707413
Ga0436389_032_35938_37950	12	77.5	91.6	3.7	90.5	89.7	430.7	555.9	588	887.4	440.6	396.6	320.3	94.1	118.8	856.9	250.7	625.4	0.025134825
Ga0436389_032_41323_42405	357.6	788.4	750.9	333	710.5	662.7	248.7	267.7	994.6	252.5	262.8	841.3	333.3	272.8	581.3	240.1	726.5	1291.1	0.877551687
Ga0436389_032_43693_44277	46.2	201.4	192	23.6	116.6	114	146.2	171	238.3	201.7	151.1	146.9	205.7	342.5	477.7	286.4	198.8	176.4	0.078239773
Ga0436389_032_46280_46882	1047.3	629.4	651.6	1009.8	456.4	430.2	1038.5	978.5	409.5	1079.5	993.3	378	1015.2	1014.4	418.9	905.4	255.6	222.3	0.826836516
Ga0436389_032_48164_49306	167.6	357.6	334.8	111.5	202.1	201.4	215.2	164	236.2	156.3	108.3	191.1	175.3	146.9	254.2	140.7	224.6	221.3	0.333382375
Ga0436389_032_49450_49980	45.1	152.9	147.5	32.4	120.5	139.8	27	37	97.5	30.7	44.1	169.9	31.7	44.6	134.3	27.9	140.4	129.5	0.880099815
Ga0436389_032_53894_55153	554.1	461.1	447	476.6	401.6	354.6	884.5	854	600.2	922.7	656.9	441.8	854.5	700.3	536.4	833.3	430.2	529.7	0.244287239
Ga0436389_032_56652_57287	563.7	532	503.2	429.9	564.9	530.1	247	365.8	329.5	210.4	349.2	358	186.8	248.2	226.1	193	302.2	218.2	0.001064225
Ga0436389_032_62224_63021	504.8	509.8	459.4	314.1	349.5	337.5	337.2	343.3	291.2	324.4	292.3	333.3	236.3	217.3	258.9	426.7	201.1	191.2	0.018854187
Ga0436389_032_6540_7061	31	87.8	81.4	46.4	145.2	114.6	25.4	39.4	153.1	81.1	53.6	107.6	47.8	30.4	157.1	94.6	107.8	112.8	0.926265459
Ga0436389_032_67162_68346	804.1	1145.9	1109.8	580.1	766.7	681.3	664.8	662.4	886.4	485	494	711.3	664.3	722.8	1118.9	485	622	483.7	0.063974878
Ga0436389_032_71407_71673	100.3	100.3	67.6	59.8	67.8	65.5	136.4	131.5	76.9	157.9	83.8	86.6	123.9	75.5	42.5	158.6	89.9	103.3	0.472830503
Ga0436389_032_76480_77469	37.9	169.4	135.6	10.9	95.6	131.7	93.8	70.3	382.7	37.3	29.8	406	96.5	193.3	626.4	37	252.5	688.3	0.7608511
Ga0436389_032_79814_80881	191.1	399.7	389.6	158.5	429	419.8	263.1	317.8	820	384.8	333.6	695.6	273.6	148.6	483.5	419.7	404.4	760.6	0.753985871
Ga0436389_032_82184_82780	57.3	93.2	92	61.1	78.3	85.2	32.2	69.1	62.6	42.3	57.6	56.8	41.8	64.8	87.4	41.7	69.5	54.1	0.421582471
Ga0436389_032_82878_83618	742.6	725.9	698	607.4	715.7	689.6	362.8	398.7	434.3	343.1	473.1	492.2	266.1	350.3	479.6	286.8	497.2	363.2	0.002186741
Ga0436389_032_8650_9396	1105.9	1215.1	1067.3	1069.9	1487.4	1531.4	876.3	976.1	1024.7	980.1	963.9	953.8	761.8	992.9	884.4	1092.7	890.1	740.7	0.042145464
Ga0436389_032_9377_10729	130.3	125.9	115.9	57.8	162.4	132.8	216.1	164.9	472.5	81.5	78.9	378.3	146.6	381.4	307.5	119.8	283.9	450.6	0.534606147
Ga0436389_033_1518_2282	3891.7	5715.3	5427.1	5873.2	5815.2	6899.1	3785.4	4152.4	5134.4	4230.7	5266.5	6493.7	3385.9	3483.7	3617	3275.7	5330.2	4317.2	0.078381735
Ga0436389_033_768_1400	113.3	43.3	44.1	64.6	50.2	48.5	78.7	72.8	53.7	71.9	80.1	48.6	57.4	60.2	26.5	52.7	66.9	49.2	0.84461268
Ga0436389_034_100615_101928	340	521.1	523.2	334.9	597.7	597.3	793.8	762.6	908.3	420.9	293.2	723.6	1406.1	994.1	670.4	907.2	788.1	789.2	0.057461534
Ga0436389_034_102442_104241	129	376.3	395.7	136.4	367.9	326.8	390.8	241.2	437.7	217.1	153.8	369	325.4	271.9	460.3	288.2	775.5	675.8	0.302730006
Ga0436389_034_105439_106125	42.5	60	70.6	33.6	46.5	49.1	51.2	54.3	79.9	41	43.1	63.6	80	104.7	105.9	47.1	141.2	126.7	0.096385027
Ga0436389_034_106258_106620	1006.1	976.1	1001.3	1092.5	1304.4	1402	955.7	919.5	884.3	785.9	928.2	1016.7	908.4	933.5	1009.1	629.6	1040.5	827.7	0.045470018
Ga0436389_034_106626_106994	132.8	84.8	83.7	92.7	73.6	88.9	320.2	357.5	397.4	348.6	294.6	368.9	735	494.8	213.3	457	303.6	597.8	0.03051462

Table B.21 continued from previous page

Protein Accession	NoN1	NoN2	NoN3	NO_3^-1	NO_3^-2	NO_3^-3	N1	N2	N3	R1	R2	R3	P1	P2	P3	E1	E2	E3	Q Value
Ga0436389_034_110575_111900	352.5	247.9	243.8	332.9	245.5	257.3	477.1	418.4	238.4	290.4	355.6	226.8	508.6	428.1	169	281.8	196.1	185.1	0.565518388
Ga0436389_034_115884_117023	2107.4	1742.8	1545.5	1140.2	1200.9	1124.9	2079	2317.3	1630.1	1401.3	1366.6	1404.8	2927.2	2809.4	1701.6	1880.4	1293.1	1362.6	0.035079555
Ga0436389_034_117107_117496	85.4	105.7	104.2	50.7	64.9	76.9	159.2	190.4	96	154.1	89.6	79.6	177.6	129.1	87.1	199.7	104.9	84.8	0.371975359
Ga0436389_034_118666_119205	1068.5	1344.9	1268.4	672.1	1416.1	1334.8	787	761	1024.3	800.1	809.8	981.5	810.2	730	1106.1	973.4	917.9	920	0.333068295
Ga0436389_034_119235_120056	269.6	449.3	437.3	359.6	452.3	432	335	332	288.6	224.7	236.3	302.7	576.8	505.8	341.6	322.7	295.8	226.3	0.068567799
Ga0436389_034_120282_120986	26.4	24.7	26.4	25.9	49.8	51.6	96.2	72.2	127.9	73.4	47	88.3	180.7	134.2	37	86.1	75.5	153.5	0.15225946
Ga0436389_034_122385_123560	80.9	366	345	83	339.1	328.8	69	43.1	262.5	31.6	19.8	269.1	11.8	27.5	250.1	44.1	207.2	242	0.649261156
Ga0436389_034_124251_126146	69.1	56.8	54.8	146.3	77.2	57.9	89.4	105.7	104	81.6	82.4	52.2	115.7	72.6	205.1	43.9	95.1	82.2	0.409673167
Ga0436389_034_126808_129276	116.8	110.8	116.2	55.8	122.6	135.7	155.1	181.7	298.4	125.5	85	109.3	172.9	142.4	407.4	1308.8	647.3	399.3	0.04160352
Ga0436389_034_132099_132971	24.8	85.5	83.9	31.4	77.7	77.6	112.4	79.6	195.6	127.1	66.8	273.8	84.5	121.7	103.7	141.9	112.8	169.8	0.383579104
Ga0436389_034_132980_134293	849.2	657.8	647.8	744.6	891.8	863.4	1005.2	913.9	704.9	730	784.3	581.3	1102	1550.3	1006	834.4	702.7	484.4	0.055219206
Ga0436389_034_13470_16556	45.2	146	133	104	90.6	100.2	116.2	116.4	99.9	216.6	180.5	90.1	155	120	111.5	213.2	147.7	136.1	0.38658556
Ga0436389_034_136032_137081	125.8	305.9	277	241.9	600.8	552.8	427.8	338.5	600.8	432.3	279.8	429.2	276	182.8	501.7	538.7	709.4	623.7	0.131149751
Ga0436389_034_137101_138090	115.7	136.3	147.7	129.7	172	172.8	189.9	164.8	109.3	113.5	112.7	108.9	174.2	187.7	115.2	112.8	85.7	57.6	0.103551455
Ga0436389_034_138836_139579	173	223.5	216.4	163.9	195.5	190.1	295.7	301.4	277.9	398.1	310.2	154.8	331.2	245.5	240.2	351.2	324.3	327.9	0.125280741
Ga0436389_034_139586_140119	71.4	194.4	183.8	134.3	318.9	286.5	181.4	221.5	335.2	192.8	178	282.4	340.5	183.9	297.8	287.5	466	427.9	0.139167839
Ga0436389_034_144476_145480	567.3	681.2	645.2	636.7	638.7	616.8	753	796	771.9	616.1	594	609	846.7	705	673.5	743.1	605.1	524.6	0.087749901
Ga0436389_034_146075_146959	89.4	127.8	124.3	124.2	90.1	122	123.3	126	103.7	116.2	100.6	98.8	135	99.2	82.8	143	73.7	73.1	0.932176973
Ga0436389_034_147623_148528	23.2	334.9	322.4	42.9	643.1	589.3	60.7	112.6	444.4	46	37.6	325.9	140.5	150.3	402.3	49.1	999.7	909.5	0.473658566
Ga0436389_034_156962_157243	134.9	257.5	232.9	79.9	148.5	134.5	133.1	125	109.2	80.9	83.6	110.7	283	188	114.2	97.6	98.2	74.7	0.090914086
Ga0436389_034_157455_157766	6659.6	6530.2	6103	3702.5	6961.9	6313.1	3565.6	3749.8	3153.9	3194.1	3542.5	3353.5	2902.8	2856.7	2468.3	3494.1	2710.5	2127	0.00283184
Ga0436389_034_157818_159230	7372.8	7326.1	7066.7	8360.4	8177.6	7302.1	7000.9	7211	7108.6	6153.1	7353.2	6585.5	5704.7	5128.1	4687.1	5960.9	7101.2	6436.5	0.00283184
Ga0436389_034_159332_159673	920.1	1100.3	1087.7	609.8	1127.1	1057	531.4	612.6	609.2	861.5	732.3	789.9	447.1	352.4	538.1	1008.6	579.3	491.3	0.045470018
Ga0436389_034_159746_162112	6400.8	5276.5	5285.5	5918.1	5335.3	5986.8	5416	5377.6	5437.5	3541.8	4745.4	4481.2	3817.3	4183.3	3816.4	3664.5	4547.5	3919.6	0.00392427
Ga0436389_034_162120_163658	196.5	385.9	362.3	155.3	295.4	281.4	159.3	174.7	516.8	182.8	157.5	271.7	147.9	150.3	1047.2	157.6	334.7	188.6	0.880099815
Ga0436389_034_163661_163975	368.5	289.5	288.2	233	318.6	294.9	253.9	268.8	241.5	188.4	220.3	239.6	207.7	222.1	138.7	144.6	204.5	187.4	0.018857035
Ga0436389_034_19011_19817	25.5	72.8	52.6	14.2	51.3	50.8	64.1	102.3	102.4	91.4	49	81.9	115.8	109.3	54.6	125.9	79.5	140	0.092484161
Ga0436389_034_51799_54357	126.9	337.4	354.2	46.4	133.5	134.7	176.8	151.6	319.3	318.3	180.9	193.3	177.6	123.1	430.3	282.2	225.9	244.1	0.572511951
Ga0436389_034_55284_56603	48.6	445.8	393	76.2	281	265.5	184.5	159.5	276.7	39.1	22.1	261.9	3.4	18.4	227.2	37.5	271.1	295.9	0.626719426
Ga0436389_034_64984_65754	367.8	550.5	520.4	396.3	779.6	780.7	322.6	358.8	562	296.3	363	583.5	284.2	306.6	638.4	310.5	609.1	434.4	0.603222152
Ga0436389_034_65806_68100	884.4	1467	1456.2	525.5	1212.1	1067.7	538.5	600.9	998.9	476.1	609.2	1488.9	505.8	596.7	1126.6	514.5	1029.4	1095.8	0.651032499

Table B.21 continued from previous page

Protein Accession	NoN1	NoN2	NoN3	NO_3^-1	NO_3^-2	NO_3^-3	N1	N2	N3	R1	R2	R3	P1	P2	P3	E1	E2	E3	Q Value
Ga0436389_034_68197_69255	159.3	491.5	460.5	122.7	419.5	432.1	288.2	304.3	443.5	378.4	316.5	308.2	278.7	160.2	303.1	523.3	279.2	297.7	0.902033568
Ga0436389_034_695_1744	8	22.4	20	3.8	19.2	22.6	140.8	219.2	302.7	260.6	137.2	196.5	130.4	46.9	24.4	310.3	142.1	456.6	0.01782435
Ga0436389_034_70361_70702	2113.1	2316.9	1801.2	4476.9	4436.7	4925	2106.8	2198.7	1636.4	1731.5	2673.5	2409.2	1346.3	1454.1	817.4	1120.2	1285.1	989.6	1.40236E-05
Ga0436389_034_71575_74205	73.6	241.8	232.5	55.6	145.6	154.7	365.5	684.6	621.7	1035.6	568.4	518.4	612.3	386.1	662.5	1864.7	555.3	904.8	0.072665902
Ga0436389_034_79200_79616	35.3	196.3	192.1	17.6	170.5	205.6	83.3	82.7	186.3	109.3	70	173.8	95.6	61.8	128.3	161.5	113.4	125.7	0.973155451
Ga0436389_034_81921_82508	697.5	560.2	559.6	479.4	581.2	521.5	517.9	762.2	484.4	475.5	647.3	502.1	633.2	641.8	328.8	536.6	483.9	452	0.877551687
Ga0436389_034_82555_83199	339.2	360.7	365.1	632.3	784.6	718.1	372.5	487	535.7	255.6	280.2	400.8	464.1	537.4	544	289.7	600.3	529.9	0.013756859
Ga0436389_034_85477_86430	963.3	594.4	569.5	1049.5	638.8	518.7	1387.3	1556.7	688.9	976.1	1163	675.6	1692.9	1598.2	924.8	945.3	600.8	523.4	0.180522374
Ga0436389_034_86538_86939	83.4	154.8	139.3	76.1	143.4	112.8	179.1	207.8	111.8	111.4	93	84.5	230.1	228.1	95.9	151.1	170.1	156.6	0.29132233
Ga0436389_034_90680_91993	457.7	364.4	323	321.4	349.1	298.5	794.8	849.3	1429.5	1306.3	921.9	425.5	524	433.4	2791.4	820.9	1137.3	619	0.55477985
Ga0436389_034_95986_96882	142.5	244.3	233.6	82.1	221	191.8	301.2	264.9	287.8	148.3	95.4	223.5	445.1	511.2	478.3	106.1	211	190	0.000985048
Ga0436389_034_97526_97777	92.1	87.4	59.6	82.6	72.3	73.9	67.6	57.5	138.6	58.8	80	107	44.3	88.1	143	42.6	87.5	122.1	0.994167314
Ga0436389_034_97778_98308	264.8	372.9	360.5	92.7	267.3	245.7	194.8	197.5	245.6	146.8	133.8	339	341.4	355	235.8	113.1	196.7	227.5	0.243497196
Ga0436389_035_10584_11219	233.5	291.7	280.3	455.5	486.2	417.8	322.4	359.6	446.8	173.7	287.7	438.9	379	304.7	516.2	139.3	276.3	178.6	0.071433399
Ga0436389_035_11212_11598	44.6	36.6	41	52.2	66.7	55	122.6	124.2	31.8	56.4	72.5	54.2	120.3	194.1	23.5	96.4	32.1	44.9	0.545815122
Ga0436389_035_12160_13143	229	315.6	308.9	272.5	267.4	233.8	233.4	236.9	512.3	190.5	252.7	397.6	212.4	188.5	662.9	181.5	891.8	681.8	0.56208721
Ga0436389_035_16419_17534	206.5	100	84.2	106.2	113.9	105.1	700.8	690.9	135.6	689.5	375.4	65.9	835.6	741.9	270.2	898.1	127	89.6	0.424833573
Ga0436389_035_17632_18282	48.5	125.2	123.7	36.7	96.6	104.1	82.3	94.7	208.2	33.3	40.6	122.7	183.3	152.7	174.8	31.3	125.4	127.3	0.30826916
Ga0436389_035_21873_22319	312.7	507.6	521.9	407.9	426.3	421.6	332.4	366.1	667	272.3	260.8	480.1	279.7	208	936.2	307.8	586.2	592.1	0.960969675
Ga0436389_035_25523_25975	733.3	279.1	285.1	49	37.3	59.2	114	105.8	153.2	126	75.4	51.4	166.1	169	112	141	61.9	95.6	0.045177535
Ga0436389_035_26926_27327	154.9	86	89.6	293.9	217.7	227.3	197.9	209.3	75.3	271.2	269.8	128.6	162.1	120	35.1	237.4	99.2	92.2	0.212501218
Ga0436389_035_30354_30695	433.1	398.2	414.6	613.7	1112.7	982.4	362.8	428.6	591.2	365.4	518.2	541.3	283.6	308.6	556.1	410.3	715.2	604.4	0.044973608
Ga0436389_035_32118_33074	3302.8	3866.2	3506.5	2616.4	2504.7	2799.4	2386.3	2702.6	2176.1	2599.4	3032	3483.9	2699.8	2451.9	2333.9	2196.9	1901.4	1783.2	0.00283184
Ga0436389_035_33266_34033	485.8	355.9	381.6	789.4	533.3	595.4	642.2	644.9	600.9	699.2	670.2	629.3	597.9	421.9	478.1	860.4	718.5	707.1	0.014632382
Ga0436389_035_3387_4148	955.1	941.2	962	1008.7	1139.6	1075.5	985.6	1035.3	898.2	940.8	978.6	978.5	866.3	847.3	528.5	873.4	722.8	1196.1	0.283569116
Ga0436389_035_4417_6156	168.2	107.2	108.3	167.3	90.3	230	160.8	145.1	64.8	130.2	169.8	143.5	192.4	177.5	57.3	149.3	74.3	69.7	0.82310684
Ga0436389_035_45633_46937	189.3	116.2	107.2	99.4	134	114.4	321.8	338.7	204.4	386	305.6	177.2	320.5	274.7	140.4	383.1	131	320.5	0.192856887
Ga0436389_035_47426_54061	1459.5	1619.5	1539.5	230.8	394.9	387.3	731.9	646.7	717.8	512.8	512.3	689.7	1442.8	600	775.8	1505.8	639.7	621.9	0.02169711
Ga0436389_035_6196_6426	809.5	1052.2	624.3	730.3	1151	1194.5	606.3	658.4	877.6	540.1	613.1	883.4	555.7	454	718.1	432.4	767	659.6	0.218490699
Ga0436389_035_6439_7116	558.9	411.7	404.4	621.2	653	756.2	890.9	1031.6	433.8	809.5	782.6	498.1	620.8	637	340	804.1	549.5	460	0.490184598
Ga0436389_035_7139_10462	378.6	738.3	750.3	268.5	786.9	886.7	828.9	972.2	2155.7	860.9	641.1	1286.6	873.1	643.7	1314.8	1285.4	2111.7	2847.2	0.092484161

Table B.21 continued from previous page

Protein Accession	NoN1	NoN2	NoN3	NO_3^-1	NO_3^-2	NO_3^-3	N1	N2	N3	R1	R2	R3	P1	P2	P3	E1	E2	E3	Q Value
Ga0436389_037_14309_15898	32	91.2	75.8	33	122.2	90.7	36.6	30.7	151.9	19.6	22.7	112.8	42.6	38.9	119.1	16.3	177.6	187	0.813285008
Ga0436389_037_15970_16170	126.7	123.5	163.3	152.4	291.2	302	82.6	118.5	113.1	76.1	190.3	188.5	48.4	58.5	99.6	63.8	138.2	118.5	0.044973608
Ga0436389_037_16239_16586	224.9	123.9	127.6	214.8	198.3	189.7	264.8	295.4	137.3	238.3	246.2	164.2	357.1	280.8	53.1	341.7	214.2	221.7	0.84461268
Ga0436389_037_16624_17187	24.5	202.4	193.9	17.4	201.6	197.7	135.1	138.6	377.9	322.3	159	221.9	174.3	80.6	916.4	283.9	321.7	244.1	0.778855474
Ga0436389_037_17206_17775	436.8	248.5	243.1	113	107.7	90	388.8	399.2	203.5	671	359.4	167.4	325.2	145.1	255.7	750.8	207.9	253.1	0.50920719
Ga0436389_037_18956_20152	1139.4	992	970.1	1058.2	1129.6	988.8	1523.3	1369.5	1084.1	1012	1276.7	1137.1	2044.7	1853.1	653	1334.7	841.1	834.8	0.595324896
Ga0436389_037_1977_2906	145.9	242.9	206.4	88.3	106.7	82.1	237.4	253.1	171	262.2	175	147.9	245.6	175.5	99.2	315.3	164.3	164.8	0.298767254
Ga0436389_037_25191_26810	136.6	259.2	219.9	129.5	185.6	176	159.9	187.2	580.3	117.8	132.1	638.8	172.7	180.6	278.8	114.5	282.4	499.1	0.907655706
Ga0436389_037_28018_29223	3563.5	2940.2	2850	1728.9	1487.6	1395.5	1661	1608.5	1231.8	1523	1670.5	1455.4	1374.7	1421.7	1038.6	1666.5	1341.8	1364.9	0.000171561
Ga0436389_037_2903_3931	37.4	129.9	142	63.5	149.3	152.4	22.6	50.9	222.5	27.6	49.9	213.9	27	46.6	113	15.2	165.4	244	0.93388418
Ga0436389_037_29285_29752	751.4	756.9	755.2	502.1	722.7	760.4	1052.1	1257.8	907.9	1180.9	1310.5	1175.2	395.7	450.6	650.1	562.1	828.3	518.3	0.001733473
Ga0436389_037_3_1904	642.7	717.7	725.5	1077.2	1402.5	1342.1	682.4	845	1157.2	601.9	758.2	683.8	446.5	553.3	1014.9	534.6	978.3	727.6	0.063328851
Ga0436389_037_31804_32256	289.2	116.3	105.6	265.4	138.2	123.8	354.5	262.7	149.8	274.3	204.5	88.9	365.4	485.2	323.1	375.6	242.5	281.4	0.148860042
Ga0436389_037_32240_32641	114	54.2	54.2	101.8	69.2	74.4	53	69.4	52.4	35.5	89.9	52	34.1	53.7	43.6	11	38.2	28.1	0.15484998
Ga0436389_037_37116_37610	636.1	687.2	671	373.5	682.3	718.7	291.1	344.2	486.2	223.4	348.6	630.6	259.6	270.8	534.1	250.6	476.9	475.7	0.214132572
Ga0436389_037_37616_38176	1349.3	1121.2	1094.9	1373	1308.4	1312.4	1109.9	1129.7	1136.5	981.1	1041.7	985.8	1033	1005.9	951.4	888.9	1279.5	1025.5	0.049687943
Ga0436389_037_38288_39391	970.7	851.6	850.2	763.7	945.1	989.6	1041.2	1007.7	748.9	955.9	1031.7	836.9	1272.1	922	658.2	823.2	641	659	0.620974786
Ga0436389_037_40445_41416	6974.6	5957.8	5974.8	8179.9	6750.6	5960.4	5398.7	6209.1	4296.7	4672	6058.5	4945.2	3282.1	3803.4	3767.4	3973.3	5079.9	3556.6	0.013756859
Ga0436389_037_4250_5704	1705.1	1505.1	1384	2813.3	2971.4	2957	1824.3	2135.4	1606.9	1247	1770.3	1597.3	1608.4	1445.7	1469.1	1376.5	1777.9	1222.5	0.000354976
Ga0436389_037_44026_44367	258.3	243.2	203.2	257.9	342.6	337.2	535.8	635	619.7	507.8	519.9	495.9	529.9	431.2	609.2	609.4	484.6	679.5	0.000810522
Ga0436389_037_45677_46801	229.1	124.5	123.7	231.3	123.9	106.5	202.5	204.2	288.1	208.1	182.2	203.9	169.2	159.6	154.1	144.3	191.8	280.9	0.536635729
Ga0436389_037_46801_47529	540.7	678.1	671.1	337.6	600.9	552.5	744	694.8	580	666.8	575.9	441.7	632.1	603.1	595.5	650.3	602.6	440.9	0.526225136
Ga0436389_037_47667_49124	65.7	173.1	155.9	112.3	100	127.2	98.6	106.1	42.9	84.3	101.8	36.7	85.8	104.9	68.6	154.7	70.2	40.4	0.649387493
Ga0436389_037_49130_50209	1029.7	2021	1945.4	961.8	1626	1633.2	1541.2	1577.2	1749.3	1562.2	1205.2	1737	2716.6	2399.8	1804.6	1797.2	1392.1	1359.4	0.210500726
Ga0436389_037_53784_54287	517.7	383.4	346.9	265	373.9	348.9	297.9	266.6	215.4	191.4	334.6	323	157.8	203.2	209.4	114.5	188.9	153.7	0.013296234
Ga0436389_037_54429_55829	122.5	102.3	100.4	107.4	88	72.2	89.5	88.6	240.5	64	99.8	121.3	48.5	69.5	148.6	38.3	114.7	174.4	0.885895212
Ga0436389_037_5741_6487	338.8	455.7	430.6	253.9	505.5	448.9	458.3	697.7	605.9	1369.8	915.8	669.6	834.9	429.8	264.1	1669.5	346.5	554.6	0.440599143
Ga0436389_037_58993_59619	102.9	139.2	136.5	83	117.4	105	249.8	263.2	107.7	82.3	74.5	115	405.7	363.4	45.3	29.1	81.7	81.5	0.210351732
Ga0436389_037_60402_61052	652.6	726.2	610.6	1166	833.7	948	823.3	939.7	790	630.2	849.7	858.9	893.3	646.9	462.1	588.1	1012.5	602.3	0.356524587
Ga0436389_037_62088_63002	1885.7	1257.5	1110.6	529.1	540.8	552.1	725.3	749.3	725.3	746.9	685.9	634.6	848.2	674.6	768.9	864	522.4	815.8	0.011410112
Ga0436389_037_63067_65535	83.2	144.6	133.4	87	123.5	101.5	766.8	750	1160.5	869.4	494.6	670.9	731.4	301.4	388.8	912.9	697.3	1451.3	0.006888804

Table B.21 continued from previous page

Protein Accession	NoN1	NoN2	NoN3	NO$_3^-$1	NO$_3^-$2	NO$_3^-$3	N1	N2	N3	R1	R2	R3	P1	P2	P3	E1	E2	E3	Q Value
Ga0436389_037_6596_7969	63.4	314.9	294.1	43.9	269	232.5	346.6	418.8	954.3	671.1	400.9	697.2	547	293.4	475.6	1020.3	1426.9	1752.4	0.003842314
Ga0436389_038_14437_15219	16.9	60.9	66.8	5.6	78.3	66.4	70.7	98.3	462.6	176.1	95.6	541.9	100.7	41.6	117.3	352.1	220.7	664	0.204800703
Ga0436389_038_18685_19071	539.8	290.7	280.7	371.7	214.9	212.4	267	243.9	276.8	197.5	216.9	191.5	256.8	342.6	422.3	167.1	204.7	148	0.159393564
Ga0436389_038_21532_23262	38.7	39.5	39.6	28.3	23	26.7	83.8	76.7	37.5	73.2	55.3	29.1	280	127.4	39.7	89.7	42.4	32.1	0.272574478
Ga0436389_038_24_332	102.9	59.5	55.5	58.7	46.1	43.9	27.4	41.1	37.4	31	33.9	50.4	15.2	23.4	20.8	11.6	20.7	26.4	0.014478567
Ga0436389_038_338_805	20.8	30.7	32.2	16.8	16.2	19.2	31.7	41.3	18.2	39.4	26.3	15.9	64.3	43.6	24	58	230.5	222.5	0.025629089
Ga0436389_038_35516_39118	91.1	43.7	37.6	87.6	28.1	28.3	58.3	76.8	60.7	36.2	37.6	53.9	75	50.4	51.5	47.5	36.2	51.3	0.802082179
Ga0436389_038_42686_43543	126.8	49.8	48.6	6	2.9	4.4	59.5	41.7	13	82.5	53.5	32	126.3	101.5	62.8	66.4	58.2	60.6	0.058925315
Ga0436389_038_47586_49181	81.5	128.1	123.6	41	117.8	93.8	150.5	130.4	208.3	102.6	60.7	78.6	170.3	166.7	368.4	176	223.6	135.9	0.099464554
Ga0436389_038_5609_6577	82.7	56.7	52.9	59.8	54.5	56.3	114.3	155	118.1	143.8	94.9	73.7	109.5	70.5	78.1	171.6	82.1	121.6	0.082892106
Ga0436389_038_60913_61758	283.1	153.9	152.9	200.3	175.5	186.4	583.1	422.8	156.6	564.6	437.3	161.6	605.1	476.4	76.1	833.4	114.2	111.6	0.84461268
Ga0436389_038_63285_63647	3913.8	2445.6	2399.6	2909.8	3227.3	3185.4	2337.6	2583	1769.2	2187.1	2765	1947.5	1938	1903.3	1252	1976.8	3269.6	2845.2	0.15225946
Ga0436389_038_8106_9233	649	863.6	932	567.7	1071.6	954.2	1335.3	1307.3	932.1	1184.6	769.9	752.8	2636.1	1957.3	861.5	1299.8	1168.9	1217.9	0.187137734
Ga0436389_040_6902_7993	6235.9	4162.5	3910.9	1148.9	4180.7	4176.5	5546.6	6884	4939.3	2873.1	6955.3	5592.9	5593.3	6710.5	4345.5	5492.8	5702.6	4283.3	0.473658566
Ga0436389_041_15810_16178	1944.3	1960.5	1849.2	1190.5	1570	1640.4	905.4	896.6	887.4	693.2	775.6	832.9	564	606	767.5	924.9	843.4	840.7	7.66039E-06
Ga0436389_041_18408_18683	2652.9	1539.9	1522	1867	1490.9	1599.4	1639.9	1603.4	1411.9	1123	1915.5	795.1	1680	1965.9	1595.1	1041	741	501.2	0.105762507
Ga0436389_041_2_2542	1257.1	2009.3	2035.2	1293	2244.6	2261.5	1309.2	1461.1	2020.3	1033.6	1203.2	1981.1	1410.2	1316.6	2418.5	1256	3166.3	2508.8	0.703010602
Ga0436389_041_26392_27222	182.2	399.3	374.9	110.1	273.2	265.4	94.4	70.1	512.2	60.1	62.4	469.8	43.6	41	201.9	47.1	345.7	527.4	0.806884893
Ga0436389_041_29844_31208	143.4	282.1	283.6	115.2	305.7	299.9	338.6	437	780.6	659.5	437.5	575.2	419.1	270.2	775.1	940.8	542.1	727.8	0.092484161
Ga0436389_041_39208_39861	174.1	439.1	420.8	138	595	536.5	109.2	61.7	238	17.8	23.7	137.7	202.5	182.7	420.8	76.3	437.7	292.5	0.242389468
Ga0436389_041_4021_4233	255.8	450.2	394.2	247.9	489.8	461.9	209.8	226.8	522.2	157.4	195.8	556.9	191.7	175.8	307.2	158.3	372.9	485.5	0.858546293
Ga0436389_042_304_510	2174.9	793.1	802.7	424.5	327.6	236	482.4	458.3	175.8	533	655.4	267.2	182.5	334.5	108.5	153	131.5	51.1	0.061794365
Ga0436389_044_15264_16370	278.6	413.8	440.6	241.9	361.8	299.7	214.4	202.7	718.7	202.1	195	434.9	155.3	184.6	841.1	211.5	591	658.5	0.931740959
Ga0436389_044_17931_19019	1535.5	1406.7	1405.8	1105.8	1229.8	1116.9	992.8	1082.6	1193.8	1017.3	1116.8	1060.5	868.1	986.9	1286.7	982.8	1021.5	1169.6	0.034253521
Ga0436389_044_22522_23193	1187	1185.5	1254.9	1163	1180.5	1135.3	1362.1	1351.8	1211.7	929.1	990.5	1032.9	1443.3	1367.7	1233.2	967.2	1300.4	1075.8	0.024882622
Ga0436389_044_23310_23921	303.2	336.6	325.4	309	326.8	308.7	254.6	293.9	263.8	190.8	277.5	247.4	177.8	198	197.3	150.3	238.5	201	0.00283184
Ga0436389_044_23975_24970	3134.6	2190.9	2170	3333.4	2314.9	2186.9	2292.1	2586.3	1788.5	1765.4	2444	1744.5	2557.3	2520.4	2213.2	2109.5	1880.2	1336.2	0.389602554
Ga0436389_044_3292_3540	115.3	325.6	353.3	78.4	276.5	273.7	103.1	88.8	202.5	97.6	100.4	176.5	153	143.3	238.9	170.8	354.6	318	0.369854624
Ga0436389_044_3813_5120	47.2	110	96.6	43.2	113.5	115	146.4	90.9	151.5	97.1	80.8	70.1	231.7	201.1	241.6	88.8	197.8	136	0.015360078
Ga0436389_044_40_924	21.5	241.8	222.7	21.6	363	332.4	54.2	43.9	534.9	92.1	40.6	456.4	53.2	33.6	200.4	139.8	327	459.5	0.877551687
Ga0436389_044_5223_6404	330.9	173.9	172.6	277.4	135.5	138	317.5	363.3	222.7	548.5	396.1	140.8	332.6	228.2	534.5	699.2	252.9	192.4	0.719114302

Table B.21 continued from previous page

Protein Accession	NoN1	NoN2	NoN3	NO_3^-1	NO_3^-2	NO_3^-3	N1	N2	N3	R1	R2	R3	P1	P2	P3	E1	E2	E3	Q Value
Ga0436389_044_7743_8291	37.2	84.6	69.3	35.7	83.5	100.2	42.6	29.3	82.5	37.3	37.2	74.7	51.6	29.7	90.9	184.4	109.7	110.9	0.111963966
Ga0436389_044_926_2398	1523	2730.7	2586.4	1050	2004.7	1869	1307.9	1232.2	1632.6	931.8	1030.5	1760.8	1316.7	1501.5	1354.8	1047.3	2382.2	1989	0.321568703
Ga0436389_045_2298_4007	2489.7	2428.3	2348.9	1412.3	1417.2	1394.4	1681.4	1889.2	2490.6	1669.6	1514.6	1985.9	1724.3	1680.3	1685.5	1701.3	1715.7	2305.1	0.025707545
Ga0436389_046_16533_17252	221.1	279.4	245.9	252.4	189.7	162.6	199.7	256.4	195.3	240.7	281.2	304.2	158.8	126.4	150.7	196.8	174.8	174.9	0.016562088
Ga0436389_046_22209_22643	219.8	282	265.7	131.4	166.6	171.5	99.1	104.3	96.5	87.4	113	119.7	209.6	138.9	69.1	63.3	62.8	53.6	0.00283184
Ga0436389_046_28833_30602	8289.6	8017.2	7693.6	961.7	2195.6	1632	2270.8	2268.6	2752.6	2124.3	2407.8	2794.4	2154.6	2246	2753	1539.1	2535.4	2330.7	7.09692E-07
Ga0436389_046_30897_31103	1129.4	829.7	866.7	284.3	317	236.7	477.1	399.1	231.6	346.1	393.5	281.3	486.9	493.5	160.4	207.4	130.1	124.4	0.001077545
Ga0436389_046_3359_3736	1743.7	2282.3	1790.8	1451.4	2737.2	3005.9	1346.9	1480.5	1417.2	1238.1	1719.9	2147.6	1140.6	1388.5	1500.2	1173.8	1664.5	1471.2	0.162677093
Ga0436389_046_5566_6492	188.6	349.7	343.7	203.3	419	492.7	299.2	211.7	321.4	265.6	247.4	445.7	202.6	164.4	265.4	171.1	471.8	275.4	0.746686275
Ga0436389_046_6649_8010	3.3	85	93.6	3.9	82.1	72.1	244	250.6	710.6	382.9	190.7	552.8	108	49.5	169.3	315.4	468.2	929.7	0.071433399
Ga0436389_046_9521_11131	1238.6	1090.8	1020.8	1251.1	1011.2	1062.3	1220.5	1249.1	1673.2	759.5	945.4	1412.6	1207.8	1173.6	824	1019.6	1056.9	1019.8	0.50920719
Ga0436389_047_10386_11198	286.3	390.2	358.7	231.6	236.1	265.6	552.9	544.1	634.6	541.1	418.4	531.6	395.9	321.1	461.3	534.1	508.5	731.7	0.003464828
Ga0436389_047_105796_109314	118.2	63.5	71.2	207.5	49	59.6	77.3	70.7	20.2	49.8	57.3	38.9	78.9	103.8	24.1	78.3	23.8	15.6	0.662091974
Ga0436389_047_112270_112707	279.3	91.5	94.1	70.4	51.5	42.6	98.1	68.4	97.2	75.5	69.4	68.2	82.5	68.4	199.3	50.8	96.9	56.7	0.440599143
Ga0436389_047_117992_119026	871.6	264.6	253.1	807.4	331.1	272.5	680.9	737.8	233.9	702.2	773.4	204.2	579	606.4	246.2	798	326.3	219.2	0.994167314
Ga0436389_047_122388_123812	21.5	561.1	519.9	7.5	393.6	385.5	133.5	95	1011	181.7	95.8	653.4	103.2	33	1027.8	261	857.2	900.9	0.888484534
Ga0436389_047_129806_130741	655.9	441	474.4	1596.2	1146.6	1676.9	2516	2636.5	1425.5	3522.3	3135.4	1716.7	2217.3	1936.7	835.8	2418.7	1351.4	1543.8	0.053564664
Ga0436389_047_135731_136105	857.9	369.3	309.3	762.2	344.9	279.3	605.8	627.3	619.7	858.1	780.3	584.9	752	643.2	392.9	1210.8	397.6	930.9	0.609161724
Ga0436389_047_136639_137808	1161.3	1026.9	1012.4	962.5	869.5	802.1	1070.7	1131.1	1051.3	844.1	1209.3	1111.9	691.8	926.7	596.4	1248.9	698.8	854.5	0.260291619
Ga0436389_047_143128_146160	88.3	249.9	222.8	60.5	191.3	174.5	504.7	685.7	563.6	1171.5	606.7	469	743.5	325.6	237.5	1378.3	493.5	729.8	0.092739695
Ga0436389_047_148951_149382	4760.4	4339.7	4322.8	3892.1	5654.3	5914.9	2818.1	2872.9	2878.6	2409.4	2885.1	2952.8	2691.9	2634.8	3931.9	1858.3	3063	2204.9	0.005942455
Ga0436389_047_152153_153259	195.1	179	159.3	149	281.7	264.7	120.5	97.6	258.2	103.4	94.2	242.5	57.6	100.6	139.7	230.9	226.3	265.8	0.211224492
Ga0436389_047_153453_154478	208.3	330.9	331.2	159.2	398	352.6	350.4	333.9	305.4	192.8	181.3	266.6	445.8	447.7	326.9	239.8	306.4	300.1	0.200619907
Ga0436389_047_155929_156594	377.4	440.8	425.8	342.5	370.1	346.6	345.9	344.6	219.5	236.7	295.7	247.1	524.2	445.8	217.8	284	218.4	135	0.142280758
Ga0436389_047_162129_163163	97.6	55.9	44.6	72.4	54.8	39.7	200.5	196	152.1	227.5	184.5	144.7	347.6	282.3	41.3	237.4	60.1	126.2	0.227140403
Ga0436389_047_166282_166944	103.6	126.1	121.7	79.2	116.4	113.9	201	188.2	160.6	158.2	112.9	136.7	158.8	261.7	217.3	202.1	236.7	270.1	0.005401718
Ga0436389_047_1764_2963	295.8	96.1	103.7	431.3	160.2	112.4	274.5	304.8	96.8	257.5	193.1	124.3	211.6	257.2	77.1	570.3	194.4	196.4	0.847671069
Ga0436389_047_24055_25611	44.5	40.6	37.6	39.2	39.1	37.1	60.4	53.7	32.6	63.3	49.2	21.9	47.8	30.2	92.7	95	28.1	16.1	0.962382441
Ga0436389_047_3494_3745	112.9	52.6	52.2	152.3	73.5	53.5	145.9	150.3	54.8	164.7	200.6	72.8	131	105.9	37.2	98.9	41.5	47.3	0.536635729
Ga0436389_047_37485_38417	80.9	115.4	121.4	113.8	105.8	91.6	164.4	129.3	123.5	91.1	93.3	102.3	249.3	232.3	135.2	138.3	108.8	82.6	0.034253521
Ga0436389_047_40433_41332	26.3	56.6	42.9	23.6	29.5	37.2	70	66	43.6	61.3	38.4	38.7	68	85	62	71.9	39.3	18.4	0.183472688

Table B.21 continued from previous page

Protein Accession	NoN1	NoN2	NoN3	NO$_3^-$1	NO$_3^-$2	NO$_3^-$3	N1	N2	N3	R1	R2	R3	P1	P2	P3	E1	E2	E3	Q Value
Ga0436389_047_44242_46971	77.9	275.5	241.4	77.1	223.4	221.7	149.9	136.9	580.5	78.6	64.7	467	194.3	197.4	434.9	67.1	451.2	638.3	0.857619793
Ga0436389_047_55922_56941	112.2	389.5	386.1	72.9	192.5	158.5	104.5	74.5	185.9	87.8	68.9	180.7	98.4	182.3	131	153.6	141.7	114.1	0.232085881
Ga0436389_047_66740_67207	421.5	1001	944.4	572	1308	1385.5	579.1	643.8	1043.8	649.1	587.2	1185.9	532.9	491.2	1419.5	619.5	1084.8	761.9	0.923665272
Ga0436389_047_98548_99237	236.8	318.4	309.6	229	310.9	307.8	132.4	126.3	413.4	87.8	138.2	447.2	146.2	143.4	767.6	79.4	380.3	324.9	0.973155451
Ga0436389_047_99576_102875	94.1	286.9	306.5	90.1	73.5	66.2	112.1	104.6	138	89.2	92.9	101.6	146.1	175.8	108.1	95.3	233.9	218.5	0.157541873
Ga0436389_048_51_539	3652	3541.3	3433.7	7862.3	7142.1	7541	4935.7	5393.2	5505.9	5139.8	6327.4	7153.9	4745.9	3797.8	3682	4396.1	4986.5	4642.4	0.000289927
Ga0436389_048_600_749	67	32	29.7	114.2	98.9	79	182.2	211.4	58.1	207.9	174.7	85.4	132.6	110.2	27.9	192.9	45.1	44.3	0.414364868
Ga0436389_049_407_1258	2356.7	2380.4	2351.2	5720.1	5278.6	5754.9	4181	4666.3	3533.6	3726.3	4279.7	4333.2	3809.9	3767.9	2663.4	3324.3	3591	3261.2	0.000354976
Ga0436389_050_11625_12095	119.7	146.8	142.4	83.4	168.3	133	612.7	455.8	565.2	109.7	97.8	217.2	142	134.2	140.2	133.2	193.6	325.3	0.000354976
Ga0436389_050_12238_13098	1394.8	1441	1421.6	1374.5	1199.2	1224.9	1551.5	1725.7	1543.1	1766.3	1629.9	1682.1	1621.3	1315.9	1256.3	1651.1	1333	1342.3	0.054350005
Ga0436389_050_18118_18513	242.3	86.4	74.2	131.2	58.5	65.5	142.9	128.1	135.8	84.3	93.1	106.7	79.9	87.3	52.5	512.7	93.9	148.4	0.507445157
Ga0436389_050_18519_19343	333.3	400.9	391	989.8	328	331.2	756.5	1008	1566.5	2450.3	1841.4	1933	850.9	665.5	341.2	1298.1	610.7	1488.5	0.005577792
Ga0436389_050_19443_20441	183.5	471.7	436	125.7	268.3	249.8	274.8	285	290.7	320.3	189.9	258.7	272.4	295.2	338.4	320.3	257.1	243.6	0.526516592
Ga0436389_050_24975_25649	266.2	242.7	211.7	509.3	256.7	304.4	227.7	300.8	222.9	292.9	301.8	220.1	272.3	182.7	116.2	288	389.5	239.1	0.351702975
Ga0436389_050_25707_26189	362.7	395.2	338.4	734.2	529	458.6	620.2	674.3	317	318.9	482.7	411.6	688.8	627.3	192	422.9	318.1	261	0.518376507
Ga0436389_050_490_747	1642.3	1560.1	1353.9	1906	2603.7	2706.7	1725.9	1711.1	1588.7	1410.2	2000.1	1975.2	1165.7	1395.2	1472.3	1011.2	1828.7	1468.6	0.03984082
Ga0436389_050_5813_7072	2301.7	1535.9	1446.8	972	749.6	668.8	1077.9	1037.2	941.6	1109.2	995	902.8	1426.9	1289.8	844.6	1013.3	851.5	895.3	0.027686861
Ga0436389_050_791_1636	2974.6	3807.5	3552.2	4759.4	4894.5	5026.6	3269	3746.5	3546.4	2885.4	3606.8	4084.3	3140.4	3017	2550.2	2538.7	3907.1	3521.6	0.018452092
Ga0436389_050_7931_9007	3920.1	3222.1	3034.3	3915.9	4824.4	4710.9	3804.6	3881.9	3336	2688.6	3112.9	2753.2	3564.7	4295.4	2145.4	2606.3	2568.9	2060.1	0.045470018
Ga0436389_050_9135_9548	53.7	65.7	68	53.1	40.6	40.7	74.2	58.1	27	77.8	54.2	26.3	85.3	70.6	28.5	94.1	34.4	22.2	0.960969675
Ga0436389_050_9873_11432	754.3	939.3	929.5	1680.3	1525.3	1891.4	910.7	775.5	1014.5	602.1	913.6	1176.7	661	604.4	974.3	595.1	1565.1	1204.1	0.03984082
Ga0436389_051_14205_14576	1003.4	631.2	629	172.7	226.4	237.6	363.9	325.2	197.6	266.6	203.3	148	389.9	365.6	214.1	242.7	153.9	159.5	0.002862583
Ga0436389_051_19732_19947	50.2	113.4	90.9	56.3	59.5	45.7	57.7	63.4	45.8	71	181.2	186.8	80.5	46.7	31.3	65.1	21.4	19.5	0.058925315
Ga0436389_051_24078_24575	123.4	264.8	263.9	191.5	27.6	27.2	130.4	110	129.4	337.1	438.3	657.3	182.8	94	117.7	206.2	48.4	80.7	0.013073998
Ga0436389_051_45516_45962	1756.5	1130.1	987	427.5	306.6	300.6	625.6	550.8	229.5	625.5	455.9	221.9	481.7	511.5	268.4	753.9	226.1	166	0.023400743
Ga0436389_051_70097_71521	207.9	227.3	234.7	414.5	339.4	303.4	293.7	403.8	256.3	302.8	318.7	247.4	440.2	353.3	111.2	369.8	216.5	155.8	0.706399868
Ga0436389_051_72247_74799	2266.3	1830.5	1780.8	2513.3	2563.1	2461.6	3367.2	3421.1	2774.8	3604.6	3087.4	2162.8	3936.8	3537.3	2906.5	5716	2328.5	2222.3	0.50228876
Ga0436389_053_3_782	25.6	28	31.6	50.1	29.9	26.1	77.4	110.9	18	107.7	99.6	36.3	98.7	73	5.6	124	74.8	74.5	0.380246122
Ga0436389_070_2_241	204.2	279.7	276.4	47.8	130.2	109.2	126.2	112	127.2	38.4	55.2	136.2	309.3	438.7	329.9	41.9	139.2	134.4	0.001054454
Ga0436389_074_3_368	3110.1	2303.3	2186.3	5558.2	4208.8	4357.9	3549.6	3858.2	2095.6	2890.3	4956.6	3070	2320.7	2548.3	1494.4	2006.6	2054.9	1472.5	0.025479647
Ga0436389_079_2_553	360.9	640.9	583	156.5	305.4	250	332.1	356.1	768.3	421.2	310.3	503.8	359.5	312.1	394.6	553.5	387.1	549.1	0.281752357

Table B.21 continued from previous page

Protein Accession	**NoN1**	**NoN2**	**NoN3**	**NO_3^-1**	**NO_3^-2**	**NO_3^-3**	**N1**	**N2**	**N3**	**R1**	**R2**	**R3**	**P1**	**P2**	**P3**	**E1**	**E2**	**E3**	**Q Value**
Ga0436389_083_3_254	222.9	184.1	171.6	678.6	425.8	473.1	202	242.7	174	151.8	364.7	266.1	225.6	162.4	104.5	152.9	285.8	180.9	0.010479808
Ga0436389_084_1_297	168.4	1603.7	1544.5	47.8	696.8	470.2	85.5	77.3	397.4	116.8	168.2	659.1	64.4	77	290.2	28.2	274.7	218.7	0.178013453
Ga0436389_096_1_843	745.7	546.3	524.3	300.6	338.8	270.7	421.5	456.4	468.2	433.7	346.9	417.8	534.2	594	361.6	401.4	334.9	415.2	0.032296347

Table B.22: **Volcano plot top hits for *Synechococcus* sp. CCAP1479/10 growth on amino acid nitrogen-substrates v. NO_3^-**. Top 10 hits based on manhattan distance (fold change + significance). Volcano plots are found in Appendix Figure A.15.

Change	Protein Accession/Name	Description	COG
Arginine v NO_3^-			
Increase	Ga0436389_005_2951_3700	DUF305 domain-containing protein	S
Increase	MinD	septum site-determining protein MinD	D
Increase	Ga0436389_051_24078_24575	DUF411 domain-containing protein	S
Decrease	UrtA	urea ABC transporter substrate-binding protein	E
Decrease	PetJ	c-type cytochrome	C
Decrease	Tig	trigger factor	D
Decrease	Ga0436389_041_15810_16178	AbrB family transcriptional regulator	K
Decrease	NrtA	ABC transporter substrate-binding protein	P
Decrease	GlnN	glutamine synthetase III	S
Decrease	Ga0436389_006_276270_277475	acetamidase/formamidase family protein	C
Asparagine v NO_3^-			
Increase	Ga0436389_050_11625_12095	2Fe-2S iron-sulfur cluster binding domain	C
Increase	FtsH1	ATP-dependent zinc metalloprotease FtsH	O
Increase	Ga0436389_047_10386_11198	META domain-containing protein	O
Increase	ManC	mannose-1-phosphate guanylyltransferase	M
Increase	AroG	3-deoxy-7-phosphoheptulonate synthase	E
Increase	ParC	DNA topoisomerase 4 subunit A	L
Decrease	PetJ	c-type cytochrome	C
Decrease	UrtA	urea ABC transporter substrate-binding protein	E
Decrease	NrtA	ABC transporter substrate-binding protein	P
Decrease	Tig	trigger factor	D
Glutamate v NO_3^-			
Increase	Ga0436389_006_247274_249397	transglutaminase family protein	E
Increase	TatA	TatA/E family twin arginine-targeting protein translocase	U
Increase	MsrB	peptide-methionine (R)-S-oxide reductase MsrB	O
Increase	ClpX	ATP-dependent protease ATP-binding subunit ClpX	O
Decrease	PetJ	c-type cytochrome	C
Decrease	UrtA	urea ABC transporter substrate-binding protein	E
Decrease	Tig	trigger factor	D
Decrease	PsbO	photosystem II manganese-stabilizing polypeptide	S
Decrease	NrtA	ABC transporter substrate-binding protein	P
Decrease	GlpX	class II fructose-bisphosphatase	G
Proline v NO_3^-			
Increase	Ga0436389_024_37846_43668	esterase-like activity of phytase family protein	S
Increase	PurH	Bifunctional purine biosynthesis protein PurH	F
Increase	Amt1	Ammonium Transporter Family	P
Increase	ManC	mannose-1-phosphate guanylyltransferase	M
Increase	KatG	catalase/peroxidase HPI	P
Decrease	PetJ	c-type cytochrome	C
Decrease	AtpA	F0F1 ATP synthase subunit alpha	C
Decrease	CpcB	phycocyanin subunit beta	C
Decrease	CbbA	fructose-bisphosphate aldolase class II	G
Decrease	Ga0436389_041_15810_16178	AbrB family transcriptional regulator	K

Table B.23: **Volcano plot top hits for *Synechococcus* sp. CCAP1479/10 growth on amino acid nitrogen-substrates v. nitrogen-starvation**. Top 10 hits based on manhattan distance (fold change + significance). Volcano plots are found in Appendix Figure A.16.

Change	Protein Accession/Name	Description	COG
Arginine v N-starvation			
Decrease	Ga0436389_020_45810_46403	hypothetical protein	-
Decrease	Ga0436389_026_21467_21841	VOC family protein	S
Decrease	Ga0436389_006_276270_277475	acetamidase/formamidase family protein	C
Decrease	Ga0436389_046_28833_30602	carbohydrate porin	S
Decrease	AstA	arylsulfatase	P
Decrease	UrtA	urea ABC transporter substrate-binding protein	E
Decrease	Ga0436389_041_15810_16178	AbrB family transcriptional regulator	K
Decrease	SbpA	sulfate ABC transporter substrate-binding protein	P
Decrease	Ga0436389_006_211053_212597	metallophosphoesterase / Calcineurin-like phosphoesterase protein	S
Decrease	CmpA	ABC transporter substrate-binding protein	P
Asparagine v N-starvation			
Decrease	Ga0436389_020_45810_46403	hypothetical protein	-
Decrease	Ga0436389_026_21467_21841	VOC family protein	S
Decrease	Ga0436389_006_276270_277475	acetamidase/formamidase family protein	C
Decrease	Ga0436389_046_28833_30602	carbohydrate porin	S
Decrease	AstA	arylsulfatase	P
Decrease	UrtA	urea ABC transporter substrate-binding protein	E
Decrease	Ga0436389_022_205356_205733	hypothetical protein	-
Decrease	Ga0436389_041_15810_16178	AbrB family transcriptional regulator	K
Decrease	Ga0436389_006_177279_178604	ScyD/ScyE family protein	G
Decrease	PetF	2Fe-2S iron-sulfur cluster-binding protein	C
Glutamate v N-starvation			
Increase	Ga0436389_006_247274_249397	transglutaminase family protein	E
Decrease	Ga0436389_020_45810_46403	hypothetical protein	-
Decrease	Ga0436389_026_21467_21841	VOC family protein	S
Decrease	Ga0436389_046_28833_30602	carbohydrate porin	S
Decrease	Ga0436389_006_276270_277475	acetamidase/formamidase family protein	C
Decrease	AstA	arylsulfatase	P
Decrease	UrtA	urea ABC transporter substrate-binding protein	E
Decrease	GroS	co-chaperone GroES	O
Decrease	SbpA	sulfate ABC transporter substrate-binding protein	P
Decrease	GlnB	P-II family nitrogen regulator	K
Proline v N-starvation			
Increase	Ga0436389_024_37846_43668	esterase-like activity of phytase family protein	S
Decrease	Ga0436389_020_45810_46403	hypothetical protein	-
Decrease	Ga0436389_026_21467_21841	VOC family protein	S
Decrease	Ga0436389_006_276270_277475	acetamidase/formamidase family protein	C
Decrease	Ga0436389_046_28833_30602	carbohydrate porin	S
Decrease	AstA	arylsulfatase	P
Decrease	Ga0436389_041_15810_16178	AbrB family transcriptional regulator	K
Decrease	GroS	co-chaperone GroES	O
Decrease	UrtA	urea ABC transporter substrate-binding protein	E
Decrease	Ga0436389_006_177279_178604	ScyD/ScyE family protein	G

Table B.24: ***Synechococcus* sp. CCAP1479/10 DEPs when grown on arginine v. NO_3^-**. All proteins were detected in triplicate using TMT quantitative proteomics. Differential expression was determined through FDR-adjusted ANOVA and Tukey's test. $\log_2$FC: $\log_2$ fold change.

Change	Protein Accession	Gene	ANOVA (Q)	Tukey (Q)	$\log_2$FC	Description	COG
Increase	Ga0436389_006_177279_178604	-	1.84716E-05	0.023630171	1.81	ScyD/ScyE family protein	G
Increase	Ga0436389_037_44026_44367	-	0.000810522	0.023710917	0.658	YbaB/EbfC family nucleoid-associated protein	S
Increase	Ga0436389_037_29285_29752	-	0.001733473	0.002146619	0.928	hypothetical protein	S
Increase	Ga0436389_005_2951_3700	-	0.00283184	6.15262E-05	1.86	DUF305 domain-containing protein	S
Increase	Ga0436389_047_10386_11198	-	0.003464828	0.00922001	0.901	META domain-containing protein	O
Increase	Ga0436389_026_246109_248088	*ftsH1*	0.003551791	0.012931663	2.37	ATP-dependent zinc metalloprotease FtsH	O
Increase	Ga0436389_029_46607_47680	*aroG*	0.003736291	0.010621468	2.11	3-deoxy-7-phosphoheptulonate synthase	E
Increase	Ga0436389_050_18519_19343	*minD*	0.005577792	0.001593595	2.05	septum site-determining protein MinD	D
Increase	Ga0436389_051_24078_24575	-	0.013073998	0.002260822	2.05	DUF411 domain-containing protein	S
Increase	Ga0436389_006_282399_284078	*pyrG*	0.013122001	0.02751852	2.44	CTP synthase	F
Increase	Ga0436389_032_35938_37950	*ndhF1*	0.025134825	0.039536187	2.35	NAD(P)H-quinone oxidoreductase subunit 5	CP
Increase	Ga0436389_006_251960_252859	-	0.029633032	0.029275252	0.864	transglutaminase family protein	E
Decrease	Ga0436389_001_54542_55054	-	0.00392427	0.006287851	-0.703	DNA recombination-mediator protein A	LU
Decrease	Ga0436389_004_47991_48395	*rbpD*	0.006370734	0.014424671	-0.964	RNA-binding protein	S
Decrease	Ga0436389_004_64037_64612	-	0.004731823	0.016530013	-0.567	hypothetical protein	S
Decrease	Ga0436389_006_182501_182869	-	0.033343677	0.034508945	-0.58	AbrB family transcriptional regulator	K
Decrease	Ga0436389_032_82878_83618	-	0.002186741	0.024038811	-0.518	MBL fold metallo-hydrolase	S
Decrease	Ga0436389_034_82555_83199	-	0.013756859	0.001475573	-1.03	L,D-transpeptidase	S
Decrease	Ga0436389_035_30354_30695	*rbpD*	0.044973608	0.041370331	-0.768	RNA-binding protein	S
Decrease	Ga0436389_041_15810_16178	-	7.66039E-06	9.7399E-05	-0.794	AbrB family transcriptional regulator	K
Decrease	Ga0436389_022_86912_88222	*glyA*	0.008332703	0.001821073	-0.601	serine hydroxymethyltransferase	E
Decrease	Ga0436389_032_56652_57287	*cysC*	0.001064225	0.011540719	-0.629	adenylyl-sulfate kinase	F
Decrease	Ga0436389_006_276270_277475	-	7.40069E-07	0.000279519	-1.44	acetamidase/formamidase family protein	C
Decrease	Ga0436389_003_41071_43239	*glnN*	0.000160037	0.000233623	-1.37	glutamine synthetase III	S
Decrease	Ga0436389_007_7050_8879	-	0.066603213	0.041490493	-0.705	V-type ATP synthase subunit A	C
Decrease	Ga0436389_008_204171_205112	*petA*	0.003400378	0.007500419	-0.688	apocytochrome f	C
Decrease	Ga0436389_047_148951_149382	*psaD*	0.005942455	0.004666912	-0.767	photosystem I reaction center subunit II	S
Decrease	Ga0436389_050_9873_11432	*psbB*	0.03984082	0.028023679	-0.834	photosystem II chlorophyll-binding protein CP47	P

Table B.24 continued from previous page

Change	Protein Accession	Gene	ANOVA (Q)	Tukey (Q)	$\log_2$FC	Description	COG
Decrease	Ga0436389_007_151376_151486	*psbD*	0.02020216	0.021563037	-0.826	Photosystem II DII subunit, Q(A) protein	C
Decrease	Ga0436389_083_3_254	*psbD*	0.010479808	0.017554315	-0.887	photosystem II D2 protein (photosystem q(a) protein)	C
Decrease	Ga0436389_022_258493_259305	*psbO*	0.000875663	0.003479142	-0.535	photosystem II manganese-stabilizing polypeptide	S
Decrease	Ga0436389_004_71443_71565	*psbY*	0.006236576	0.003818481	-0.948	photosystem II protein Y	U
Decrease	Ga0436389_006_67639_68166	*rplJ*	0.017711586	0.019010876	-0.52	50S ribosomal protein L10	J
Decrease	Ga0436389_024_74509_74919	*rplP*	0.019132506	0.030222528	-0.838	50S ribosomal protein L16	J
Decrease	Ga0436389_001_50769_51245	*rplS*	0.010905627	0.002329835	-0.949	50S ribosomal protein L19	J
Decrease	Ga0436389_008_100726_101124	*rplU*	0.01336335	0.00605191	-0.905	50S ribosomal protein L21	J
Decrease	Ga0436389_006_67191_67577	*rplL*	0.00283184	0.006013339	-0.621	50S ribosomal protein L7/L12	J
Decrease	Ga0436389_050_7931_9007	*rps1a*	0.045470018	0.03559398	-0.531	30S ribosomal protein S1	J
Decrease	Ga0436389_024_76096_76371	*rpsS*	0.018780364	0.027931909	-0.706	30S ribosomal protein S19	J
Decrease	Ga0436389_024_19563_20033	*rps7*	0.006236576	0.003951548	-0.594	30S ribosomal protein S7	J
Decrease	Ga0436389_024_72207_72581	*rps8*	0.004776308	0.014460749	-0.646	30S ribosomal protein S8	J
Decrease	Ga0436389_022_100421_100738	*tatA*	3.80939E-06	0.048263313	-0.747	TatA/E family twin arginine-targeting protein translocase	U
Decrease	Ga0436389_037_4250_5704	*tig*	0.000354976	4.27403E-05	-0.81	trigger factor	D
Decrease	Ga0436389_026_300596_301927	*prc*	0.015030853	0.03574391	-0.564	S41 family peptidase	M
Decrease	Ga0436389_034_157455_157766	*ccmK2*	0.00283184	0.034793244	-0.582	BMC domain-containing protein	CQ
Decrease	Ga0436389_034_70361_70702	*petJ*	1.40236E-05	1.25232E-05	-0.919	c-type cytochrome	C
Decrease	Ga0436389_026_120479_121483	*glpX*	0.000389029	0.002250783	-0.571	class II fructose-bisphosphatase	G
Decrease	Ga0436389_006_270480_271775	*urtA*	6.12887E-06	1.61464E-05	-1.7	urea ABC transporter substrate-binding protein	E
Decrease	Ga0436389_037_3_1904	*rnj*	0.063328851	0.03057534	-0.768	RNase J family beta-CASP ribonuclease	J
Decrease	Ga0436389_029_19276_20589	*nrtA*	0.000222036	0.000114249	-1.79	ABC transporter substrate-binding protein	P
Decrease	Ga0436389_008_73445_74017	-	0.000736838	0.017794927	-0.525	peroxiredoxin	O

Table B.25: ***Synechococcus* sp. CCAP1479/10 DEPs when grown on asparagine v. NO_3^-**. All proteins were detected in triplicate using TMT quantitative proteomics. Differential expression was determined through FDR-adjusted ANOVA and Tukey's test. $\log_2$FC: $\log_2$ fold change.

Change	Protein Accession	Gene	ANOVA (Q)	Tukey (Q)	$\log_2$FC	Description	COG
Increase	Ga0436389_001_107869_108948	sbpA	1.84716E-05	0.021087146	1.53	sulfate ABC transporter substrate-binding protein	P
Increase	Ga0436389_001_48625_50067	*gltX*	0.010905627	0.040261702	0.942	glutamate–tRNA ligase	J
Increase	Ga0436389_003_22524_23453	-	0.025134825	0.031575949	1.11	family 2A encapsulin nanocompartment shell protein	T
Increase	Ga0436389_004_46165_47994	-	0.004139985	0.03298931	0.946	ABC transporter	V
Increase	Ga0436389_005_2951_3700	-	0.00283184	0.023589743	1.08	DUF305 domain-containing protein	S
Increase	Ga0436389_008_14236_16140	*rpoC1*	0.00154649	0.003204607	1.01	DNA-directed RNA polymerase subunit gamma	K
Increase	Ga0436389_008_212849_213967	*nfrC*	0.03561983	0.042692852	0.84	UDP-N-acetylglucosamine 2-epimerase (non-hydrolyzing)	G
Increase	Ga0436389_008_242971_244134	*guaB*	0.001077545	0.028769042	0.602	GuaB3 family IMP dehydrogenase-related protein	F
Increase	Ga0436389_008_97722_99221	*kaiC*	0.035792517	0.046100312	0.863	circadian clock protein KaiC	T
Increase	Ga0436389_020_41281_42729	*manC*	0.000558886	0.002930525	1.63	mannose-1-phosphate guanylyltransferase	M
Increase	Ga0436389_022_59508_60977	-	0.013822851	0.009157397	0.831	phosphoglucomutase/phosphomannomutase family protein	G
Increase	Ga0436389_022_77174_78718	*purH*	0.000116372	0.005719125	0.56	Bifunctional purine biosynthesis protein PurH	F
Increase	Ga0436389_024_37846_43668	-	5.55925E-05	0.037264606	2.25	esterase-like activity of phytase family protein	S
Increase	Ga0436389_024_71253_71618	*rplR*	0.014478567	0.016637443	0.905	50S ribosomal protein L18	J
Increase	Ga0436389_025_71305_72870	-	0.02042856	0.012186274	1.26	DUF1957 domain-containing protein	G
Increase	Ga0436389_026_116581_117309	-	0.001077545	0.029513977	0.969	LPS export ABC transporter ATP-binding protein	S
Increase	Ga0436389_026_246109_248088	*ftsH1*	0.003551791	0.001141784	2.64	ATP-dependent zinc metalloprotease FtsH	O
Increase	Ga0436389_026_319856_320947	*acsF*	0.004663153	0.022439127	0.759	magnesium-protoporphyrin IX monomethyl ester	H
Increase	Ga0436389_026_416356_417222	*dapB*	0.025479647	0.003791854	0.744	4-hydroxy-tetrahydrodipicolinate reductase	E
Increase	Ga0436389_029_46607_47680	*aroG*	0.003736291	0.006945769	2.12	3-deoxy-7-phosphoheptulonate synthase	E
Increase	Ga0436389_037_29285_29752	-	0.001733473	0.02172252	0.612	hypothetical protein	S
Increase	Ga0436389_037_44026_44367	-	0.000810522	0.001389699	0.74	YbaB/EbfC family nucleoid-associated protein	S
Increase	Ga0436389_037_63067_65535	*parC*	0.006888804	0.009113149	2.4	DNA topoisomerase 4 subunit A	L
Increase	Ga0436389_047_10386_11198	-	0.003464828	0.000999047	0.973	META domain-containing protein	O
Increase	Ga0436389_050_11625_12095	-	0.000354976	2.87447E-05	1.57	2Fe-2S iron-sulfur cluster binding domain	C
Decrease	Ga0436389_001_54542_55054	-	0.00392427	0.006287851	-0.721	DNA recombination-mediator protein A	LU
Decrease	Ga0436389_004_47991_48395	*rbpD*	0.006370734	0.014424671	-1.27	RNA-binding protein	S
Decrease	Ga0436389_004_64037_64612	-	0.004731823	0.016530013	-0.709	hypothetical protein	S

Table B.25 continued from previous page

Change	**Protein Accession**	**Gene**	**ANOVA (Q)**	**Tukey (Q)**	**$\log_2$FC**	**Description**	**COG**
Decrease	Ga0436389_032_82878_83618	-	0.002186741	0.024038811	-0.759	MBL fold metallo-hydrolase	S
Decrease	Ga0436389_034_82555_83199	-	0.013756859	0.001475573	-0.642	L,D-transpeptidase	S
Decrease	Ga0436389_035_30354_30695	*rbpD*	0.044973608	0.041370331	-0.95	RNA-binding protein	S
Decrease	Ga0436389_041_15810_16178	-	7.66039E-06	9.7399E-05	-0.703	AbrB family transcriptional regulator	K
Decrease	Ga0436389_007_47651_48175	-	0.023400743	0.093029729	-1.24	hypothetical protein	S
Decrease	Ga0436389_025_95217_95657	-	0.001541807	0.733981599	-0.567	cAMP phosphodiesterase	S
Decrease	Ga0436389_026_175590_176126	-	0.034625813	0.105138532	-0.521	hypothetical protein	S
Decrease	Ga0436389_026_294297_294674	-	0.012752538	0.226149263	-1.13	hypothetical protein	S
Decrease	Ga0436389_032_56652_57287	*cysC*	0.001064225	0.011540719	-0.719	adenylyl-sulfate kinase	F
Decrease	Ga0436389_006_47679_48884	*pgk*	0.001280805	0.004344298	-0.676	phosphoglycerate kinase	F
Decrease	Ga0436389_027_49480_51609	*pnp*	0.000538839	0.009470396	-0.56	polyribonucleotide nucleotidyltransferase	J
Decrease	Ga0436389_006_276270_277475	-	7.40069E-07	0.000279519	-1.11	acetamidase/formamidase family protein	C
Decrease	Ga0436389_003_41071_43239	*glnN*	0.000160037	0.000233623	-0.977	glutamine synthetase III	S
Decrease	Ga0436389_008_62779_63327	*atpD*	0.001541807	0.00430454	-0.666	F0F1 ATP synthase subunit delta	C
Decrease	Ga0436389_048_51_539	*cpcA*	0.000289927	0.087442212	-0.512	phycocyanin subunit alpha	C
Decrease	Ga0436389_003_35798_36316	*cpcB*	0.000167728	0.024829542	-0.616	phycocyanin subunit beta	C
Decrease	Ga0436389_008_204171_205112	*petA*	0.003400378	0.007500419	-0.85	apocytochrome f	C
Decrease	Ga0436389_047_148951_149382	*psaD*	0.005942455	0.004666912	-0.827	photosystem I reaction center subunit II	S
Decrease	Ga0436389_050_9873_11432	*psbB*	0.03984082	0.028023679	-0.916	photosystem II chlorophyll-binding protein CP47	P
Decrease	Ga0436389_007_151376_151486	*psbD*	0.02020216	0.021563037	-1.09	Photosystem II DII subunit, Q(A) protein	C
Decrease	Ga0436389_083_3_254	*psbD*	0.010479808	0.017554315	-1.26	photosystem II D2 protein (photosystem q(a) protein)	C
Decrease	Ga0436389_022_258493_259305	*psbO*	0.000875663	0.003479142	-0.722	photosystem II manganese-stabilizing polypeptide	S
Decrease	Ga0436389_004_71443_71565	*psbY*	0.006236576	0.003818481	-0.721	photosystem II protein Y	U
Decrease	Ga0436389_006_67639_68166	*rplJ*	0.017711586	0.019010876	-0.605	50S ribosomal protein L10	J
Decrease	Ga0436389_024_74509_74919	*rplP*	0.019132506	0.030222528	-0.994	50S ribosomal protein L16	J
Decrease	Ga0436389_001_50769_51245	*rplS*	0.010905627	0.002329835	-0.824	50S ribosomal protein L19	J
Decrease	Ga0436389_008_100726_101124	*rplU*	0.01336335	0.00605191	-0.833	50S ribosomal protein L21	J
Decrease	Ga0436389_037_15970_16170	*rpmI*	0.044973608	0.220523312	-1.12	50S ribosomal protein L35	J
Decrease	Ga0436389_006_67191_67577	*rplL*	0.00283184	0.006013339	-0.766	50S ribosomal protein L7/L12	J
Decrease	Ga0436389_024_76096_76371	*rpsS*	0.018780364	0.027931909	-0.858	30S ribosomal protein S19	J
Decrease	Ga0436389_024_19563_20033	*rps7*	0.006236576	0.003951548	-0.576	30S ribosomal protein S7	J

Table B.25 continued from previous page

Change	Protein Accession	Gene	ANOVA (Q)	Tukey (Q)	$\log_2$FC	Description	COG
Decrease	Ga0436389_024_72207_72581	*rps8*	0.004776308	0.014460749	-0.997	30S ribosomal protein S8	J
Decrease	Ga0436389_007_15768_16574	-	0.014632382	0.033218129	-0.539	mechanosensitive ion channel family protein	M
Decrease	Ga0436389_037_4250_5704	*tig*	0.000354976	4.27403E-05	-0.663	trigger factor	D
Decrease	Ga0436389_034_157455_157766	*ccmK2*	0.00283184	0.034793244	-0.651	BMC domain-containing protein	CQ
Decrease	Ga0436389_034_70361_70702	*petJ*	1.40236E-05	1.25232E-05	-1.22	c-type cytochrome	C
Decrease	Ga0436389_032_8650_9396	*rpaB*	0.042145464	0.045103328	-0.507	response regulator transcription factor	K
Decrease	Ga0436389_026_120479_121483	*glpX*	0.000389029	0.002250783	-0.579	class II fructose-bisphosphatase	G
Decrease	Ga0436389_006_270480_271775	*urtA*	6.12887E-06	1.61464E-05	-1.6	urea ABC transporter substrate-binding protein	E
Decrease	Ga0436389_029_19276_20589	*nrtA*	0.000222036	0.000114249	-1.47	ABC transporter substrate-binding protein	P
Decrease	Ga0436389_008_73445_74017	-	0.000736838	0.017794927	-0.669	peroxiredoxin	O

Table B.26: ***Synechococcus* sp. CCAP1479/10 DEPs when grown on glutamate v. NO_3^-**. All proteins were detected in triplicate using TMT quantitative proteomics. Differential expression was determined through FDR-adjusted ANOVA and Tukey's test. $\log_2$FC: $\log_2$ fold change.

Change	**Protein Accession**	**Gene**	**ANOVA (Q)**	**Tukey (Q)**	**$\log_2$FC**	**Description**	**COG**
Increase	Ga0436389_022_100421_100738	*tatA*	3.80939E-06	5.7029E-07	0.829	TatA/E family twin arginine-targeting protein translocase	U
Increase	Ga0436389_006_247274_249397	-	5.76564E-06	5.41039E-08	2.53	transglutaminase family protein	E
Increase	Ga0436389_006_177279_178604	-	1.84716E-05	0.006496581	1.88	ScyD/ScyE family protein	G
Increase	Ga0436389_008_121728_122231	*msrB*	0.000538839	2.06635E-05	1.17	peptide-methionine (R)-S-oxide reductase MsrB	O
Increase	Ga0436389_037_44026_44367	-	0.000810522	0.001650761	0.741	YbaB/EbfC family nucleoid-associated protein	S
Increase	Ga0436389_008_14236_16140	*rpoC1*	0.00154649	0.004377593	0.972	DNA-directed RNA polymerase subunit gamma	K
Increase	Ga0436389_005_2951_3700	-	0.00283184	0.00768571	1.23	DUF305 domain-containing protein	S
Increase	Ga0436389_025_49774_50304	-	0.00283184	0.00012104	0.598	peptide deformylase	J
Increase	Ga0436389_006_109737_112418	*alaS*	0.00283184	0.000154911	0.814	alanine-tRNA ligase	J
Increase	Ga0436389_047_10386_11198	-	0.003464828	0.000688558	1.02	META domain-containing protein	O
Increase	Ga0436389_026_246109_248088	*ftsH1*	0.003551791	0.035360772	2.05	ATP-dependent zinc metalloprotease FtsH	O
Increase	Ga0436389_029_46607_47680	*aroG*	0.003736291	0.008543894	2.07	3-deoxy-7-phosphoheptulonate synthase	E
Increase	Ga0436389_037_6596_7969	*clpX*	0.003842314	0.000361245	2.64	ATP-dependent protease ATP-binding subunit ClpX	O
Increase	Ga0436389_004_46165_47994	-	0.004139985	0.002543954	1.34	ABC transporter	V
Increase	Ga0436389_026_319856_320947	*acsF*	0.004663153	0.001834352	1.09	magnesium-protoporphyrin IX monomethyl ester	H
Increase	Ga0436389_004_4451_6448	*cmpC*	0.004731823	0.006754513	0.898	ABC transporter substrate-binding protein	P
Increase	Ga0436389_047_166282_166944	*cpcF*	0.005401718	0.001536953	0.725	HEAT repeat domain-containing protein	C
Increase	Ga0436389_026_185765_186919	*dapF*	0.006236576	0.005648923	0.701	diaminopimelate epimerase	E
Increase	Ga0436389_037_63067_65535	*parC*	0.006888804	0.002846475	2.57	DNA topoisomerase 4 subunit A	L
Increase	Ga0436389_008_146464_147237	-	0.007926666	0.000659048	1.8	hypothetical protein	S
Increase	Ga0436389_031_32334_33683	*accC*	0.008424705	0.009275411	0.792	acetyl-CoA carboxylase biotin carboxylase subunit	I
Increase	Ga0436389_007_16624_17823	*cefD*	0.010266945	0.000891112	1.01	aminotransferase class V-fold PLP-dependent enzyme	E
Increase	Ga0436389_006_282399_284078	*pyrG*	0.013122001	0.005809378	2.55	CTP synthase	F
Increase	Ga0436389_022_59508_60977	-	0.013822851	0.001287574	1.02	phosphoglucomutase/phosphomannomutase family protein	G
Increase	Ga0436389_024_71253_71618	*rplR*	0.014478567	0.03172043	0.833	50S ribosomal protein L18	J
Increase	Ga0436389_020_38119_40017	*glmS*	0.016250918	0.002038598	1.69	glutamine–fructose-6-phosphate transaminase	M
Increase	Ga0436389_034_695_1744	-	0.01782435	0.008586514	1.94	hypothetical protein	-
Increase	Ga0436389_032_35938_37950	*ndhF1*	0.025134825	0.038286171	2.19	NAD(P)H-quinone oxidoreductase subunit 5	CP
Increase	Ga0436389_038_338_805	-	0.025629089	0.006835415	1.22	bacterioferritin	P

Table B.26 continued from previous page

Change	Protein Accession	Gene	ANOVA (Q)	Tukey (Q)	$\log_2$FC	Description	COG
Increase	Ga0436389_006_251960_252859	-	0.029633032	0.005765917	0.959	transglutaminase family protein	E
Increase	Ga0436389_022_132852_134576	-	0.029633032	0.011890055	0.94	patatin-like phospholipase family protein	S
Increase	Ga0436389_034_106626_106994	*arsC*	0.03051462	0.033095286	1.66	Spx/MgsR family RNA polymerase-binding regulatory protein	P
Increase	Ga0436389_026_265589_266674	-	0.034876657	0.027977269	3.15	fructose-bisphosphate aldolase class I	G
Increase	Ga0436389_006_79714_80949	*argJ*	0.035079555	0.011069398	0.836	arginine biosynthesis bifunctional protein ArgJ	E
Increase	Ga0436389_008_97722_99221	*kaiC*	0.035792517	0.042903558	0.887	circadian clock protein KaiC	T
Increase	Ga0436389_007_21606_22052	-	0.036138173	0.017927233	0.857	PH domain-containing protein	J
Increase	Ga0436389_019_10490_11986	*lysS*	0.038351239	0.008378353	1.3	lysine−tRNA ligase	J
Increase	Ga0436389_008_200112_200912	-	0.03864506	0.018023357	0.738	helix-turn-helix domain-containing protein	S
Increase	Ga0436389_006_62459_63187	-	0.038960755	0.046192514	0.844	PilN domain-containing protein	-
Increase	Ga0436389_034_126808_129276	*ppsA*	0.04160352	0.014780625	2.14	phosphoenolpyruvate synthase	H
Increase	Ga0436389_029_10758_11861	*coaE*	0.041697393	0.009916797	1.95	oxygenase MpaB family protein	H
Decrease	Ga0436389_006_276270_277475	-	7.40069E-07	0.003939052	-0.947	acetamidase/formamidase family protein	C
Decrease	Ga0436389_006_270480_271775	*urtA*	6.12887E-06	3.45813E-05	-1.57	urea ABC transporter substrate-binding protein	E
Decrease	Ga0436389_041_15810_16178	-	7.66039E-06	0.000440315	-0.729	AbrB family transcriptional regulator	K
Decrease	Ga0436389_034_70361_70702	*petJ*	1.40236E-05	1.63881E-07	-1.99	c-type cytochrome	C
Decrease	Ga0436389_008_63390_64907	*atpA*	5.3563E-05	6.14694E-05	-0.542	F0F1 ATP synthase subunit alpha	C
Decrease	Ga0436389_029_89839_90942	*thrC*	0.000137732	0.000126929	-0.87	threonine synthase	E
Decrease	Ga0436389_003_41071_43239	*glnN*	0.000160037	0.001472576	-1.06	glutamine synthetase III	S
Decrease	Ga0436389_003_35798_36316	*cpcB*	0.000167728	5.75419E-05	-0.9	phycocyanin subunit beta	C
Decrease	Ga0436389_029_19276_20589	*nrtA*	0.000222036	0.000152628	-1.78	ABC transporter substrate-binding protein	P
Decrease	Ga0436389_003_33834_34154	*glnB*	0.00025251	0.035656671	-1.45	P-II family nitrogen regulator	K
Decrease	Ga0436389_048_51_539	*cpcA*	0.000289927	0.000300127	-0.681	phycocyanin subunit alpha	C
Decrease	Ga0436389_049_407_1258	*cpeC*	0.000354976	0.000241528	-0.713	phycobilisome rod-core linker polypeptide	H
Decrease	Ga0436389_037_4250_5704	*tig*	0.000354976	2.41352E-05	-0.997	trigger factor	D
Decrease	Ga0436389_026_120479_121483	*glpX*	0.000389029	7.99361E-05	-1.09	class II fructose-bisphosphatase	G
Decrease	Ga0436389_027_49480_51609	*pnp*	0.000538839	0.001397985	-0.635	polyribonucleotide nucleotidyltransferase	J
Decrease	Ga0436389_008_73445_74017	-	0.000736838	0.000966634	-1.06	peroxiredoxin	O
Decrease	Ga0436389_022_258493_259305	*psbO*	0.000875663	4.4791E-05	-1.21	photosystem II manganese-stabilizing polypeptide	S
Decrease	Ga0436389_032_56652_57287	*cysC*	0.001064225	0.001159222	-1.02	adenylyl-sulfate kinase	F
Decrease	Ga0436389_006_47679_48884	*pgk*	0.001280805	0.000541505	-0.759	phosphoglycerate kinase	F

Table B.26 continued from previous page

Change	Protein Accession	Gene	ANOVA (Q)	Tukey (Q)	$\log_2$FC	Description	COG
Decrease	Ga0436389_025_95217_95657	-	0.001541807	0.014330794	-0.686	cAMP phosphodiesterase	-
Decrease	Ga0436389_008_62779_63327	*atpD*	0.001541807	0.001062292	-0.714	F0F1 ATP synthase subunit delta	C
Decrease	Ga0436389_032_82878_83618	-	0.002186741	0.00561671	-0.801	MBL fold metallo-hydrolase	S
Decrease	Ga0436389_046_22209_22643	-	0.00283184	0.038343268	-1.05	SRPBCC family protein	S
Decrease	Ga0436389_044_23310_23921	*clpP3*	0.00283184	0.002778717	-0.658	ATP-dependent Clp protease proteolytic subunit	O
Decrease	Ga0436389_006_67191_67577	*rplL*	0.00283184	0.000233348	-1.2	50S ribosomal protein L7/L12	J
Decrease	Ga0436389_034_157455_157766	*ccmK2*	0.00283184	0.007494421	-0.995	BMC domain-containing protein	CQ
Decrease	Ga0436389_008_204171_205112	*petA*	0.003400378	0.000498451	-1.22	apocytochrome f	C
Decrease	Ga0436389_001_54542_55054	-	0.00392427	0.004032391	-0.88	DNA recombination-mediator protein A	LU
Decrease	Ga0436389_034_159746_162112	*csoS2*	0.00392427	0.005572325	-0.508	carboxysome shell protein	S
Decrease	Ga0436389_004_64037_64612	-	0.004731823	0.009246642	-0.783	hypothetical protein	-
Decrease	Ga0436389_024_72207_72581	*rps8*	0.004776308	0.001245064	-1.12	30S ribosomal protein S8	J
Decrease	Ga0436389_047_148951_149382	*psaD*	0.005942455	0.001426408	-1.11	photosystem I reaction center subunit II	S
Decrease	Ga0436389_004_71443_71565	*psbY*	0.006236576	0.000378942	-1.49	photosystem II protein Y	U
Decrease	Ga0436389_024_19563_20033	*rps7*	0.006236576	0.002009456	-0.771	30S ribosomal protein S7	J
Decrease	Ga0436389_004_47991_48395	*rbpD*	0.006370734	0.002634748	-1.54	RNA-binding protein	S
Decrease	Ga0436389_008_124196_125833	*groEL*	0.007606261	0.006383912	-0.673	chaperonin GroEL	O
Decrease	Ga0436389_008_242405_242728	*trxA*	0.007811318	0.01600614	-0.855	thioredoxin	O
Decrease	Ga0436389_083_3_254	*psbD*	0.010479808	0.004509715	-1.22	photosystem II D2 protein (photosystem q(a) protein)	C
Decrease	Ga0436389_001_50769_51245	*rplS*	0.010905627	0.004273093	-0.962	50S ribosomal protein L19	J
Decrease	Ga0436389_026_294297_294674	-	0.012752538	0.002441384	-1.46	hypothetical protein	-
Decrease	Ga0436389_037_53784_54287	-	0.013296234	0.032646118	-0.99	RNA-binding protein	S
Decrease	Ga0436389_008_100726_101124	*rplU*	0.01336335	0.001309669	-1.27	50S ribosomal protein L21	J
Decrease	Ga0436389_037_40445_41416	*gap2*	0.013756859	0.010856786	-0.726	type I glyceraldehyde-3-phosphate dehydrogenase	C
Decrease	Ga0436389_007_15768_16574	-	0.014632382	0.005673359	-0.65	mechanosensitive ion channel family protein	M
Decrease	Ga0436389_008_54688_55173	*apcA*	0.014911347	0.005594539	-0.607	allophycocyanin subunit alpha apoprotein	C
Decrease	Ga0436389_006_67639_68166	*rplJ*	0.017711586	0.004163031	-0.847	50S ribosomal protein L10	J
Decrease	Ga0436389_050_791_1636	*cpeC*	0.018452092	0.010710383	-0.577	phycobilisome rod-core linker polypeptide	H
Decrease	Ga0436389_034_163661_163975	-	0.018857035	0.039725433	-0.626	carboxysome peptide A	CQ
Decrease	Ga0436389_024_74509_74919	*rplP*	0.019132506	0.041587811	-0.828	50S ribosomal protein L16	J
Decrease	Ga0436389_008_55558_58584	*apcE*	0.020802167	0.006987698	-0.559	phycobilisome rod-core linker polypeptide	G

Table B.26 continued from previous page

Change	Protein Accession	Gene	ANOVA (Q)	Tukey (Q)	$\log_2$FC	Description	COG
Decrease	Ga0436389_007_47651_48175	-	0.023400743	0.008164209	-1.48	hypothetical protein	S
Decrease	Ga0436389_008_209555_210064	*psaF*	0.025134825	0.006924563	-1.04	Photosystem I reaction center subunit III	S
Decrease	Ga0436389_074_3_368	*cpcB*	0.025479647	0.005694793	-1.34	phycocyanin subunit beta	C
Decrease	Ga0436389_026_175590_176126	-	0.034625813	0.037404315	-0.522	hypothetical protein	-
Decrease	Ga0436389_024_15652_15972	*rpsJ*	0.036515336	0.004234932	-0.783	30S ribosomal protein S10	J
Decrease	Ga0436389_050_490_747	-	0.03984082	0.016272115	-0.756	CpcD/allophycocyanin linker domain-containing protein	S
Decrease	Ga0436389_032_8650_9396	*rpaB*	0.042145464	0.019375547	-0.58	response regulator transcription factor	K
Decrease	Ga0436389_015_2539_2748	*psaE*	0.044344903	0.027294837	-1.12	photosystem I reaction center subunit IV	U
Decrease	Ga0436389_037_15970_16170	*rpmI*	0.044973608	0.036157493	-1.02	50S ribosomal protein L35	J
Decrease	Ga0436389_001_97188_98384	*natI*	0.045470018	0.021545154	-0.751	ABC transporter substrate-binding protein	E
Decrease	Ga0436389_034_106258_106620	-	0.045470018	0.007840838	-0.616	(2Fe-2S)-binding protein	C
Decrease	Ga0436389_050_7931_9007	*rps1a*	0.045470018	0.007090303	-0.887	30S ribosomal protein S1	J

Table B.27: ***Synechococcus* sp. CCAP1479/10 DEPs when grown on proline v. NO_3^-**. All proteins were detected in triplicate using TMT quantitative proteomics. Differential expression was determined through FDR-adjusted ANOVA and Tukey's test. log_2FC: log_2 fold change.

Change	Protein Accession	Gene	ANOVA (Q)	Tukey (Q)	log_2FC	Description	COG
Increase	Ga0436389_006_247274_249397	-	5.76564E-06	0.026745147	1.07	transglutaminase family protein	E
Increase	Ga0436389_001_107869_108948	*sbpA*	1.84716E-05	0.003049822	1.75	sulfate ABC transporter substrate-binding protein	P
Increase	Ga0436389_006_211053_212597	-	4.78812E-05	0.003779327	0.71	metallophosphoesterase	S
Increase	Ga0436389_024_37846_43668	-	5.55925E-05	1.76922E-06	3.69	esterase-like activity of phytase family protein	S
Increase	Ga0436389_022_77174_78718	*purH*	0.000116372	1.95612E-06	1.1	Bifunctional purine biosynthesis protein PurH	F
Increase	Ga0436389_022_81418_82803	*amt1*	0.000137732	5.77191E-06	2.55	Ammonium Transporter Family	P
Increase	Ga0436389_004_7569_8897	*cmpA*	0.000137732	0.033515075	1.4	ABC transporter substrate-binding protein	P
Increase	Ga0436389_008_121728_122231	-	0.000538839	0.011801475	0.623	peptide-methionine (R)-S-oxide reductase MsrB	O
Increase	Ga0436389_020_41281_42729	*manC*	0.000558886	3.32985E-05	2.23	mannose-1-phosphate guanylyltransferase	M
Increase	Ga0436389_022_230442_230735	*gatC*	0.000795241	6.30996E-05	0.954	Asp-tRNA(Asn)/Glu-tRNA(Gln) amidotransferase subunit GatC	J
Increase	Ga0436389_037_44026_44367	-	0.000810522	0.014216916	0.568	YbaB/EbfC family nucleoid-associated protein	S
Increase	Ga0436389_034_95986_96882	*argB*	0.000985048	0.000127192	1.21	acetylglutamate kinase	F
Increase	Ga0436389_006_232805_235027	*katG*	0.000985048	3.36018E-05	2.32	catalase/peroxidase HPI	P
Increase	Ga0436389_070_2_241	*glgP*	0.001054454	0.000278935	1.31	glycogen phosphorylase	F
Increase	Ga0436389_026_116581_117309	-	0.001077545	0.000580282	1.44	LPS export ABC transporter ATP-binding protein	S
Increase	Ga0436389_024_71651_72190	*rpl6*	0.001184822	0.000564531	1.12	50S ribosomal protein L6	J
Increase	Ga0436389_008_14236_16140	*rpoC1*	0.00154649	0.001753655	1.05	DNA-directed RNA polymerase subunit gamma	K
Increase	Ga0436389_005_2951_3700	-	0.00283184	0.018980752	1.12	DUF305 domain-containing protein	S
Increase	Ga0436389_006_109737_112418	*alaS*	0.00283184	0.003768827	0.573	alanine-tRNA ligase	J
Increase	Ga0436389_024_76407_77270	*rpl2*	0.00283184	0.008673997	0.504	50S ribosomal protein L2	J
Increase	Ga0436389_032_18824_19840	-	0.00283184	0.001405127	0.532	aromatic ring-hydroxylating dioxygenase subunit alpha	P
Increase	Ga0436389_022_227772_229340	-	0.002919058	0.002724542	0.649	alpha/beta hydrolase	S
Increase	Ga0436389_029_46607_47680	*aroG*	0.003736291	0.003091735	2.23	3-deoxy-7-phosphoheptulonate synthase	E
Increase	Ga0436389_004_46165_47994	-	0.004139985	0.003733239	1.28	ABC transporter	V
Increase	Ga0436389_001_77097_77483	-	0.004820086	0.001642777	1.23	hypothetical protein	-
Increase	Ga0436389_047_166282_166944	*cpcF*	0.005401718	0.007396693	0.58	HEAT repeat domain-containing protein	C
Increase	Ga0436389_026_185765_186919	*dapF*	0.006236576	0.004595668	0.705	diaminopimelate epimerase	E
Increase	Ga0436389_031_32334_33683	*accC*	0.008424705	0.003900343	0.879	acetyl-CoA carboxylase biotin carboxylase subunit	I
Increase	Ga0436389_001_48625_50067	*gltX*	0.010905627	0.002605844	1.27	glutamate–tRNA ligase	J

Table B.27 continued from previous page

Change	Protein Accession	Gene	ANOVA (Q)	Tukey (Q)	$\log_2$FC	Description	COG
Increase	Ga0436389_007_63822_66329	-	0.011410112	0.001536022	0.85	bile acid beta-glucosidase	G
Increase	Ga0436389_026_378378_378947	-	0.013792556	0.005332027	0.543	thylakoid membrane photosystem I accumulation factor	CO
Increase	Ga0436389_022_59508_60977	-	0.013822851	0.029290584	0.723	phosphoglucomutase/phosphomannomutase family protein	G
Increase	Ga0436389_024_65535_65885	*rplQ*	0.014911347	0.010106018	0.842	50S ribosomal protein L17	J
Increase	Ga0436389_044_3813_5120	*murA*	0.015360078	0.00580071	0.792	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	M
Increase	Ga0436389_024_74025_74276	*rpsQ*	0.02169711	0.036316304	1.04	30S ribosomal protein S17	J
Increase	Ga0436389_026_11964_14078	*prlC*	0.023400743	0.00464923	1.47	M3 family metallopeptidase	E
Increase	Ga0436389_006_66146_66901	-	0.025134825	0.005677507	1.41	DUF3747 domain-containing protein	M
Increase	Ga0436389_006_252844_253824	-	0.030052607	0.004864543	1.55	alpha-E domain-containing protein	S
Increase	Ga0436389_034_106626_106994	*arsC*	0.03051462	0.020733135	1.64	Spx/MgsR family RNA polymerase-binding regulatory protein	P
Increase	Ga0436389_047_37485_38417	-	0.034253521	0.015596646	0.524	DUF262 domain-containing protein	-
Increase	Ga0436389_034_115884_117023	*tal*	0.035079555	0.006857672	0.995	transaldolase	H
Increase	Ga0436389_006_79714_80949	*argJ*	0.035079555	0.049477709	0.715	arginine biosynthesis bifunctional protein ArgJ	E
Increase	Ga0436389_024_5314_6540	*aspB*	0.035079555	0.006978579	1.09	LL-diaminopimelate aminotransferase	H
Increase	Ga0436389_030_38087_39844	-	0.035258266	0.009980267	1.89	ABC-F family ATP-binding cassette domain-containing protein	S
Increase	Ga0436389_008_212849_213967	*nfrC*	0.03561983	0.015738964	0.968	UDP-N-acetylglucosamine 2-epimerase (non-hydrolyzing)	G
Increase	Ga0436389_008_97722_99221	*kaiC*	0.035792517	0.022659777	0.954	circadian clock protein KaiC	T
Increase	Ga0436389_008_200112_200912	-	0.03864506	0.048588874	0.675	helix-turn-helix domain-containing protein	S
Increase	Ga0436389_017_1370_2116	-	0.042972213	0.014348328	1.45	hypothetical protein	V
Decrease	Ga0436389_006_276270_277475	-	7.40069E-07	0.00429163	-0.953	acetamidase/formamidase family protein	C
Decrease	Ga0436389_022_100421_100738	*tatA*	3.80939E-06	0.044253763	-0.776	TatA/E family twin arginine-targeting protein translocase	U
Decrease	Ga0436389_006_270480_271775	*urtA*	6.12887E-06	5.84162E-05	-1.45	urea ABC transporter substrate-binding protein	E
Decrease	Ga0436389_041_15810_16178	-	7.66039E-06	1.93406E-05	-1.15	AbrB family transcriptional regulator	K
Decrease	Ga0436389_034_70361_70702	*petJ*	1.40236E-05	2.09566E-07	-1.95	c-type cytochrome	C
Decrease	Ga0436389_008_125887_126198	*groS*	4.05456E-05	0.00859336	-1.45	co-chaperone GroES	O
Decrease	Ga0436389_008_63390_64907	*atpA*	5.3563E-05	4.11049E-07	-1.02	F0F1 ATP synthase subunit alpha	C
Decrease	Ga0436389_029_89839_90942	*thrC*	0.000137732	0.000676773	-0.708	threonine synthase	E
Decrease	Ga0436389_003_35798_36316	*cpcB*	0.000167728	2.81389E-06	-1.38	phycocyanin subunit beta	C
Decrease	Ga0436389_029_19276_20589	*nrtA*	0.000222036	0.002392248	-1.07	ABC transporter substrate-binding protein	P
Decrease	Ga0436389_048_51_539	*cpcA*	0.000289927	4.6332E-05	-0.9	phycocyanin subunit alpha	C
Decrease	Ga0436389_049_407_1258	*cpeC*	0.000354976	0.000265197	-0.737	phycobilisome rod-core linker polypeptide	H

Table B.27 continued from previous page

Change	Protein Accession	Gene	ANOVA (Q)	Tukey (Q)	$\log_2$FC	Description	COG
Decrease	Ga0436389_037_4250_5704	*tig*	0.000354976	3.4218E-05	-0.952	trigger factor	D
Decrease	Ga0436389_026_120479_121483	*glpX*	0.000389029	7.15964E-05	-1.11	class II fructose-bisphosphatase	G
Decrease	Ga0436389_027_49480_51609	*pnp*	0.000538839	0.000392143	-0.775	polyribonucleotide nucleotidyltransferase	J
Decrease	Ga0436389_008_73445_74017	-	0.000736838	0.001532026	-0.994	peroxiredoxin	O
Decrease	Ga0436389_026_339934_341007	*cbbA*	0.000874633	2.62343E-05	-1.19	fructose-bisphosphate aldolase class II	G
Decrease	Ga0436389_022_258493_259305	*psbO*	0.000875663	4.95613E-05	-1.19	photosystem II manganese-stabilizing polypeptide	S
Decrease	Ga0436389_032_56652_57287	*cysC*	0.001064225	0.000666797	-1.11	adenylyl-sulfate kinase	F
Decrease	Ga0436389_006_47679_48884	*pgk*	0.001280805	0.000241438	-0.865	phosphoglycerate kinase	F
Decrease	Ga0436389_008_62779_63327	*atpD*	0.001541807	4.08888E-05	-1.14	F0F1 ATP synthase subunit delta	C
Decrease	Ga0436389_032_82878_83618	-	0.002186741	0.003579773	-0.877	MBL fold metallo-hydrolase	S
Decrease	Ga0436389_044_23310_23921	*clpP3*	0.00283184	0.001897364	-0.687	ATP-dependent Clp protease proteolytic subunit	O
Decrease	Ga0436389_034_157818_159230	*cbbL*	0.00283184	0.000114735	-0.633	form I ribulose bisphosphate carboxylase large subunit	H
Decrease	Ga0436389_006_67191_67577	*rplL*	0.00283184	0.000274578	-1.17	50S ribosomal protein L7/L12	J
Decrease	Ga0436389_034_157455_157766	*ccmK2*	0.00283184	0.006853891	-1	BMC domain-containing protein	CQ
Decrease	Ga0436389_008_204171_205112	*petA*	0.003400378	0.000729122	-1.18	apocytochrome f	C
Decrease	Ga0436389_001_54542_55054	-	0.00392427	0.006521987	-0.837	DNA recombination-mediator protein A	LU
Decrease	Ga0436389_034_159746_162112	*csoS2*	0.00392427	0.003485818	-0.556	carboxysome shell protein	S
Decrease	Ga0436389_004_64037_64612	-	0.004731823	0.001838391	-0.969	hypothetical protein	-
Decrease	Ga0436389_024_72207_72581	*rps8*	0.004776308	0.000385438	-1.32	30S ribosomal protein S8	J
Decrease	Ga0436389_024_17375_19450	*fusA*	0.005942455	0.000231215	-0.789	elongation factor G	J
Decrease	Ga0436389_047_148951_149382	*psaD*	0.005942455	0.014201049	-0.75	photosystem I reaction center subunit II	S
Decrease	Ga0436389_004_71443_71565	*psbY*	0.006236576	0.0064162	-0.918	photosystem II protein Y	U
Decrease	Ga0436389_024_19563_20033	*rps7*	0.006236576	0.000623528	-0.932	30S ribosomal protein S7	J
Decrease	Ga0436389_004_47991_48395	*rbpD*	0.006370734	0.001332342	-1.76	RNA-binding protein	S
Decrease	Ga0436389_008_124196_125833	*groEL*	0.007606261	0.001949682	-0.821	chaperonin GroEL	O
Decrease	Ga0436389_008_242405_242728	*trxA*	0.007811318	0.002671064	-1.24	thioredoxin	O
Decrease	Ga0436389_008_54156_54644	*apcB*	0.008543143	0.001286088	-0.792	allophycocyanin subunit beta	C
Decrease	Ga0436389_083_3_254	*psbD*	0.010479808	0.001628164	-1.5	photosystem II D2 protein (photosystem q(a) protein)	C
Decrease	Ga0436389_001_50769_51245	*rplS*	0.010905627	0.002551911	-1.01	50S ribosomal protein L19	J
Decrease	Ga0436389_026_294297_294674	-	0.012752538	0.004486449	-1.33	hypothetical protein	-
Decrease	Ga0436389_008_100726_101124	*rplU*	0.01336335	0.007806143	-0.947	50S ribosomal protein L21	J

Table B.27 continued from previous page

Change	Protein Accession	Gene	ANOVA (Q)	Tukey (Q)	$\log_2$FC	Description	COG
Decrease	Ga0436389_037_40445_41416	*gap2*	0.013756859	0.002463087	-0.943	type I glyceraldehyde-3-phosphate dehydrogenase	C
Decrease	Ga0436389_007_15768_16574	-	0.014632382	0.001292398	-0.81	mechanosensitive ion channel family protein	M
Decrease	Ga0436389_008_54688_55173	*apcA*	0.014911347	0.004213386	-0.648	allophycocyanin subunit alpha apoprotein	C
Decrease	Ga0436389_006_67639_68166	*rplJ*	0.017711586	0.002897437	-0.921	50S ribosomal protein L10	J
Decrease	Ga0436389_050_791_1636	*cpeC*	0.018452092	0.001670544	-0.768	phycobilisome rod-core linker polypeptide	H
Decrease	Ga0436389_024_76096_76371	*rpsS*	0.018780364	0.00132825	-1.41	30S ribosomal protein S19	J
Decrease	Ga0436389_024_74509_74919	*rplP*	0.019132506	0.001362578	-1.61	50S ribosomal protein L16	J
Decrease	Ga0436389_007_151376_151486	*psbD*	0.02020216	0.002581947	-1.22	Photosystem II DII subunit, Q(A) protein	C
Decrease	Ga0436389_008_55558_58584	*apcE*	0.020802167	0.010486053	-0.524	phycobilisome rod-core linker polypeptide	G
Decrease	Ga0436389_007_47651_48175	-	0.023400743	0.003895386	-1.74	hypothetical protein	S
Decrease	Ga0436389_008_209555_210064	*psaF*	0.025134825	0.005201013	-1.1	Photosystem I reaction center subunit III	S
Decrease	Ga0436389_074_3_368	*cpcB*	0.025479647	0.012027355	-1.17	phycocyanin subunit beta	C
Decrease	Ga0436389_006_182501_182869	-	0.033343677	0.008273393	-0.91	AbrB family transcriptional regulator	K
Decrease	Ga0436389_026_175590_176126	-	0.034625813	0.00342846	-0.814	hypothetical protein	-
Decrease	Ga0436389_050_490_747	-	0.03984082	0.008586892	-0.832	CpcD/allophycocyanin linker domain-containing protein	S
Decrease	Ga0436389_050_9873_11432	*psbB*	0.03984082	0.008668261	-1.19	photosystem II chlorophyll-binding protein CP47	P
Decrease	Ga0436389_032_8650_9396	*rpaB*	0.042145464	0.012876361	-0.628	response regulator transcription factor	K
Decrease	Ga0436389_022_259490_260659	*sat*	0.04376215	0.018439442	-0.878	sulfate adenylyltransferase	P
Decrease	Ga0436389_035_30354_30695	*rbpD*	0.044973608	0.01171385	-1.2	RNA-binding protein	S
Decrease	Ga0436389_037_15970_16170	*rpmI*	0.044973608	0.007285694	-1.43	50S ribosomal protein L35	J
Decrease	Ga0436389_034_159332_159673	*rbcS*	0.045470018	0.044474729	-1	ribulose bisphosphate carboxylase small subunit	C

Table B.28: ***Synechococcus* sp. CCAP1479/10 DEPs when grown on arginine v. nitrogen-starvation**. All proteins were detected in triplicate using TMT quantitative proteomics. Differential expression was determined through FDR-adjusted ANOVA and Tukey's test. $\log_2$FC: $\log_2$ fold change.

Change	Protein Accession	Gene	ANOVA (Q)	Tukey (Q)	$\log_2$FC	Description	COG
Increase	Ga0436389_003_35798_36316	*cpcB*	0.000167728	0.028342312	0.624	phycocyanin subunit beta	C
Increase	Ga0436389_048_51_539	*cpcA*	0.000289927	0.000541164	0.947	phycocyanin subunit alpha	C
Increase	Ga0436389_049_407_1258	*cpeC*	0.000354976	0.001824095	0.94	phycobilisome rod-core linker polypeptide	H
Increase	Ga0436389_037_44026_44367	-	0.000810522	0.001961604	1.07	YbaB/EbfC family nucleoid-associated protein	S
Increase	Ga0436389_008_14236_16140	*rpoC1*	0.00154649	0.02560096	1.07	DNA-directed RNA polymerase subunit gamma	K
Increase	Ga0436389_037_29285_29752	-	0.001733473	0.008864124	0.798	hypothetical protein	-
Increase	Ga0436389_032_18824_19840	-	0.00283184	0.043466982	0.536	aromatic ring-hydroxylating dioxygenase subunit alpha	P
Increase	Ga0436389_026_246109_248088	*ftsH1*	0.003551791	0.010046564	2.63	ATP-dependent zinc metalloprotease FtsH	O
Increase	Ga0436389_029_46607_47680	*aroG*	0.003736291	0.012612551	2.1	3-deoxy-7-phosphoheptulonate synthase	E
Increase	Ga0436389_026_319856_320947	*acsF*	0.004663153	0.036431818	1.1	magnesium-protoporphyrin IX monomethyl ester	H
Increase	Ga0436389_050_18519_19343	*minD*	0.005577792	0.000614939	2.44	septum site-determining protein MinD	D
Increase	Ga0436389_051_24078_24575	-	0.013073998	0.043863469	1.06	DUF411 domain-containing protein	S
Increase	Ga0436389_006_282399_284078	*pyrG*	0.013122001	0.032342226	2.42	CTP synthase	F
Increase	Ga0436389_035_33266_34033	*cpcG1*	0.014632382	0.020522663	0.778	phycobilisome rod-core linker polypeptide	G
Increase	Ga0436389_032_35938_37950	*ndhF1*	0.025134825	0.039115122	2.46	NAD(P)H-quinone oxidoreductase subunit 5	CP
Decrease	Ga0436389_020_45810_46403	-	1.68921E-08	9.11931E-11	-2.16	hypothetical protein	-
Decrease	Ga0436389_026_21467_21841	-	2.09994E-07	2.62395E-09	-2.39	VOC family protein	S
Decrease	Ga0436389_046_28833_30602	-	7.09692E-07	1.76624E-08	-1.54	carbohydrate porin	S
Decrease	Ga0436389_006_276270_277475	-	7.40069E-07	6.02163E-09	-2.49	acetamidase/formamidase family protein	C
Decrease	Ga0436389_007_49464_52028	*astA*	2.40887E-06	2.45986E-08	-1.65	arylsulfatase	P
Decrease	Ga0436389_022_100421_100738	*tatA*	3.80939E-06	0.017248081	-0.673	TatA/E family twin arginine-targeting protein translocase	U
Decrease	Ga0436389_006_270480_271775	*urtA*	6.12887E-06	3.63309E-07	-2.02	urea ABC transporter substrate-binding protein	E
Decrease	Ga0436389_041_15810_16178	-	7.66039E-06	5.11795E-07	-1.12	AbrB family transcriptional regulator	K
Decrease	Ga0436389_006_177279_178604	-	1.84716E-05	5.03026E-06	-1.39	ScyD/ScyE family protein	G
Decrease	Ga0436389_001_107869_108948	*sbpA*	1.84716E-05	6.47216E-07	-2.19	sulfate ABC transporter substrate-binding protein	P
Decrease	Ga0436389_022_205356_205733	-	4.05456E-05	1.92856E-06	-1.44	hypothetical protein	-
Decrease	Ga0436389_008_125887_126198	*groS*	4.05456E-05	3.09308E-06	-1.68	co-chaperone GroES	O

Table B.28 continued from previous page

Change	Protein Accession	Gene	ANOVA (Q)	Tukey (Q)	$\log_2$FC	Description	COG
Decrease	Ga0436389_006_211053_212597	-	4.78812E-05	1.46428E-06	-1.64	metallophosphoesterase	S
Decrease	Ga0436389_029_89839_90942	*thrC*	0.000137732	0.000119564	-0.557	threonine synthase	E
Decrease	Ga0436389_004_7569_8897	*cmpA*	0.000137732	5.91913E-06	-2.19	ABC transporter substrate-binding protein	P
Decrease	Ga0436389_003_41071_43239	*glnN*	0.000160037	7.6783E-06	-1.65	glutamine synthetase III	S
Decrease	Ga0436389_037_28018_29223	*dhsS*	0.000171561	3.10187E-05	-0.826	alanine–glyoxylate aminotransferase family protein	E
Decrease	Ga0436389_029_19276_20589	*nrtA*	0.000222036	7.56775E-05	-1.78	ABC transporter substrate-binding protein	P
Decrease	Ga0436389_003_33834_34154	*glnB*	0.00025251	1.21227E-05	-2.15	P-II family nitrogen regulator	K
Decrease	Ga0436389_008_108347_108646	*petF*	0.000289927	3.73479E-05	-1.87	2Fe-2S iron-sulfur cluster-binding protein	C
Decrease	Ga0436389_018_1_504	-	0.000354976	0.000242993	-0.965	iron uptake porin	S
Decrease	Ga0436389_023_3_1748	-	0.000538839	5.02217E-05	-2.2	hypothetical protein	Q
Decrease	Ga0436389_006_51533_52327	-	0.000585037	7.22232E-05	-1.03	hypothetical protein	-
Decrease	Ga0436389_008_73445_74017	-	0.000736838	0.001668579	-0.67	peroxiredoxin	O
Decrease	Ga0436389_026_39316_39780	-	0.000840212	0.001484014	-0.582	peroxiredoxin	O
Decrease	Ga0436389_026_242104_242433	-	0.000980459	0.000243304	-1.17	nucleoside triphosphate pyrophosphohydrolase family protein	S
Decrease	Ga0436389_070_2_241	*glgP*	0.001054454	0.008163385	-1.3	glycogen phosphorylase	F
Decrease	Ga0436389_032_56652_57287	*cysC*	0.001064225	0.004906991	-0.628	adenylyl-sulfate kinase	F
Decrease	Ga0436389_046_30897_31103	-	0.001077545	0.000531519	-1.21	porin	M
Decrease	Ga0436389_008_242971_244134	*guaB*	0.001077545	0.00051856	-0.593	GuaB3 family IMP dehydrogenase-related protein	F
Decrease	Ga0436389_006_214742_215242	*moaB*	0.001136206	0.000123204	-0.842	molybdenum cofactor biosynthesis protein B	H
Decrease	Ga0436389_006_249428_251854	*glgP*	0.001251338	0.000207192	-1.29	glycogen/starch/alpha-glucan phosphorylase	G
Decrease	Ga0436389_006_284265_284750	*dps*	0.001293886	0.000447868	-1.45	DNA starvation/stationary phase protection protein	P
Decrease	Ga0436389_001_17070_17834	-	0.001754392	0.000621217	-1.51	hypothetical protein	-
Decrease	Ga0436389_032_82878_83618	-	0.002186741	0.005998005	-0.555	MBL fold metallo-hydrolase	S
Decrease	Ga0436389_026_45716_46756	*trpD*	0.002498765	0.000229239	-1.13	anthranilate phosphoribosyltransferase	F
Decrease	Ga0436389_046_22209_22643	-	0.00283184	0.001633924	-0.911	SRPBCC family protein	S
Decrease	Ga0436389_034_157455_157766	*ccmK2*	0.00283184	0.00465414	-0.765	BMC domain-containing protein	CQ
Decrease	Ga0436389_051_14205_14576	-	0.002862583	0.000501124	-1.53	AbrB family transcriptional regulator	K
Decrease	Ga0436389_026_406021_406554	-	0.00291824	0.000539602	-1.32	ferritin	P
Decrease	Ga0436389_030_26116_26676	-	0.003508896	0.001072079	-0.691	alpha/beta hydrolase	S
Decrease	Ga0436389_032_293_556	-	0.003697608	0.000488428	-1.64	hypothetical protein	S
Decrease	Ga0436389_001_54542_55054	-	0.00392427	0.007618067	-0.648	DNA recombination-mediator protein A	LU

Table B.28 continued from previous page

Change	Protein Accession	Gene	ANOVA (Q)	Tukey (Q)	$\log_2$FC	Description	COG
Decrease	Ga0436389_026_170827_171744	*IlvE*	0.004514614	0.001172131	-1.02	branched-chain amino acid transaminase	E
Decrease	Ga0436389_004_4451_6448	*cmpC*	0.004731823	0.00454305	-0.861	ABC transporter substrate-binding protein	P
Decrease	Ga0436389_024_2079_4040	-	0.005942455	0.002666748	-2.03	PFAM Hemolysin-type calcium-binding repeat (2 copies)	Q
Decrease	Ga0436389_047_148951_149382	*psaD*	0.005942455	0.044724014	-0.538	photosystem I reaction center subunit II	S
Decrease	Ga0436389_008_5382_6095	*rpiA*	0.006236576	0.023555848	-0.537	ribose-5-phosphate isomerase RpiA	G
Decrease	Ga0436389_019_5093_5488	*ssb*	0.006886424	0.000914987	-1.28	single-stranded DNA-binding protein	L
Decrease	Ga0436389_006_274310_275056	*urtD*	0.007640158	0.000566822	-1.52	urea ABC transporter ATP-binding protein UrtD	S
Decrease	Ga0436389_022_86912_88222	*glyA*	0.008332703	0.002578551	-0.53	serine hydroxymethyltransferase	E
Decrease	Ga0436389_022_177048_177530	-	0.008889556	0.001804057	-2.41	MSMEG_0572 family nitrogen starvation response protein	S
Decrease	Ga0436389_003_15567_16925	*gor*	0.010479808	0.00321012	-1.05	glutathione-disulfide reductase	C
Decrease	Ga0436389_037_62088_63002	-	0.011410112	0.00511637	-0.812	Tetratricopeptide repeat	S
Decrease	Ga0436389_008_179502_179948	-	0.014632382	0.002520895	-0.782	AbrB family transcriptional regulator	K
Decrease	Ga0436389_003_83842_84417	-	0.015030853	0.014161283	-0.872	ferritin	P
Decrease	Ga0436389_026_300596_301927	*prc*	0.015030853	0.003100296	-0.741	S41 family peptidase	M
Decrease	Ga0436389_035_47426_54061	-	0.02169711	0.01233494	-1.22	tandem-95 repeat protein	MQU
Decrease	Ga0436389_051_45516_45962	*cynS*	0.023400743	0.01276616	-1.39	cyanase	H
Decrease	Ga0436389_006_26936_27742	-	0.026703127	0.006752664	-1.07	glucose 1-dehydrogenase	IQ
Decrease	Ga0436389_050_5813_7072	*metK*	0.027686861	0.024116677	-0.608	methionine adenosyltransferase	H
Decrease	Ga0436389_022_50756_51397	*hisB*	0.03470805	0.02363697	-1.22	imidazoleglycerol-phosphate dehydratase HisB	E
Decrease	Ga0436389_006_48929_49336	-	0.035792517	0.007184131	-1.72	hypothetical protein	-
Decrease	Ga0436389_031_41189_42022	-	0.039326564	0.01825758	-0.755	DUF3747 domain-containing protein	C
Decrease	Ga0436389_022_220013_220420	-	0.03984082	0.026330504	-0.789	DUF1348 family protein	S
Decrease	Ga0436389_007_41358_41957	*sodB*	0.044531423	0.020847158	-1.25	superoxide dismutase	C
Decrease	Ga0436389_035_25523_25975	-	0.045177535	0.020944682	-1.68	hypothetical protein	-
Decrease	Ga0436389_026_411109_412116	-	0.048914251	0.040871833	-1.32	phosphate ABC transporter substrate-binding protein PstS	P

Table B.29: ***Synechococcus* sp. CCAP1479/10 DEPs when grown on asparagine v. nitrogen-starvation**. All proteins were detected in triplicate using TMT quantitative proteomics. Differential expression was determined through FDR-adjusted ANOVA and Tukey's test. $\log_2$FC: $\log_2$ fold change.

Change	Protein Accession	Gene	ANOVA (Q)	Tukey (Q)	$\log_2$FC	Description	COG
Increase	Ga0436389_048_51_539	*cpcA*	0.000289927	0.016682163	0.608	phycocyanin subunit alpha	C
Increase	Ga0436389_049_407_1258	*cpeC*	0.000354976	0.001706047	0.82	phycobilisome rod-core linker polypeptide	H
Increase	Ga0436389_050_11625_12095	-	0.000354976	3.51003E-05	1.6	2Fe-2S iron-sulfur cluster binding domain	C
Increase	Ga0436389_020_41281_42729	manC	0.000558886	0.009053535	1.31	mannose-1-phosphate guanylyltransferase	M
Increase	Ga0436389_037_44026_44367	-	0.000810522	0.00015214	1.15	YbaB/EbfC family nucleoid-associated protein	S
Increase	Ga0436389_026_116581_117309	-	0.001077545	0.005343926	1.43	LPS export ABC transporter ATP-binding protein	S
Increase	Ga0436389_024_71651_72190	*rpl6*	0.001184822	0.017398613	1.16	50S ribosomal protein L6	J
Increase	Ga0436389_008_14236_16140	*rpoC1*	0.00154649	0.001151209	1.27	DNA-directed RNA polymerase subunit gamma	K
Increase	Ga0436389_022_227772_229340	-	0.002919058	0.028354048	0.675	alpha/beta hydrolase	S
Increase	Ga0436389_047_10386_11198	-	0.003464828	0.0166929	0.649	META domain-containing protein	O
Increase	Ga0436389_026_246109_248088	*ftsH1*	0.003551791	0.000906866	2.91	ATP-dependent zinc metalloprotease FtsH	O
Increase	Ga0436389_029_46607_47680	*aroG*	0.003736291	0.008233688	2.11	3-deoxy-7-phosphoheptulonate synthase	E
Increase	Ga0436389_004_46165_47994	-	0.004139985	0.030025378	1.17	ABC transporter	V
Increase	Ga0436389_026_319856_320947	*acsF*	0.004663153	0.008344652	1.12	magnesium-protoporphyrin IX monomethyl ester	H
Increase	Ga0436389_037_63067_65535	*parC*	0.006888804	0.010608785	2.38	DNA topoisomerase 4 subunit A	L
Increase	Ga0436389_024_71253_71618	*rplR*	0.014478567	0.004645252	1.28	50S ribosomal protein L18	J
Increase	Ga0436389_024_65535_65885	*rplQ*	0.014911347	0.036652147	0.952	50S ribosomal protein L17	J
Increase	Ga0436389_025_71305_72870	-	0.02042856	0.01375888	1.34	DUF1957 domain-containing protein	G
Increase	Ga0436389_024_73649_74014	*rplN*	0.025134825	0.00428666	0.959	50S ribosomal protein L14	J
Increase	Ga0436389_026_416356_417222	*dapB*	0.025479647	0.022121465	0.591	4-hydroxy-tetrahydrodipicolinate reductase	E
Decrease	Ga0436389_020_45810_46403	-	1.68921E-08	5.18939E-11	-2.57	hypothetical protein	-
Decrease	Ga0436389_026_21467_21841	-	2.09994E-07	2.5871E-09	-2.55	VOC family protein	S
Decrease	Ga0436389_046_28833_30602	-	7.09692E-07	1.72695E-08	-1.66	carbohydrate porin	S
Decrease	Ga0436389_006_276270_277475	-	7.40069E-07	1.30522E-08	-2.16	acetamidase/formamidase family protein	C
Decrease	Ga0436389_007_49464_52028	*astA*	2.40887E-06	6.36505E-08	-1.51	arylsulfatase	P
Decrease	Ga0436389_006_270480_271775	*urtA*	6.12887E-06	6.12634E-07	-1.92	urea ABC transporter substrate-binding protein	E
Decrease	Ga0436389_041_15810_16178	-	7.66039E-06	1.89872E-06	-1.03	AbrB family transcriptional regulator	K

Table B.29 continued from previous page

Change	Protein Accession	Gene	ANOVA (Q)	Tukey (Q)	$\log_2$FC	Description	COG
Decrease	Ga0436389_006_177279_178604	-	1.84716E-05	2.05762E-06	-1.81	ScyD/ScyE family protein	G
Decrease	Ga0436389_001_107869_108948	*sbpA*	1.84716E-05	9.13634E-06	-1.37	sulfate ABC transporter substrate-binding protein	P
Decrease	Ga0436389_022_205356_205733	-	4.05456E-05	1.40019E-06	-1.73	hypothetical protein	-
Decrease	Ga0436389_008_125887_126198	*groS*	4.05456E-05	7.18022E-06	-1.53	co-chaperone GroES	O
Decrease	Ga0436389_006_211053_212597	-	4.78812E-05	4.03642E-05	-1.15	metallophosphoesterase	S
Decrease	Ga0436389_029_89839_90942	*thrC*	0.000137732	0.00119129	-0.512	threonine synthase	E
Decrease	Ga0436389_004_7569_8897	*cmpA*	0.000137732	2.18458E-05	-1.74	ABC transporter substrate-binding protein	P
Decrease	Ga0436389_003_41071_43239	*glnN*	0.000160037	5.16333E-05	-1.25	glutamine synthetase III	S
Decrease	Ga0436389_037_28018_29223	*dhsS*	0.000171561	2.2591E-05	-1	alanine–glyoxylate aminotransferase family protein	E
Decrease	Ga0436389_029_19276_20589	*nrtA*	0.000222036	0.000248318	-1.46	ABC transporter substrate-binding protein	P
Decrease	Ga0436389_003_33834_34154	*glnB*	0.00025251	6.54448E-05	-1.62	P-II family nitrogen regulator	K
Decrease	Ga0436389_008_108347_108646	*petF*	0.000289927	2.51617E-05	-2.16	2Fe-2S iron-sulfur cluster-binding protein	C
Decrease	Ga0436389_018_1_504	-	0.000354976	0.001171099	-0.894	iron uptake porin	S
Decrease	Ga0436389_023_3_1748	-	0.000538839	0.000183398	-1.73	hypothetical protein	Q
Decrease	Ga0436389_027_49480_51609	*pnp*	0.000538839	0.000618095	-0.61	polyribonucleotide nucleotidyltransferase	J
Decrease	Ga0436389_006_51533_52327	-	0.000585037	6.53157E-05	-1.22	hypothetical protein	-
Decrease	Ga0436389_008_73445_74017	-	0.000736838	0.00135847	-0.813	peroxiredoxin	O
Decrease	Ga0436389_026_39316_39780	-	0.000840212	0.003223896	-0.671	peroxiredoxin	O
Decrease	Ga0436389_026_242104_242433	-	0.000980459	0.000161258	-1.48	nucleoside triphosphate pyrophosphohydrolase family protein	S
Decrease	Ga0436389_032_56652_57287	*cysC*	0.001064225	0.006513164	-0.717	adenylyl-sulfate kinase	F
Decrease	Ga0436389_046_30897_31103	-	0.001077545	0.000828787	-1.28	porin	M
Decrease	Ga0436389_026_100307_100966	*tsf*	0.001136206	0.000648566	-0.53	translation elongation factor Ts	J
Decrease	Ga0436389_006_214742_215242	*moaB*	0.001136206	0.001700231	-0.737	molybdenum cofactor biosynthesis protein B	H
Decrease	Ga0436389_006_249428_251854	*glgP*	0.001251338	0.000144799	-1.55	glycogen/starch/alpha-glucan phosphorylase	G
Decrease	Ga0436389_006_284265_284750	*dps*	0.001293886	0.000370587	-1.6	DNA starvation/stationary phase protection protein	P
Decrease	Ga0436389_025_95217_95657	-	0.001541807	0.000313304	-0.872	cAMP phosphodiesterase	-
Decrease	Ga0436389_001_17070_17834	-	0.001754392	0.000737271	-1.58	hypothetical protein	-
Decrease	Ga0436389_032_82878_83618	-	0.002186741	0.002244621	-0.797	MBL fold metallo-hydrolase	S
Decrease	Ga0436389_026_45716_46756	*trpD*	0.002498765	0.001383718	-0.988	anthranilate phosphoribosyltransferase	F
Decrease	Ga0436389_035_32118_33074	-	0.00283184	0.002816945	-0.512	orange carotenoid-binding protein	S
Decrease	Ga0436389_046_22209_22643	-	0.00283184	0.001116856	-1.14	SRPBCC family protein	S

Table B.29 continued from previous page

Change	Protein Accession	Gene	ANOVA (Q)	Tukey (Q)	$\log_2$FC	Description	COG
Decrease	Ga0436389_034_157455_157766	*ccmK2*	0.00283184	0.006432968	-0.834	BMC domain-containing protein	CQ
Decrease	Ga0436389_051_14205_14576	-	0.002862583	0.002373472	-1.24	AbrB family transcriptional regulator	K
Decrease	Ga0436389_026_406021_406554	-	0.00291824	0.002929313	-1.06	ferritin	P
Decrease	Ga0436389_030_26116_26676	-	0.003508896	0.015762607	-0.515	alpha/beta hydrolase	S
Decrease	Ga0436389_032_293_556	-	0.003697608	0.001034146	-1.49	hypothetical protein	S
Decrease	Ga0436389_001_54542_55054	-	0.00392427	0.017829309	-0.666	DNA recombination-mediator protein A	LU
Decrease	Ga0436389_026_170827_171744	*ilvE*	0.004514614	0.007754559	-0.82	branched-chain amino acid transaminase	E
Decrease	Ga0436389_004_64037_64612	-	0.004731823	0.025176039	-0.599	hypothetical protein	-
Decrease	Ga0436389_004_4451_6448	*cmpC*	0.004731823	0.00979445	-0.91	ABC transporter substrate-binding protein	P
Decrease	Ga0436389_024_2079_4040	-	0.005942455	0.002613501	-2.17	PFAM Hemolysin-type calcium-binding repeat (2 copies)	Q
Decrease	Ga0436389_008_5382_6095	*rpiA*	0.006236576	0.043404769	-0.583	ribose-5-phosphate isomerase RpiA	G
Decrease	Ga0436389_019_5093_5488	*ssb*	0.006886424	0.001619382	-1.3	single-stranded DNA-binding protein	L
Decrease	Ga0436389_006_274310_275056	*urtD*	0.007640158	0.007563035	-1.02	urea ABC transporter ATP-binding protein UrtD	S
Decrease	Ga0436389_022_177048_177530	-	0.008889556	0.007532927	-1.71	MSMEG_0572 family nitrogen starvation response protein	S
Decrease	Ga0436389_016_16760_17248	*ribH*	0.008889556	0.015253183	-0.659	6,7-dimethyl-8-ribityllumazine synthase	H
Decrease	Ga0436389_003_15567_16925	*gor*	0.010479808	0.012474474	-0.916	glutathione-disulfide reductase	C
Decrease	Ga0436389_037_62088_63002	-	0.011410112	0.008208693	-0.853	Tetratricopeptide repeat	S
Decrease	Ga0436389_022_30940_31545	*ndhJ*	0.013879759	0.012998196	-0.692	NAD(P)H-quinone oxidoreductase subunit J	C
Decrease	Ga0436389_038_24_332	*ccmK2*	0.014478567	0.035448709	-0.775	BMC domain-containing protein	CQ
Decrease	Ga0436389_008_179502_179948	-	0.014632382	0.007008639	-0.795	AbrB family transcriptional regulator	K
Decrease	Ga0436389_003_83842_84417	-	0.015030853	0.007367506	-1.25	ferritin	P
Decrease	Ga0436389_032_62224_63021	*ycf29*	0.018854187	0.039176186	-0.56	response regulator	K
Decrease	Ga0436389_035_47426_54061	-	0.02169711	0.031235032	-1.07	tandem-95 repeat protein	MQU
Decrease	Ga0436389_051_45516_45962	*cynS*	0.023400743	0.016895525	-1.42	cyanase	H
Decrease	Ga0436389_006_26936_27742	-	0.026703127	0.021994463	-0.922	glucose 1-dehydrogenase	IQ
Decrease	Ga0436389_050_5813_7072	*metK*	0.027686861	0.027737385	-0.708	methionine adenosyltransferase	H
Decrease	Ga0436389_022_50756_51397	*hisB*	0.03470805	0.047820973	-1.07	imidazoleglycerol-phosphate dehydratase HisB	E
Decrease	Ga0436389_031_41189_42022	-	0.039326564	0.009971533	-1.01	DUF3747 domain-containing protein	C
Decrease	Ga0436389_022_220013_220420	-	0.03984082	0.020904271	-0.971	DUF1348 family protein	S
Decrease	Ga0436389_006_179061_179315	-	0.040298695	0.040828235	-0.784	hypothetical protein	S
Decrease	Ga0436389_007_41358_41957	*sodB*	0.044531423	0.028588045	-1.24	superoxide dismutase	C

Table B.29 continued from previous page

Change	Protein Accession	Gene	ANOVA (Q)	Tukey (Q)	$\log_2$FC	Description	COG
Decrease	Ga0436389_035_25523_25975	-	0.045177535	0.044450193	-1.43	hypothetical protein	-

Table B.30: ***Synechococcus* sp. CCAP1479/10 DEPs when grown on glutamate v. nitrogen-starvation**. All proteins were detected in triplicate using TMT quantitative proteomics. Differential expression was determined through FDR-adjusted ANOVA and Tukey's test. $\log_2$FC: $\log_2$ fold change.

Change	Protein Accession	Gene	ANOVA (Q)	Tukey (Q)	$\log_2$FC	Description	COG
Increase	Ga0436389_022_100421_100738	*tatA*	3.80939E-06	1.01478E-06	0.903	TatA/E family twin arginine-targeting protein translocase	U
Increase	Ga0436389_006_247274_249397	-	5.76564E-06	1.86971E-07	2.01	transglutaminase family protein	E
Increase	Ga0436389_008_121728_122231	-	0.000538839	5.88608E-05	1.11	peptide-methionine (R)-S-oxide reductase MsrB	O
Increase	Ga0436389_037_44026_44367	-	0.000810522	0.000177128	1.16	YbaB/EbfC family nucleoid-associated protein	S
Increase	Ga0436389_008_14236_16140	*rpoC1*	0.00154649	0.001551116	1.23	DNA-directed RNA polymerase subunit gamma	K
Increase	Ga0436389_006_109737_112418	*alaS*	0.00283184	0.004441576	0.556	alanine-tRNA ligase	J
Increase	Ga0436389_047_10386_11198	-	0.003464828	0.011039405	0.692	META domain-containing protein	O
Increase	Ga0436389_026_246109_248088	*ftsH1*	0.003551791	0.027423079	2.32	ATP-dependent zinc metalloprotease FtsH	O
Increase	Ga0436389_029_46607_47680	*aroG*	0.003736291	0.01013734	2.05	3-deoxy-7-phosphoheptulonate synthase	E
Increase	Ga0436389_037_6596_7969	*clpX*	0.003842314	0.000500577	2.49	ATP-dependent protease ATP-binding subunit ClpX	O
Increase	Ga0436389_004_46165_47994	-	0.004139985	0.002327569	1.56	ABC transporter	V
Increase	Ga0436389_026_319856_320947	*acsF*	0.004663153	0.000741032	1.46	magnesium-protoporphyrin IX monomethyl ester	H
Increase	Ga0436389_047_166282_166944	*cpcF*	0.005401718	0.003835127	0.718	HEAT repeat domain-containing protein	C
Increase	Ga0436389_026_185765_186919	*dapF*	0.006236576	0.008472745	0.763	diaminopimelate epimerase	E
Increase	Ga0436389_037_63067_65535	*parC*	0.006888804	0.003295076	2.55	DNA topoisomerase 4 subunit A	L
Increase	Ga0436389_008_146464_147237	-	0.007926666	0.002293719	1.48	hypothetical protein	-
Increase	Ga0436389_031_32334_33683	*accC*	0.008424705	0.012344969	0.832	acetyl-CoA carboxylase biotin carboxylase subunit	I
Increase	Ga0436389_007_16624_17823	*cefD*	0.010266945	0.002695564	0.958	aminotransferase class V-fold PLP-dependent enzyme	E
Increase	Ga0436389_006_282399_284078	*pyrG*	0.013122001	0.006803286	2.52	CTP synthase	F
Increase	Ga0436389_024_71253_71618	*rplR*	0.014478567	0.008730748	1.21	50S ribosomal protein L18	J
Increase	Ga0436389_035_33266_34033	*cpcG1*	0.014632382	0.001919229	0.848	phycobilisome rod-core linker polypeptide	G
Increase	Ga0436389_020_38119_40017	*glmS*	0.016250918	0.006371524	1.4	glutamine–fructose-6-phosphate transaminase	M
Increase	Ga0436389_034_695_1744	-	0.01782435	0.008944411	2.1	hypothetical protein	-
Increase	Ga0436389_024_32741_35002	*psaA*	0.019454238	0.040647416	0.793	photosystem I core protein PsaA	C
Increase	Ga0436389_032_35938_37950	*ndhF1*	0.025134825	0.037878198	2.3	NAD(P)H-quinone oxidoreductase subunit 5	CP
Increase	Ga0436389_038_338_805	-	0.025629089	0.011498117	1.23	bacterioferritin	P
Increase	Ga0436389_026_236904_238823	*dxs*	0.029037399	0.01216682	1.12	1-deoxy-D-xylulose-5-phosphate synthase	H

Table B.30 continued from previous page

Change	Protein Accession	Gene	ANOVA (Q)	Tukey (Q)	$\log_2$FC	Description	COG
Increase	Ga0436389_022_132852_134576	-	0.029633032	0.007344048	1.12	patatin-like phospholipase family protein	S
Increase	Ga0436389_034_106626_106994	*arsC*	0.03051462	0.042656249	1.66	Spx/MgsR family RNA polymerase-binding regulatory protein	P
Increase	Ga0436389_026_113845_115653	*typA*	0.031800823	0.003639287	1.55	translational GTPase TypA	T
Increase	Ga0436389_026_265589_266674	-	0.034876657	0.034888593	2.86	fructose-bisphosphate aldolase class I	G
Increase	Ga0436389_007_21606_22052	-	0.036138173	0.012575848	1.02	PH domain-containing protein	J
Increase	Ga0436389_019_10490_11986	*lysS*	0.038351239	0.038252035	0.976	lysine–tRNA ligase	J
Increase	Ga0436389_008_200112_200912	-	0.03864506	0.037398875	0.747	helix-turn-helix domain-containing protein	S
Increase	Ga0436389_006_62459_63187	-	0.038960755	0.016273491	1.21	PilN domain-containing protein	-
Increase	Ga0436389_034_126808_129276	*ppsA*	0.04160352	0.016336703	2.15	phosphoenolpyruvate synthase	H
Increase	Ga0436389_029_10758_11861	*coaE*	0.041697393	0.011062366	2.04	oxygenase MpaB family protein	H
Increase	Ga0436389_006_69080_69787	*rpl1*	0.04376215	0.00917161	1.45	50S ribosomal protein L1	J
Decrease	Ga0436389_020_45810_46403	-	1.68921E-08	2.66516E-11	-2.73	hypothetical protein	-
Decrease	Ga0436389_026_21467_21841	-	2.09994E-07	1.52346E-09	-2.87	VOC family protein	S
Decrease	Ga0436389_046_28833_30602	-	7.09692E-07	9.90803E-09	-1.86	carbohydrate porin	S
Decrease	Ga0436389_006_276270_277475	-	7.40069E-07	1.76802E-08	-2	acetamidase/formamidase family protein	C
Decrease	Ga0436389_007_49464_52028	*astA*	2.40887E-06	4.08009E-08	-1.62	arylsulfatase	P
Decrease	Ga0436389_006_270480_271775	*urtA*	6.12887E-06	6.45784E-07	-1.89	urea ABC transporter substrate-binding protein	E
Decrease	Ga0436389_041_15810_16178	-	7.66039E-06	1.43174E-06	-1.06	AbrB family transcriptional regulator	K
Decrease	Ga0436389_034_70361_70702	*petJ*	1.40236E-05	0.030979437	-0.807	c-type cytochrome	C
Decrease	Ga0436389_006_177279_178604	-	1.84716E-05	1.19183E-05	-1.32	ScyD/ScyE family protein	G
Decrease	Ga0436389_001_107869_108948	*sbpA*	1.84716E-05	2.32313E-06	-1.73	sulfate ABC transporter substrate-binding protein	P
Decrease	Ga0436389_022_205356_205733	-	4.05456E-05	2.75824E-06	-1.43	hypothetical protein	-
Decrease	Ga0436389_008_125887_126198	*groS*	4.05456E-05	2.26992E-06	-1.9	co-chaperone GroES	O
Decrease	Ga0436389_006_211053_212597	-	4.78812E-05	2.73309E-05	-1.07	metallophosphoesterase	S
Decrease	Ga0436389_029_89839_90942	*thrC*	0.000137732	5.11183E-06	-1.01	threonine synthase	E
Decrease	Ga0436389_004_7569_8897	*cmpA*	0.000137732	4.13028E-05	-1.55	ABC transporter substrate-binding protein	P
Decrease	Ga0436389_003_41071_43239	*glnN*	0.000160037	3.31651E-05	-1.33	glutamine synthetase III	S
Decrease	Ga0436389_037_28018_29223	*dhsS*	0.000171561	1.72522E-05	-1.03	alanine–glyoxylate aminotransferase family protein	E
Decrease	Ga0436389_029_19276_20589	*nrtA*	0.000222036	0.000100333	-1.78	ABC transporter substrate-binding protein	P
Decrease	Ga0436389_003_33834_34154	*glnB*	0.00025251	8.98537E-06	-2.46	P-II family nitrogen regulator	K
Decrease	Ga0436389_008_108347_108646	*petF*	0.000289927	3.44894E-05	-1.99	2Fe-2S iron-sulfur cluster-binding protein	C

Table B.30 continued from previous page

Change	Protein Accession	Gene	ANOVA (Q)	Tukey (Q)	log$_2$FC	Description	COG
Decrease	Ga0436389_018_1_504	-	0.000354976	0.000325994	-1.08	iron uptake porin	S
Decrease	Ga0436389_026_120479_121483	*glpX*	0.000389029	0.000286097	-0.945	class II fructose-bisphosphatase	G
Decrease	Ga0436389_023_3_1748	-	0.000538839	0.000199193	-1.69	hypothetical protein	Q
Decrease	Ga0436389_027_49480_51609	*pnp*	0.000538839	0.000284659	-0.685	polyribonucleotide nucleotidyltransferase	J
Decrease	Ga0436389_006_51533_52327	-	0.000585037	0.000303558	-0.907	hypothetical protein	-
Decrease	Ga0436389_008_73445_74017	-	0.000736838	0.000120916	-1.2	peroxiredoxin	O
Decrease	Ga0436389_026_39316_39780	-	0.000840212	2.33194E-05	-1.21	peroxiredoxin	O
Decrease	Ga0436389_026_242104_242433	-	0.000980459	0.000188382	-1.26	nucleoside triphosphate pyrophosphohydrolase family protein	S
Decrease	Ga0436389_070_2_241	*glgP*	0.001054454	0.027251673	-1.05	glycogen phosphorylase	F
Decrease	Ga0436389_032_56652_57287	*cysC*	0.001064225	0.000532587	-1.02	adenylyl-sulfate kinase	F
Decrease	Ga0436389_046_30897_31103	-	0.001077545	3.92354E-05	-2.27	porin	M
Decrease	Ga0436389_008_242971_244134	*guaB*	0.001077545	0.001203567	-0.63	GuaB3 family IMP dehydrogenase-related protein	F
Decrease	Ga0436389_026_100307_100966	*tsf*	0.001136206	0.000134523	-0.651	translation elongation factor Ts	J
Decrease	Ga0436389_006_214742_215242	*moaB*	0.001136206	6.73833E-05	-0.98	molybdenum cofactor biosynthesis protein B	H
Decrease	Ga0436389_006_249428_251854	*glgP*	0.001251338	0.000439823	-1.21	glycogen/starch/alpha-glucan phosphorylase	G
Decrease	Ga0436389_006_47679_48884	*pgk*	0.001280805	0.005976909	-0.56	phosphoglycerate kinase	F
Decrease	Ga0436389_006_284265_284750	*dps*	0.001293886	0.000176102	-1.81	DNA starvation/stationary phase protection protein	P
Decrease	Ga0436389_025_95217_95657	-	0.001541807	0.000128667	-0.992	cAMP phosphodiesterase	-
Decrease	Ga0436389_001_17070_17834	-	0.001754392	0.001107085	-1.54	hypothetical protein	-
Decrease	Ga0436389_032_82878_83618	-	0.002186741	0.001486505	-0.839	MBL fold metallo-hydrolase	S
Decrease	Ga0436389_026_45716_46756	*trpD*	0.002498765	0.001599496	-0.88	anthranilate phosphoribosyltransferase	F
Decrease	Ga0436389_035_32118_33074	-	0.00283184	0.00013196	-0.802	orange carotenoid-binding protein	S
Decrease	Ga0436389_046_22209_22643	-	0.00283184	0.000134839	-1.5	SRPBCC family protein	S
Decrease	Ga0436389_044_23310_23921	*clpP3*	0.00283184	0.001733319	-0.608	ATP-dependent Clp protease proteolytic subunit	O
Decrease	Ga0436389_034_157455_157766	*ccmK2*	0.00283184	0.00109275	-1.18	BMC domain-containing protein	CQ
Decrease	Ga0436389_051_14205_14576	-	0.002862583	0.000356663	-1.74	AbrB family transcriptional regulator	K
Decrease	Ga0436389_026_406021_406554	-	0.00291824	0.000338773	-1.62	ferritin	P
Decrease	Ga0436389_008_204171_205112	*petA*	0.003400378	0.020368423	-0.834	apocytochrome f	C
Decrease	Ga0436389_030_26116_26676	-	0.003508896	0.000428565	-0.94	alpha/beta hydrolase	S
Decrease	Ga0436389_032_293_556	-	0.003697608	0.000700093	-1.59	hypothetical protein	S
Decrease	Ga0436389_001_54542_55054	-	0.00392427	0.004872012	-0.825	DNA recombination-mediator protein A	LU

Table B.30 continued from previous page

Change	Protein Accession	Gene	ANOVA (Q)	Tukey (Q)	$\log_2$FC	Description	COG
Decrease	Ga0436389_026_170827_171744	*IlvE*	0.004514614	0.001880459	-1.06	branched-chain amino acid transaminase	E
Decrease	Ga0436389_004_64037_64612	-	0.004731823	0.01742691	-0.673	hypothetical protein	-
Decrease	Ga0436389_024_2079_4040	-	0.005942455	0.001284021	-2.54	PFAM Hemolysin-type calcium-binding repeat (2 copies)	Q
Decrease	Ga0436389_047_148951_149382	*psaD*	0.005942455	0.012789989	-0.883	photosystem I reaction center subunit II	S
Decrease	Ga0436389_019_5093_5488	*ssb*	0.006886424	0.004198889	-1.06	single-stranded DNA-binding protein	L
Decrease	Ga0436389_006_274310_275056	*urtD*	0.007640158	0.002842339	-1.23	urea ABC transporter ATP-binding protein UrtD	S
Decrease	Ga0436389_008_242405_242728	*trxA*	0.007811318	0.018845939	-0.764	thioredoxin	O
Decrease	Ga0436389_022_177048_177530	-	0.008889556	0.002415089	-2.24	MSMEG_0572 family nitrogen starvation response protein	S
Decrease	Ga0436389_016_16760_17248	*ribH*	0.008889556	0.003326182	-0.817	6,7-dimethyl-8-ribityllumazine synthase	H
Decrease	Ga0436389_003_15567_16925	*gor*	0.010479808	0.00772323	-0.968	glutathione-disulfide reductase	C
Decrease	Ga0436389_037_62088_63002	-	0.011410112	0.008276803	-0.864	Tetratricopeptide repeat	S
Decrease	Ga0436389_037_53784_54287	-	0.013296234	0.001756546	-1.19	RNA-binding protein	S
Decrease	Ga0436389_022_30940_31545	*ndhJ*	0.013879759	0.003231941	-0.881	NAD(P)H-quinone oxidoreductase subunit J	C
Decrease	Ga0436389_038_24_332	*ccmK2*	0.014478567	0.002912548	-0.897	BMC domain-containing protein	CQ
Decrease	Ga0436389_003_83842_84417	-	0.015030853	0.038899483	-0.763	ferritin	P
Decrease	Ga0436389_026_300596_301927	*prc*	0.015030853	0.015166775	-0.626	S41 family peptidase	M
Decrease	Ga0436389_025_91047_92216	*petH*	0.017384523	0.002331119	-0.889	FAD-binding oxidoreductase	C
Decrease	Ga0436389_032_62224_63021	*ycf29*	0.018854187	0.006539669	-0.807	response regulator	K
Decrease	Ga0436389_034_163661_163975	-	0.018857035	0.006013424	-0.675	carboxysome peptide A	CQ
Decrease	Ga0436389_051_45516_45962	*cynS*	0.023400743	0.008330229	-1.81	cyanase	H
Decrease	Ga0436389_008_163251_164024	*pspA*	0.025479647	0.011967624	-0.601	PspA/IM30 family protein	KT
Decrease	Ga0436389_006_26936_27742	-	0.026703127	0.006227453	-1.22	glucose 1-dehydrogenase	IQ
Decrease	Ga0436389_050_5813_7072	*metK*	0.027686861	0.012057081	-0.835	methionine adenosyltransferase	H
Decrease	Ga0436389_096_1_843	*glgP*	0.032296347	0.036588008	-0.563	glycogen/starch/alpha-glucan family phosphorylase	F
Decrease	Ga0436389_022_50756_51397	*hisB*	0.03470805	0.034172954	-1.19	imidazoleglycerol-phosphate dehydratase HisB	E
Decrease	Ga0436389_031_41189_42022	-	0.039326564	0.049006232	-0.691	DUF3747 domain-containing protein	C
Decrease	Ga0436389_030_27016_27483	*bcp*	0.039326564	0.025929699	-0.623	thioredoxin-dependent thiol peroxidase	O
Decrease	Ga0436389_022_220013_220420	-	0.03984082	0.017955957	-0.948	DUF1348 family protein	S
Decrease	Ga0436389_006_179061_179315	-	0.040298695	0.011478476	-0.991	hypothetical protein	S
Decrease	Ga0436389_015_2539_2748	*psaE*	0.044344903	0.038023827	-1.06	photosystem I reaction center subunit IV	U
Decrease	Ga0436389_007_41358_41957	*sodB*	0.044531423	0.013365893	-1.57	superoxide dismutase	C

Table B.30 continued from previous page

Change	Protein Accession	Gene	ANOVA (Q)	Tukey (Q)	$\log_2$FC	Description	COG
Decrease	Ga0436389_035_25523_25975	-	0.045177535	0.027894537	-1.6	hypothetical protein	-
Decrease	Ga0436389_026_132651_134327	*ilvD*	0.045470018	0.005396703	-1.04	dihydroxy-acid dehydratase	H

Table B.31: ***Synechococcus* sp. CCAP1479/10 DEPs when grown on proline v. nitrogen-starvation**. All proteins were detected in triplicate using TMT quantitative proteomics. Differential expression was determined through FDR-adjusted ANOVA and Tukey's test. $\log_2$FC: $\log_2$ fold change.

Change	Protein Accession	Gene	ANOVA (Q)	Tukey (Q)	$\log_2$FC	Description	COG
Increase	Ga0436389_024_37846_43668	-	5.55925E-05	2.56915E-06	3.35	esterase-like activity of phytase family protein	S
Increase	Ga0436389_022_77174_78718	*purH*	0.000116372	1.35234E-05	0.896	Bifunctional purine biosynthesis protein PurH	F
Increase	Ga0436389_022_81418_82803	*amt1*	0.000137732	9.56876E-06	2.34	Ammonium Transporter Family	P
Increase	Ga0436389_008_121728_122231	-	0.000538839	0.048683818	0.564	peptide-methionine (R)-S-oxide reductase MsrB	O
Increase	Ga0436389_020_41281_42729	*manC*	0.000558886	8.02587E-05	1.91	mannose-1-phosphate guanylyltransferase	M
Increase	Ga0436389_022_230442_230735	*gatC*	0.000795241	7.10191E-05	1.03	Asp-tRNA(Asn)/Glu-tRNA(Gln) amidotransferase subunit GatC	J
Increase	Ga0436389_037_44026_44367	-	0.000810522	0.001221901	0.983	YbaB/EbfC family nucleoid-associated protein	S
Increase	Ga0436389_034_95986_96882	*argB*	0.000985048	0.000493139	1.02	acetylglutamate kinase	F
Increase	Ga0436389_006_232805_235027	*katG*	0.000985048	0.002281928	1.05	catalase/peroxidase HPI	P
Increase	Ga0436389_026_116581_117309	-	0.001077545	0.000135878	1.9	LPS export ABC transporter ATP-binding protein	S
Increase	Ga0436389_024_71651_72190	*rpl6*	0.001184822	5.76075E-05	1.76	50S ribosomal protein L6	J
Increase	Ga0436389_008_14236_16140	*rpoC1*	0.00154649	0.000648226	1.31	DNA-directed RNA polymerase subunit gamma	K
Increase	Ga0436389_024_76407_77270	*rpl2*	0.00283184	0.000100158	1.06	50S ribosomal protein L2	J
Increase	Ga0436389_032_18824_19840	-	0.00283184	0.000428986	0.745	aromatic ring-hydroxylating dioxygenase subunit alpha	P
Increase	Ga0436389_022_227772_229340	-	0.002919058	0.000158371	1.11	alpha/beta hydrolase	S
Increase	Ga0436389_029_46607_47680	*aroG*	0.003736291	0.003648218	2.22	3-deoxy-7-phosphoheptulonate synthase	E
Increase	Ga0436389_004_46165_47994	-	0.004139985	0.003410912	1.5	ABC transporter	V
Increase	Ga0436389_001_77097_77483	-	0.004820086	0.000922161	1.41	hypothetical protein	-
Increase	Ga0436389_047_166282_166944	*cpcF*	0.005401718	0.019361815	0.573	HEAT repeat domain-containing protein	C
Increase	Ga0436389_026_185765_186919	*dapF*	0.006236576	0.00687529	0.767	diaminopimelate epimerase	E
Increase	Ga0436389_031_32334_33683	*accC*	0.008424705	0.005154613	0.919	acetyl-CoA carboxylase biotin carboxylase subunit	I
Increase	Ga0436389_001_48625_50067	*gltX*	0.010905627	0.006330149	1.16	glutamate–tRNA ligase	J
Increase	Ga0436389_007_63822_66329	-	0.011410112	0.006224153	0.739	bile acid beta-glucosidase	G
Increase	Ga0436389_024_65535_65885	*rplQ*	0.014911347	0.001807524	1.27	50S ribosomal protein L17	J
Increase	Ga0436389_044_3813_5120	*murA*	0.015360078	0.00413661	0.953	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	M
Increase	Ga0436389_024_74025_74276	*rpsQ*	0.02169711	0.008373913	1.39	30S ribosomal protein S17	J
Increase	Ga0436389_026_11964_14078	*prlC*	0.023400743	0.016195823	1.15	M3 family metallopeptidase	E

Table B.31 continued from previous page

Change	Protein Accession	Gene	ANOVA (Q)	Tukey (Q)	$\log_2$FC	Description	COG
Increase	Ga0436389_006_66146_66901	-	0.025134825	0.00612965	1.56	DUF3747 domain-containing protein	M
Increase	Ga0436389_029_21478_23514	*nrtC*	0.025629089	0.002986753	0.728	ABC transporter substrate-binding protein	P
Increase	Ga0436389_026_236904_238823	*dxs*	0.029037399	0.02250793	1.09	1-deoxy-D-xylulose-5-phosphate synthase	H
Increase	Ga0436389_006_252844_253824	-	0.030052607	0.02318249	1.17	alpha-E domain-containing protein	S
Increase	Ga0436389_034_106626_106994	*arsC*	0.03051462	0.02674473	1.64	Spx/MgsR family RNA polymerase-binding regulatory protein	P
Increase	Ga0436389_047_37485_38417	-	0.034253521	0.018036001	0.616	DUF262 domain-containing protein	-
Increase	Ga0436389_024_5314_6540	*aspB*	0.035079555	0.013257968	1.02	LL-diaminopimelate aminotransferase	H
Increase	Ga0436389_030_38087_39844	-	0.035258266	0.009335588	2.06	ABC-F family ATP-binding cassette domain-containing protein	S
Increase	Ga0436389_008_212849_213967	*nfrC*	0.03561983	0.034169196	0.893	UDP-N-acetylglucosamine 2-epimerase (non-hydrolyzing)	G
Increase	Ga0436389_017_1370_2116	-	0.042972213	0.012365949	1.65	hypothetical protein	V
Decrease	Ga0436389_020_45810_46403	-	1.68921E-08	3.74114E-11	-2.57	hypothetical protein	-
Decrease	Ga0436389_026_21467_21841	-	2.09994E-07	2.22572E-09	-2.61	VOC family protein	S
Decrease	Ga0436389_046_28833_30602	-	7.09692E-07	1.57903E-08	-1.7	carbohydrate porin	S
Decrease	Ga0436389_006_276270_277475	-	7.40069E-07	1.83136E-08	-2.01	acetamidase/formamidase family protein	C
Decrease	Ga0436389_007_49464_52028	*astA*	2.40887E-06	7.23731E-08	-1.49	arylsulfatase	P
Decrease	Ga0436389_022_100421_100738	*tatA*	3.80939E-06	0.01580808	-0.702	TatA/E family twin arginine-targeting protein translocase	U
Decrease	Ga0436389_006_270480_271775	*urtA*	6.12887E-06	9.55491E-07	-1.77	urea ABC transporter substrate-binding protein	E
Decrease	Ga0436389_041_15810_16178	-	7.66039E-06	1.6494E-07	-1.49	AbrB family transcriptional regulator	K
Decrease	Ga0436389_006_177279_178604	-	1.84716E-05	1.51382E-06	-1.88	ScyD/ScyE family protein	G
Decrease	Ga0436389_001_107869_108948	*sbpA*	1.84716E-05	3.66148E-05	-1.16	sulfate ABC transporter substrate-binding protein	P
Decrease	Ga0436389_022_205356_205733	-	4.05456E-05	1.54113E-06	-1.56	hypothetical protein	-
Decrease	Ga0436389_008_125887_126198	*groS*	4.05456E-05	6.7492E-07	-2.5	co-chaperone GroES	O
Decrease	Ga0436389_006_211053_212597	-	4.78812E-05	0.000166443	-0.844	metallophosphoesterase	S
Decrease	Ga0436389_008_63390_64907	*atpA*	5.3563E-05	1.76773E-05	-0.744	F0F1 ATP synthase subunit alpha	C
Decrease	Ga0436389_029_89839_90942	*thrC*	0.000137732	1.94887E-05	-0.851	threonine synthase	E
Decrease	Ga0436389_004_7569_8897	*cmpA*	0.000137732	0.000310754	-1.13	ABC transporter substrate-binding protein	P
Decrease	Ga0436389_003_41071_43239	*glnN*	0.000160037	0.000718933	-0.854	glutamine synthetase III	S
Decrease	Ga0436389_037_28018_29223	*dhsS*	0.000171561	5.86196E-06	-1.23	alanine–glyoxylate aminotransferase family protein	E
Decrease	Ga0436389_029_19276_20589	*nrtA*	0.000222036	0.001468185	-1.06	ABC transporter substrate-binding protein	P
Decrease	Ga0436389_003_33834_34154	*glnB*	0.00025251	0.001628257	-1.01	P-II family nitrogen regulator	K
Decrease	Ga0436389_008_108347_108646	*petF*	0.000289927	1.35547E-05	-2.64	2Fe-2S iron-sulfur cluster-binding protein	C

Table B.31 continued from previous page

Change	Protein Accession	Gene	ANOVA (Q)	Tukey (Q)	$\log_2$FC	Description	COG
Decrease	Ga0436389_018_1_504	-	0.000354976	1.01304E-05	-1.79	iron uptake porin	S
Decrease	Ga0436389_026_120479_121483	*glpX*	0.000389029	0.000253915	-0.966	class II fructose-bisphosphatase	G
Decrease	Ga0436389_023_3_1748	-	0.000538839	0.000186083	-1.72	hypothetical protein	Q
Decrease	Ga0436389_027_49480_51609	*pnp*	0.000538839	8.9017E-05	-0.825	polyribonucleotide nucleotidyltransferase	J
Decrease	Ga0436389_006_51533_52327	-	0.000585037	6.13057E-05	-1.2	hypothetical protein	-
Decrease	Ga0436389_008_73445_74017	-	0.000736838	0.000181943	-1.14	peroxiredoxin	O
Decrease	Ga0436389_026_39316_39780	-	0.000840212	9.23416E-05	-0.994	peroxiredoxin	O
Decrease	Ga0436389_026_242104_242433	-	0.000980459	5.6586E-05	-1.5	nucleoside triphosphate pyrophosphohydrolase family protein	S
Decrease	Ga0436389_032_56652_57287	*cysC*	0.001064225	0.000313414	-1.11	adenylyl-sulfate kinase	F
Decrease	Ga0436389_046_30897_31103	-	0.001077545	0.00098383	-1.3	porin	M
Decrease	Ga0436389_008_242971_244134	*guaB*	0.001077545	0.000319688	-0.774	GuaB3 family IMP dehydrogenase-related protein	F
Decrease	Ga0436389_026_100307_100966	*tsf*	0.001136206	7.09099E-05	-0.728	translation elongation factor Ts	J
Decrease	Ga0436389_006_214742_215242	*moaB*	0.001136206	0.003709065	-0.575	molybdenum cofactor biosynthesis protein B	H
Decrease	Ga0436389_006_249428_251854	*glgP*	0.001251338	0.000223805	-1.45	glycogen/starch/alpha-glucan phosphorylase	G
Decrease	Ga0436389_006_47679_48884	*pgk*	0.001280805	0.002413436	-0.666	phosphoglycerate kinase	F
Decrease	Ga0436389_006_284265_284750	*dps*	0.001293886	0.00087635	-1.33	DNA starvation/stationary phase protection protein	P
Decrease	Ga0436389_025_95217_95657	-	0.001541807	0.000815365	-0.754	cAMP phosphodiesterase	-
Decrease	Ga0436389_008_62779_63327	*atpD*	0.001541807	0.011436211	-0.671	F0F1 ATP synthase subunit delta	C
Decrease	Ga0436389_001_17070_17834	-	0.001754392	0.000905451	-1.48	hypothetical protein	-
Decrease	Ga0436389_032_82878_83618	-	0.002186741	0.000971317	-0.914	MBL fold metallo-hydrolase	S
Decrease	Ga0436389_026_45716_46756	*trpD*	0.002498765	0.000970754	-0.972	anthranilate phosphoribosyltransferase	F
Decrease	Ga0436389_046_22209_22643	-	0.00283184	0.011203903	-0.783	SRPBCC family protein	S
Decrease	Ga0436389_044_23310_23921	*clpP3*	0.00283184	0.001193193	-0.637	ATP-dependent Clp protease proteolytic subunit	O
Decrease	Ga0436389_034_157455_157766	*ccmK2*	0.00283184	0.001006149	-1.19	BMC domain-containing protein	CQ
Decrease	Ga0436389_051_14205_14576	-	0.002862583	0.003926106	-1.1	AbrB family transcriptional regulator	K
Decrease	Ga0436389_026_406021_406554	-	0.00291824	0.022623041	-0.696	ferritin	P
Decrease	Ga0436389_008_204171_205112	*petA*	0.003400378	0.031372252	-0.792	apocytochrome f	C
Decrease	Ga0436389_032_293_556	-	0.003697608	0.001780278	-1.47	hypothetical protein	S
Decrease	Ga0436389_001_54542_55054	-	0.00392427	0.007903331	-0.782	DNA recombination-mediator protein A	LU
Decrease	Ga0436389_026_170827_171744	*IlvE*	0.004514614	0.012975787	-0.761	branched-chain amino acid transaminase	E
Decrease	Ga0436389_004_64037_64612	-	0.004731823	0.003356635	-0.858	hypothetical protein	-

Table B.31 continued from previous page

Change	Protein Accession	Gene	ANOVA (Q)	Tukey (Q)	$\log_2$FC	Description	COG
Decrease	Ga0436389_024_2079_4040	-	0.005942455	0.00783776	-1.72	PFAM Hemolysin-type calcium-binding repeat (2 copies)	Q
Decrease	Ga0436389_008_5382_6095	*rpiA*	0.006236576	0.000307155	-1.42	ribose-5-phosphate isomerase RpiA	G
Decrease	Ga0436389_004_47991_48395	*rbpD*	0.006370734	0.031945116	-1.36	RNA-binding protein	S
Decrease	Ga0436389_019_5093_5488	*ssb*	0.006886424	0.003746223	-1.08	single-stranded DNA-binding protein	L
Decrease	Ga0436389_006_274310_275056	*urtD*	0.007640158	0.00642342	-0.987	urea ABC transporter ATP-binding protein UrtD	S
Decrease	Ga0436389_008_242405_242728	*trxA*	0.007811318	0.003120264	-1.14	thioredoxin	O
Decrease	Ga0436389_016_16760_17248	*ribH*	0.008889556	0.008460515	-0.7	6,7-dimethyl-8-ribityllumazine synthase	H
Decrease	Ga0436389_003_15567_16925	*gor*	0.010479808	0.011916799	-0.883	glutathione-disulfide reductase	C
Decrease	Ga0436389_037_62088_63002	-	0.011410112	0.011430383	-0.797	Tetratricopeptide repeat	S
Decrease	Ga0436389_037_53784_54287	-	0.013296234	0.006093629	-0.944	RNA-binding protein	S
Decrease	Ga0436389_037_40445_41416	*gap2*	0.013756859	0.01320873	-0.761	type I glyceraldehyde-3-phosphate dehydrogenase	C
Decrease	Ga0436389_022_30940_31545	*ndhJ*	0.013879759	0.001715498	-0.959	NAD(P)H-quinone oxidoreductase subunit J	C
Decrease	Ga0436389_038_24_332	*ccmK2*	0.014478567	0.003018745	-0.872	BMC domain-containing protein	CQ
Decrease	Ga0436389_008_179502_179948	-	0.014632382	0.007463424	-0.739	AbrB family transcriptional regulator	K
Decrease	Ga0436389_003_83842_84417	-	0.015030853	0.002120685	-1.34	ferritin	P
Decrease	Ga0436389_046_16533_17252	*gst1*	0.016562088	0.013329952	-0.638	Glutathione S-transferase	O
Decrease	Ga0436389_025_91047_92216	*petH*	0.017384523	0.011974969	-0.682	FAD-binding oxidoreductase	C
Decrease	Ga0436389_032_62224_63021	*ycf29*	0.018854187	0.001964414	-0.911	response regulator	K
Decrease	Ga0436389_034_163661_163975	-	0.018857035	0.010960992	-0.627	carboxysome peptide A	CQ
Decrease	Ga0436389_051_45516_45962	*cynS*	0.023400743	0.011395878	-1.49	cyanase	H
Decrease	Ga0436389_008_163251_164024	*pspA*	0.025479647	0.024471382	-0.524	PspA/IM30 family protein	KT
Decrease	Ga0436389_022_50756_51397	*hisB*	0.03470805	0.004998768	-1.7	imidazoleglycerol-phosphate dehydratase HisB	E
Decrease	Ga0436389_022_220013_220420	-	0.03984082	0.013512913	-1.02	DUF1348 family protein	S
Decrease	Ga0436389_006_179061_179315	-	0.040298695	0.011514719	-0.98	hypothetical protein	S
Decrease	Ga0436389_007_41358_41957	*sodB*	0.044531423	0.025394135	-1.29	superoxide dismutase	C
Decrease	Ga0436389_034_159332_159673	*rbcS*	0.045470018	0.01276639	-1.13	ribulose bisphosphate carboxylase small subunit	C
Decrease	Ga0436389_026_132651_134327	*ilvD*	0.045470018	0.045041043	-0.688	dihydroxy-acid dehydratase	H

Table B.32: ***Synechococcus* sp. CCAP1479/10 DEPs when grown on arginine v. asparagine**. All proteins were detected in triplicate using TMT quantitative proteomics. Differential expression was determined through FDR-adjusted ANOVA and Tukey's test. $\log_2$FC: $\log_2$ fold change.

Up-regulated in ...	**Protein Accession**	**Gene**	**ANOVA (Q)**	**Tukey (Q)**	**log2FC**	**Description**	**COG**
Arginine	Ga0436389_051_24078_24575	-	0.013073998	0.005449952	1.76	DUF411 domain-containing protein	S
Arginine	Ga0436389_005_2951_3700	-	0.00283184	0.019417618	0.779	DUF305 domain-containing protein	S
Arginine	Ga0436389_050_18519_19343	*minD*	0.005577792	0.043367626	1.05	septum site-determining protein MinD	D
Arginine	Ga0436389_032_18824_19840	-	0.00283184	0.049852075	0.644	aromatic ring-hydroxylating dioxygenase subunit alpha	P
Asparagine	Ga0436389_050_11625_12095	-	0.000354976	4.00494E-05	1.58	2Fe-2S iron-sulfur cluster binding domain	C
Asparagine	Ga0436389_025_71305_72870	-	0.02042856	0.005208707	1.9	DUF1957 domain-containing protein	G
Asparagine	Ga0436389_026_116581_117309	-	0.001077545	0.011693872	1.19	LPS export ABC transporter ATP-binding protein	S

Table B.33: ***Synechococcus* sp. CCAP1479/10 DEPs when grown on arginine v. glutamate**. All proteins were detected in triplicate using TMT quantitative proteomics. Differential expression was determined through FDR-adjusted ANOVA and Tukey's test. $\log_2$FC: $\log_2$ fold change.

Up-regulated in ...	**Protein Accession**	**Gene**	**ANOVA (Q)**	**Tukey (Q)**	**log2FC**	**Description**	**COG**
Arginine	Ga0436389_048_51_539	*cpcA*	0.00029	0.036963	0.508	phycocyanin subunit alpha	C
Arginine	Ga0436389_008_55558_58584	*apcE*	0.020802	0.047787	0.56	phycobilisome linker protein	G
Arginine	Ga0436389_046_16533_17252	*gst1*	0.016562	0.026734	0.584	Glutathione S-transferase	O
Arginine	Ga0436389_003_35798_36316	*cpcB*	0.000168	0.016903	0.639	phycocyanin subunit beta	C
Arginine	Ga0436389_035_32118_33074		0.002832	0.004315	0.728	orange carotenoid-binding protein	S
Arginine	Ga0436389_037_29285_29752		0.001733	0.001474	1.03	hypothetical protein	S
Arginine	Ga0436389_050_18519_19343	*minD*	0.005578	0.049507	1.04	septum site-determining protein MinD	D
Arginine	Ga0436389_034_70361_70702	*petJ*	1.4E-05	0.008778	1.07	cytochrome c6	C
Arginine	Ga0436389_051_24078_24575		0.013074	0.00425	1.91	DUF411 domain-containing protein	S
Glutamate	Ga0436389_034_126808_129276	*ppsA*	0.041604	0.015067	2.21	phosphoenolpyruvate synthase	H
Glutamate	Ga0436389_006_247274_249397		5.77E-06	1.15E-07	2.18	transglutaminase family protein	E
Glutamate	Ga0436389_022_100421_100738	*tatA*	3.81E-06	3.41E-08	1.58	TatA/E family twin arginine-targeting protein translocase	U
Glutamate	Ga0436389_038_338_805		0.025629	0.011104	1.38	bacterioferritin	P
Glutamate	Ga0436389_019_10490_11986	*lysS*	0.038351	0.012011	1.23	lysine–tRNA ligase	J
Glutamate	Ga0436389_008_146464_147237		0.007927	0.011332	1.1	hypothetical protein	S
Glutamate	Ga0436389_037_6596_7969	*clpX*	0.003842	0.010635	1.08	ATP-dependent protease ATP-binding subunit ClpX	O
Glutamate	Ga0436389_020_38119_40017	*glmS*	0.016251	0.02338	0.89	glutamine–fructose-6-phosphate transaminase	M
Glutamate	Ga0436389_008_121728_122231		0.000539	0.000271	0.855	peptide-methionine (R)-S-oxide reductase MsrB	O
Glutamate	Ga0436389_026_113845_115653	*typA*	0.031801	0.027937	0.853	translational GTPase TypA	T
Glutamate	Ga0436389_026_236904_238823	*dxs*	0.029037	0.039454	0.752	1-deoxy-D-xylulose-5-phosphate synthase	H
Glutamate	Ga0436389_025_49774_50304		0.002832	0.001036	0.657	peptide deformylase	J
Glutamate	Ga0436389_026_185765_186919	*dapF*	0.006237	0.025766	0.604	diaminopimelate epimerase	E
Glutamate	Ga0436389_007_16624_17823	*cefD*	0.010267	0.046981	0.578	aminotransferase class V-fold PLP-dependent enzyme	E
Glutamate	Ga0436389_047_166282_166944	*cpcF*	0.005402	0.013789	0.542	HEAT repeat domain-containing protein	C

Table B.34: ***Synechococcus* sp. CCAP1479/10 DEPs when grown on arginine v. proline**. All proteins were detected in triplicate using TMT quantitative proteomics. Differential expression was determined through FDR-adjusted ANOVA and Tukey's test. $\log_2$FC: $\log_2$ fold change.

Up-regulated in ...	**Protein Accession**	**Gene**	**ANOVA (Q)**	**Tukey (Q)**	**log2FC**	**Description**	**COG**
Arginine	Ga0436389_034_157818_159230	cbbL	0.002832	0.017384	0.5	form I ribulose bisphosphate carboxylase large subunit	H
Arginine	Ga0436389_024_17375_19450	*fusA*	0.005942	0.023439	0.608	elongation factor G	J
Arginine	Ga0436389_008_63390_64907	*atpA*	5.36E-05	0.001304	0.665	F0F1 ATP synthase subunit alpha	C
Arginine	Ga0436389_048_51_539	*cpcA*	0.00029	0.003658	0.728	phycocyanin subunit alpha	C
Arginine	Ga0436389_005_2951_3700		0.002832	0.024133	0.736	DUF305 domain-containing protein	S
Arginine	Ga0436389_046_16533_17252	*gst1*	0.016562	0.002325	0.844	Glutathione S-transferase	O
Arginine	Ga0436389_008_54156_54644	*apcB*	0.008543	0.002389	0.846	allophycocyanin subunit beta	C
Arginine	Ga0436389_026_339934_341007	*cbbA*	0.000875	0.001559	0.998	fructose-bisphosphate aldolase class II	G
Arginine	Ga0436389_034_70361_70702	*petJ*	1.4E-05	0.014185	1.03	c-type cytochrome	C
Arginine	Ga0436389_003_35798_36316	*cpcB*	0.000168	0.00028	1.12	phycocyanin subunit beta	C
Arginine	Ga0436389_037_29285_29752		0.001733	0.000218	1.36	hypothetical protein	S
Arginine	Ga0436389_051_24078_24575		0.013074	0.006521	1.64	DUF411 domain-containing protein	S
Arginine	Ga0436389_050_18519_19343	*minD*	0.005578	0.002358	1.87	septum site-determining protein MinD	D
Proline	Ga0436389_024_37846_43668		5.56E-05	3.76E-06	2.97	esterase-like activity of phytase family protein	S
Proline	Ga0436389_022_81418_82803	*amt1*	0.000138	5.57E-06	2.66	Ammonium Transporter Family	P
Proline	Ga0436389_024_74025_74276	*rpsQ*	0.021697	0.002731	1.78	30S ribosomal protein S17	J
Proline	Ga0436389_070_2_241	*glgP*	0.001054	0.000143	1.7	glycogen phosphorylase	F
Proline	Ga0436389_006_232805_235027	*katG*	0.000985	0.000119	1.65	catalase/peroxidase HPI	P
Proline	Ga0436389_026_116581_117309		0.001078	0.000262	1.65	LPS export ABC transporter ATP-binding protein	S
Proline	Ga0436389_030_38087_39844		0.035258	0.027801	1.57	ABC-F family ATP-binding cassette domain-containing protein	S
Proline	Ga0436389_006_252844_253824		0.030053	0.015758	1.38	alpha-E domain-containing protein	S
Proline	Ga0436389_026_11964_14078	*prlC*	0.023401	0.006564	1.37	M3 family metallopeptidase	E
Proline	Ga0436389_034_95986_96882	*argB*	0.000985	9.57E-05	1.32	acetylglutamate kinase	F
Proline	Ga0436389_020_41281_42729	*manC*	0.000559	0.000389	1.32	mannose-1-phosphate guanylyltransferase/mannose-6-phosphate isomerase	M
Proline	Ga0436389_001_77097_77483		0.00482	0.001489	1.26	hypothetical protein	S
Proline	Ga0436389_017_1370_2116		0.042972	0.043917	1.25	hypothetical protein	V
Proline	Ga0436389_006_66146_66901		0.025135	0.020104	1.23	DUF3747 domain-containing protein	M
Proline	Ga0436389_024_71651_72190	*rpl6*	0.001185	0.000269	1.2	50S ribosomal protein L6	J
Proline	Ga0436389_003_33834_34154		0.000253	0.032256	1.14	P-II family nitrogen regulator	K
Proline	Ga0436389_001_48625_50067	*gltX*	0.010906	0.008394	1.05	glutamate–tRNA ligase	J
Proline	Ga0436389_001_107869_108948	*sbpA*	1.85E-05	0.02619	1.04	sulfate ABC transporter substrate-binding protein	P
Proline	Ga0436389_044_3813_5120	*murA*	0.01536	0.003711	0.979	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	M
Proline	Ga0436389_022_230442_230735	*gatC*	0.000795	0.000185	0.89	Asp-tRNA(Asn)/Glu-tRNA(Gln) amidotransferase subunit GatC	J
Proline	Ga0436389_026_378378_378947		0.013793	0.00244	0.88	thylakoid membrane photosystem I accumulation factor	CO
Proline	Ga0436389_006_211053_212597		4.79E-05	0.015684	0.798	metallophosphoesterase	S
Proline	Ga0436389_003_41071_43239	*glnN*	0.00016	0.04027	0.794	glutamine synthetase III	S
Proline	Ga0436389_022_77174_78718	*purH*	0.000116	1.21E-05	0.782	Bifunctional purine biosynthesis protein PurH	F
Proline	Ga0436389_007_63822_66329		0.01141	0.00258	0.78	bile acid beta-glucosidase	G
Proline	Ga0436389_024_5314_6540	*dapL*	0.03508	0.027874	0.768	LL-diaminopimelate aminotransferase	H
Proline	Ga0436389_047_37485_38417		0.034254	0.009046	0.706	DUF262 domain-containing protein	S
Proline	Ga0436389_034_115884_117023	*tal*	0.03508	0.026628	0.646	transaldolase	H
Proline	Ga0436389_022_227772_229340		0.002919	0.001935	0.64	alpha/beta hydrolase	S
Proline	Ga0436389_024_76407_77270	*rpl2*	0.002832	0.000933	0.632	50S ribosomal protein L2	J
Proline	Ga0436389_026_185765_186919	*dapF*	0.006237	0.020831	0.608	diaminopimelate epimerase	E
Proline	Ga0436389_026_300596_301927	*prc*	0.015031	0.023774	0.58	S41 family peptidase	M
Proline	Ga0436389_024_65535_65885	*rplQ*	0.014911	0.038311	0.568	50S ribosomal protein L17	J
Proline	Ga0436389_031_32334_33683	*accC*	0.008425	0.037664	0.534	acetyl-CoA carboxylase biotin carboxylase subunit	I
Proline	Ga0436389_029_21478_23514	*nrtC*	0.025629	0.010601	0.528	ABC transporter substrate-binding protein	P

Table B.35: ***Synechococcus* sp. CCAP1479/10 DEPs when grown on asparagine v. glutamate**. All proteins were detected in triplicate using TMT quantitative proteomics. Differential expression was determined through FDR-adjusted ANOVA and Tukey's test. $\log_2$FC: $\log_2$ fold change.

Up-regulated in ...	**Protein Accession**	**Gene**	**ANOVA (Q)**	**Tukey (Q)**	**log2FC**	**Description**	**COG**
Asparagine	Ga0436389_029_89839_90942	*thrC*	0.000138	0.012352	0.501	threonine synthase	E
Asparagine	Ga0436389_026_39316_39780		0.00084	0.039677	0.544	peroxiredoxin	O
Asparagine	Ga0436389_008_124196_125833	*groEL*	0.007606	0.020643	0.579	chaperonin GroEL	O
Asparagine	Ga0436389_024_73649_74014	*rplN*	0.025135	0.022425	0.694	50S ribosomal protein L14	J
Asparagine	Ga0436389_037_29285_29752		0.001733	0.014508	0.711	hypothetical protein	S
Asparagine	Ga0436389_050_11625_12095		0.000355	0.000307	1.19	2Fe-2S iron-sulfur cluster binding domain	C
Asparagine	Ga0436389_025_71305_72870		0.020429	0.007424	1.73	DUF1957 domain-containing protein	G
Glutamate	Ga0436389_034_126808_129276	*ppsA*	0.041604	0.043684	1.63	phosphoenolpyruvate synthase	H
Glutamate	Ga0436389_006_247274_249397		5.77E-06	9.39E-07	1.61	transglutaminase family protein	E
Glutamate	Ga0436389_038_338_805		0.025629	0.013025	1.55	bacterioferritin	P
Glutamate	Ga0436389_037_6596_7969	*clpX*	0.003842	0.009193	1.35	ATP-dependent protease ATP-binding subunit ClpX	O
Glutamate	Ga0436389_022_100421_100738	*tatA*	3.81E-06	3.27E-07	1.31	TatA/E family twin arginine-targeting protein translocase	U
Glutamate	Ga0436389_008_146464_147237		0.007927	0.017006	1.07	hypothetical protein	S
Glutamate	Ga0436389_008_121728_122231		0.000539	0.000369	0.992	peptide-methionine (R)-S-oxide reductase MsrB	O
Glutamate	Ga0436389_007_16624_17823	*cefD*	0.010267	0.015425	0.935	aminotransferase class V-fold PLP-dependent enzyme	E
Glutamate	Ga0436389_025_49774_50304		0.002832	0.003927	0.727	peptide deformylase	J

Table B.36: ***Synechococcus* sp. CCAP1479/10 DEPs when grown on asparagine v. proline**. All proteins were detected in triplicate using TMT quantitative proteomics. Differential expression was determined through FDR-adjusted ANOVA and Tukey's test. $\log_2$FC: $\log_2$ fold change.

Up-regulated in ...	**Protein Accession**	**Gene**	**ANOVA (Q)**	**Tukey (Q)**	**log2FC**	**Description**	**COG**
Asparagine	Ga0436389_008_63390_64907	*atpA*	5.36E-05	0.00143	0.548	F0F1 ATP synthase subunit alpha	C
Asparagine	Ga0436389_008_54156_54644	*apcB*	0.008543	0.011937	0.629	allophycocyanin subunit beta	C
Asparagine	Ga0436389_008_124196_125833	*groEL*	0.007606	0.006049	0.727	chaperonin GroEL	O
Asparagine	Ga0436389_003_35798_36316	*cpcB*	0.000168	0.007139	0.766	phycocyanin subunit beta	C
Asparagine	Ga0436389_026_339934_341007	*cbbA*	0.000875	0.002011	0.859	fructose-bisphosphate aldolase class II	G
Asparagine	Ga0436389_018_1_504		0.000355	0.035453	0.9	iron uptake porin	S
Asparagine	Ga0436389_037_29285_29752		0.001733	0.001762	1.05	hypothetical protein	S
Asparagine	Ga0436389_024_32741_35002	*psaA*	0.019454	0.009544	1.07	photosystem I core protein PsaA	C
Asparagine	Ga0436389_026_246109_248088	*ftsH1*	0.003552	0.01636	1.47	ATP-dependent zinc metalloprotease FtsH	O
Asparagine	Ga0436389_050_11625_12095		0.000355	3.74E-05	1.63	2Fe-2S iron-sulfur cluster binding domain	C
Proline	Ga0436389_024_37846_43668		5.56E-05	0.000107	1.44	esterase-like activity of phytase family protein	S
Proline	Ga0436389_022_81418_82803	*amt1*	0.000138	0.000154	1.38	Ammonium Transporter Family	P
Proline	Ga0436389_006_66146_66901		0.025135	0.024021	1.37	DUF3747 domain-containing protein	M
Proline	Ga0436389_070_2_241	*glgP*	0.001054	0.000726	1.33	glycogen phosphorylase	F
Proline	Ga0436389_006_232805_235027	*katG*	0.000985	0.002962	0.992	catalase/peroxidase HPI	P
Proline	Ga0436389_026_11964_14078	*prlC*	0.023401	0.035336	0.988	M3 family metallopeptidase	E
Proline	Ga0436389_032_18824_19840		0.002832	0.000484	0.853	aromatic ring-hydroxylating dioxygenase subunit alpha	P
Proline	Ga0436389_001_77097_77483		0.00482	0.01568	0.844	hypothetical protein	S
Proline	Ga0436389_034_95986_96882	*argB*	0.000985	0.008178	0.694	acetylglutamate kinase	F
Proline	Ga0436389_026_185765_186919	*dapF*	0.006237	0.043791	0.694	diaminopimelate epimerase	E
Proline	Ga0436389_024_71651_72190	*rpl6*	0.001185	0.024157	0.593	50S ribosomal protein L6	J
Proline	Ga0436389_007_63822_66329		0.01141	0.03992	0.572	bile acid beta-glucosidase	G
Proline	Ga0436389_022_230442_230735	*gatC*	0.000795	0.013259	0.567	Asp-tRNA(Asn)/Glu-tRNA(Gln) amidotransferase subunit GatC	J
Proline	Ga0436389_022_77174_78718	*purH*	0.000116	0.000599	0.538	Bifunctional purine biosynthesis protein PurH	F
Proline	Ga0436389_024_76407_77270	*rpl2*	0.002832	0.00817	0.528	50S ribosomal protein L2	J

Table B.37: ***Synechococcus* sp. CCAP1479/10 DEPs when grown on glutamate v. proline**. All proteins were detected in triplicate using TMT quantitative proteomics. Differential expression was determined through FDR-adjusted ANOVA and Tukey's test. $\log_2$FC: $\log_2$ fold change.

Up-regulated in ...	Protein Accession	Gene	ANOVA (Q)	Tukey (Q)	log2FC	Description	COG
Glutamate	Ga0436389_006_282399_284078	*pyrG*	0.013122	0.025385	1.62	CTP synthase	F
Glutamate	Ga0436389_022_100421_100738	*tatA*	3.81E-06	3.28E-08	1.6	TatA/E family twin arginine-targeting protein translocase	U
Glutamate	Ga0436389_037_6596_7969	*clpX*	0.003842	0.002867	1.59	ATP-dependent protease ATP-binding subunit ClpX	O
Glutamate	Ga0436389_034_695_1744		0.017824	0.032831	1.59	hypothetical protein	S
Glutamate	Ga0436389_006_247274_249397		5.77E-06	1.43E-06	1.46	transglutaminase family protein	E
Glutamate	Ga0436389_008_146464_147237		0.007927	0.006264	1.25	hypothetical protein	S
Glutamate	Ga0436389_038_338_805		0.025629	0.025713	1.14	bacterioferritin	P
Glutamate	Ga0436389_024_32741_35002	*psaA*	0.019454	0.006129	1.11	photosystem I core protein PsaA	C
Glutamate	Ga0436389_006_62459_63187		0.038961	0.031061	1.02	PilN domain-containing protein	S
Glutamate	Ga0436389_019_10490_11986	*lysS*	0.038351	0.041825	0.999	lysine–tRNA ligase	J
Glutamate	Ga0436389_007_21606_22052		0.036138	0.024068	0.982	PH domain-containing protein	J
Glutamate	Ga0436389_026_339934_341007	*cbbA*	0.000875	0.001345	0.891	fructose-bisphosphate aldolase class II	G
Glutamate	Ga0436389_026_113845_115653	*typA*	0.031801	0.046016	0.864	translational GTPase TypA	T
Glutamate	Ga0436389_007_16624_17823	*cefD*	0.010267	0.018033	0.791	aminotransferase class V-fold PLP-dependent enzyme	E
Glutamate	Ga0436389_025_49774_50304		0.002832	0.001592	0.647	peptide deformylase	J
Glutamate	Ga0436389_035_33266_34033	*cpcG1*	0.014632	0.01843	0.586	phycobilisome rod-core linker polypeptide	G
Glutamate	Ga0436389_008_121728_122231		0.000539	0.008904	0.548	peptide-methionine (R)-S-oxide reductase MsrB	O
Glutamate	Ga0436389_047_10386_11198		0.003465	0.044288	0.543	META domain-containing protein	O
Glutamate	Ga0436389_024_17375_19450	*fusA*	0.005942	0.018191	0.517	elongation factor G	J
Proline	Ga0436389_024_76407_77270	*rpl2*	0.002832	0.004642	0.581	50S ribosomal protein L2	J
Proline	Ga0436389_007_63822_66329		0.01141	0.02323	0.581	bile acid beta-glucosidase	G
Proline	Ga0436389_030_27016_27483	*bcp*	0.039327	0.023181	0.648	thioredoxin-dependent thiol peroxidase	O
Proline	Ga0436389_047_37485_38417		0.034254	0.023596	0.657	DUF262 domain-containing protein	S
Proline	Ga0436389_022_227772_229340		0.002919	0.002247	0.77	alpha/beta hydrolase	S
Proline	Ga0436389_022_77174_78718	*purH*	0.000116	3.19E-05	0.782	Bifunctional purine biosynthesis protein PurH	F
Proline	Ga0436389_022_230442_230735	*gatC*	0.000795	0.000523	0.823	Asp-tRNA(Asn)/Glu-tRNA(Gln) amidotransferase subunit GatC	J
Proline	Ga0436389_001_48625_50067	*gltX*	0.010906	0.0289	0.876	glutamate–tRNA ligase	J
Proline	Ga0436389_024_71651_72190	*rpl6*	0.001185	0.003572	0.905	50S ribosomal protein L6	J
Proline	Ga0436389_026_116581_117309		0.001078	0.001951	1.18	LPS export ABC transporter ATP-binding protein	S
Proline	Ga0436389_026_11964_14078	*prlC*	0.023401	0.017479	1.21	M3 family metallopeptidase	E
Proline	Ga0436389_024_74025_74276	*rpsQ*	0.021697	0.021344	1.24	30S ribosomal protein S17	J
Proline	Ga0436389_034_95986_96882	*argB*	0.000985	0.000144	1.29	acetylglutamate kinase	F
Proline	Ga0436389_020_41281_42729	*manC*	0.000559	0.000753	1.29	mannose-1-phosphate guanylyltransferase/mannose-6-phosphate isomerase	M
Proline	Ga0436389_006_232805_235027	*katG*	0.000985	0.000645	1.34	catalase/peroxidase HPI	P
Proline	Ga0436389_006_66146_66901		0.025135	0.02522	1.35	DUF3747 domain-containing protein	M
Proline	Ga0436389_001_77097_77483		0.00482	0.001068	1.4	hypothetical protein	S
Proline	Ga0436389_003_33834_34154		0.000253	0.020732	1.44	P-II family nitrogen regulator	K
Proline	Ga0436389_070_2_241	*glgP*	0.001054	0.000392	1.45	glycogen phosphorylase	F
Proline	Ga0436389_022_81418_82803	*amt1*	0.000138	7.67E-05	1.65	Ammonium Transporter Family	P
Proline	Ga0436389_024_37846_43668		5.56E-05	9.63E-06	2.37	esterase-like activity of phytase family protein	S

Appendix C

Supplementary Text

C.1 Attempts to Produce an Axenic Culture

C.1.1 Use of Antibiotics

The first technique utilised to produce axenic cultures was antibiotic isolation, based on Hong et al [559]. Imipenem, cycloheximide, and kanamycin have previously been used for cyanobacteria filamentous cyanobacteria purification. Imipenem inhibits peptidoglycan biosynthesis and kanamycin inhibits protein synthesis, impacting only bacteria which actively divide and grow when incubated in the dark. Meanwhile, eukaryotes are inhibited through cycloheximide addition, preventing translational elongation.

Picocyanobacterial culture (1.5 mL) was harvested, centrifuged for 15 mins at 3,000 x g, and streaked onto BG-11 agar containing imipenem (100 μg mL^{-1}) (Generon, UK) and cycloheximide (20 μg mL^{-1}) (Insight Biotechnology, UK). Plates were incubated in darkness to prevent picocyanobacteria growth. After 24 hours, plates were transferred to a 16 h light: 8 h dark cycle at 20 °C and checked daily for growth. Upon visible growth, picocyanobacteria were transferred to fresh BG-11 plates (no antibiotic supplement). The continued presence of contamination was determined through transfer onto LB [560] and R2A [561] agar plates. These plates were incubated in darkness for 3 + days to ascertain axenic status. Non-axenic cultures were transferred to BG-11 agar supplemented with kanamycin (100 μg mL^{-1}) (Merck, Germany) and incubated for 24 hours in darkness. The plates were then transferred to a 16 h light: 8 h dark cycle and purification steps repeated. However, after several cycles, no viable picocyanobacteria remained and other purification techniques were followed.

This technique for culture purification has not previously been successfully carried out with picocyanobacteria. Reductions in antibiotic concentration administered may be necessary for the survival of exposed picocyanobacteria, however this must be balanced with sufficient antibiotic for contaminating-bacteria removal. Further research is required to find an appropriate dosage.

C.1.2 Serial Dilution and Solid-Liquid Alternate Incubation

The next method to generate axenic cultures utilised dilution and mixed-media incubation [562]. Well-grown cultures of *Synechococcus* spp. were serially diluted to 10^{-2}, 10^{-5}, and 10^{-8} dilutions and each dilution streaked onto BG-11 agar plates. The plates were incubated for 15 days under a 16 h light: 8 h dark cycle at 20 °C. Following incubation, streaked picocyanobacteria were transferred to a flask containing 10 mL BG-11 medium. 0.1 mL of picocyanobacteria culture was streaked onto an LB agar plate once the culture had become green to test for bacterial contamination. These plates were incubated for 3 days at the above conditions to detect bacterial growth. If bacterial growth was detected, the purification process was repeated with serial dilutions and picocyanobacteria picking. This was repeated for four cycles, however bacterial growth was consistently present alongside picocyanobacteria.

Bibliography

[1] Cristiana Callieri and John G. Stockner.
Freshwater autotrophic picoplankton: A review.
Journal of Limnology, 61(1):1–14, 2002.

[2] Ho Won Lee, Jae Hoon Noh, Dong Han Choi, Misun Yun, P. S. Bhavya, Jae Joong Kang, Jae Hyung Lee, Kwan Woo Kim, Hyo Keun Jang, and Sang Heon Lee.
Picocyanobacterial contribution to the total primary production in the northwestern pacific ocean.
Water (Switzerland), 13(11):1–15, 2021.

[3] Boglárka Somogyi, Tamás Felföldi, László G. Tóth, Gábor Bernát, and Lajos Vörös.
Photoautotrophic picoplankton – a review on their occurrence, role and diversity in Lake Balaton.
Biologia Futura, 71(4):371–382, 2020.

[4] Joachim Ruber, Juergen Geist, Manuela Hartmann, Andrew Millard, Uta Raeder, Mikhail Zubkov, and Katrin Zwirglmaier.
Spatio-temporal distribution pattern of the picocyanobacterium *Synechococcus* in lakes of different trophic states: a comparison of flow cytometry and sequencing approaches.
Hydrobiologia, 811(1):77–92, 2018.

[5] Hanyan Li, Anwar Alsanea, Michael Barber, and Ramesh Goel.
High-throughput DNA sequencing reveals the dominance of pico- and other filamentous cyanobacteria in an urban freshwater Lake.

Science of the Total Environment, 661:465–480, 2019.

[6] James J. Elser, Matthew E S Bracken, Elsa E. Cleland, Daniel S. Gruner, W. Stanley Harpole, Helmut Hillebrand, Jacqueline T. Ngai, Eric W. Seabloom, Jonathan B. Shurin, and Jennifer E. Smith.
Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems.
Ecology Letters, 10(12):1135–1142, 2007.

[7] C. M. Moore, M. M. Mills, K. R. Arrigo, I. Berman-Frank, L. Bopp, P. W. Boyd, E. D. Galbraith, R. J. Geider, C. Guieu, S. L. Jaccard, T. D. Jickells, J. La Roche, T. M. Lenton, N. M. Mahowald, E. Marañón, I. Marinov, J. K. Moore, T. Nakatsuka, A. Oschlies, M. A. Saito, T. F. Thingstad, A. Tsuda, and O. Ulloa.
Processes and patterns of oceanic nutrient limitation.
Nature Geoscience, 6(9):701–710, 2013.

[8] Stefanie Müller and Simon M. Mitrovic.
Phytoplankton co-limitation by nitrogen and phosphorus in a shallow reservoir: progressing from the phosphorus limitation paradigm.
Hydrobiologia, 744(1):255–269, 2015.

[9] Jian Zhou, Xiaoxia Han, Justin D. Brookes, and Boqiang Qin.
High probability of nitrogen and phosphorus co-limitation occurring in eutrophic lakes.
Environmental Pollution, 292(PA):118276, 2022.

[10] D. J. Scanlan, M. Ostrowski, S. Mazard, A. Dufresne, L. Garczarek, W. R. Hess, A. F. Post, M. Hagemann, I. Paulsen, and F. Partensky.
Ecological Genomics of Marine Picocyanobacteria.
Microbiology and Molecular Biology Reviews, 73(2):249–299, 2009.

[11] J. Díez, A. López-Lozano, M. A. Domínguez-Martín, G. Gómez-Baena, M. C. Muñoz-Marín, Y. Melero-Rubio, and J. M. García-Fernández.
Regulatory and metabolic adaptations in the nitrogen assimilation of marine picocyanobacteria.

FEMS microbiology reviews, 47(1):1–23, 2023.

[12] Kevin J. Erratt, Irena F. Creed, and Charles G. Trick.
Comparative effects of ammonium, nitrate and urea on growth and photosynthetic efficiency of three bloom-forming cyanobacteria.
Freshwater Biology, 63(7):626–638, 2018.

[13] P.V. Bhaskar, Jane T. Bhaskar, Shridhar Jawak, R. Jyothibabu, and Neelima Mishra.
Mixotrophic plankton and *Synechococcus* distribution in waters around Svalbard, Norway during June 2019.
Polar Science, 30:100697, 2021.

[14] Leilei Bai, Chicheng Cao, Changhui Wang, Huacheng Xu, Hui Zhang, Vera I. Slaveykova, and Helong Jiang.
Toward Quantitative Understanding of the Bioavailability of Dissolved Organic Matter in Freshwater Lake during Cyanobacteria Blooming.
Environmental Science and Technology, 51(11):6018–6026, 2017.

[15] Cristiana Callieri.
Single cells and microcolonies of freshwater picocyanobacteria: A common ecology.
Journal of Limnology, 69(2):257–277, 2010.

[16] Sylwia Śliwińska-Wilczewska, Jakub Maculewicz, Aldo Barreiro Felpeto, and Adam Latała.
Allelopathic and bloom-forming picocyanobacteria in a changing world.
Toxins, 10(1):1–20, 2018.

[17] Patrick Van Hove, Warwick F. Vincent, Pierre E. Galand, and Annick Wilmotte.
Abundance and diversity of picocyanobacteria in High Arctic lakes and fjords.
Algological Studies, 126(1):209–227, 2009.

[18] Paris Lavin, Bernardo González, J. Francisco Santibáñez, David J. Scanlan, and Osvaldo Ulloa.
Novel lineages of *Prochlorococcus* thrive within the oxygen minimum zone of the eastern tropical South Pacific.
Environmental Microbiology Reports, 2(6):728–738, 2010.

[19] Andrea Mózes, Mátyás Présing, and Lajos Vörös.
Seasonal dynamics of picocyanobacteria and picoeukaryotes in a large shallow lake (Lake Balaton, Hungary).
International Review of Hydrobiology, 91(1):38–50, 2006.

[20] Brigitte F. Lavallée and Frances R. Pick.
Picocyanobacteria abundance in relation to growth and loss rates in oligotrophic to mesotrophic lakes.
Aquatic Microbial Ecology, 27(1):37–46, 2002.

[21] Ben K. Bum and Frances R. Pick.
Factors regulating phytoplankton and zooplankton biomass in temperate rivers.
Limnology and Oceanography, 41(7):1572–1577, 1996.

[22] Toshi Nagata, Yasuhiro Satoh, Yasunori Watanabe, Tatyana Kostrnova, Nina Bondarenko, Valentine Drucker, Ryoji Nakazato, Nina Guselnikova, Kyoko Kawanobe, Oliga Mologawaya, Kenji Takai, and Dong-Sup Kim.
Autotrophic picoplankton in southern Lake Baikal: abundance, growth and grazing mortality during summer.
Journal of Plankton Research, 16(8):945–959, 1994.

[23] J G Stockner and K S Shortreed.
Autotrophic picoplankton community dynamics in a pre-alpine lake in British Columbia, Canada.
Hydrobiologia, 274:133–144, 1994.

[24] Nisha H. Motwani and Elena Gorokhova.
Mesozooplankton grazing on picocyanobacteria in the baltic sea as inferred from molecular diet analysis.
PLoS ONE, 8(11), 2013.

[25] Katrin Schmidt, Antony J. Birchill, Angus Atkinson, Robert J.W. Brewin, James R. Clark, Anna E. Hickman, David G. Johns, Maeve C. Lohan, Angela Milne, Silvia Pardo, Luca Polimene, Tim J. Smyth, Glen A. Tarran, Claire E. Widdicombe, E. Malcolm S. Woodward, and Simon J. Ussher.

Increasing picocyanobacteria success in shelf waters contributes to long-term food web degradation.
Global Change Biology, 26(10):5574–5587, 2020.

[26] Susanna A. Wood, Marcia Y. Maier, Jonathan Puddick, Xavier Pochon, Anastasija Zaiko, Daniel R. Dietrich, and David P. Hamilton.
Trophic state and geographic gradients influence planktonic cyanobacterial diversity and distribution in New Zealand lakes.
FEMS Microbiology Ecology, 93(2):1–13, 2017.

[27] Lajos Vörös, Cristiana Callieri, Katalin V-Balogh, and Roberto Bertoni.
Freshwater picocyanobacteria along a trophic gradient and light quality range.
Hydrobiologia, 369-370:117–125, 1998.

[28] Xiuren Ning, James E. Cloern, and Brian E. Cole.
Spatial and temporal variability of picocyanobacteria *Synechococcus* sp. in San Francisco Bay.
Limnology and Oceanography, 45(3):695–702, 2000.

[29] Heather A. Bouman, Osvaldo Ulloa, David J. Scanlan, Katrin Zwirglmaier, William K.W. Li, Trevor Platt, Venetia Stuart, Ray Barlow, Ole Leth, Lesley Clementson, Vivian Lutz, Masao Fukasawa, Shuichi Watanabe, and Shubha Sathyendranath.
Oceanographic Basis of the Global Surface Distribution of *Prochlorococcus* Ecotypes.
Science, 312(5775):918–921, 2006.

[30] Wei Yan, Shuzhen Wei, Qiong Wang, Xilin Xiao, Qinglu Zeng, Nianzhi Jiao, and Rui Zhang.
Genome Rearrangement Shapes *Prochlorococcus* Ecological Adaptation.
Applied and Environmental Microbiology, 84(17):1–13, 2018.

[31] Laia Armengol, Albert Calbet, Gara Franchy, Adriana Rodríguez-Santos, and Santiago Hernández-León.
Planktonic food web structure and trophic transfer efficiency along a productivity gradient in the tropical and subtropical Atlantic Ocean.
Scientific Reports, 9(1):1–19, 2019.

[32] Michael J. Behrenfeld, Robert T. O'Malley, David A. Siegel, Charles R. McClain, Jorge L. Sarmiento, Gene C. Feldman, Allen J. Milligan, Paul G. Falkowski, Ricardo M. Letelier, and Emmanuel S. Boss.
Climate-driven trends in contemporary ocean productivity.
Nature, 444(7120):752–755, 2006.

[33] Andrea Di Cesare, Pedro J. Cabello-Yeves, Nathan A.M. Chrismas, Patricia Sánchez-Baracaldo, Michaela M. Salcher, and Cristiana Callieri.
Genome analysis of the freshwater planktonic *Vulcanococcus limneticus* sp. nov. reveals horizontal transfer of nitrogenase operon and alternative pathways of nitrogen utilization.
BMC Genomics, 19(1):1–12, 2018.

[34] Elliot Druce, Michele Grego, Henk Bolhuis, Penny J Johnes, and Patricia Sánchez-baracaldo.
Draft Genome Sequences of *Synechococcus* sp. strains CCAP1479/9, CCAP1479/10, CCAP1479/13, CCY0621, and CCY9618: Five Freshwater Syn/Pro Clade Picocyanobacteria.
Journal of Genomics, 11:26–36, 2023.

[35] Marc Schallenberg and Carolyn W. Burns.
Tests of autotrophic picoplankton as early indicators of nutrient enrichment in an ultra-oligotrophic lake.
Freshwater Biology, 46(1):27–37, 2001.

[36] Naoshi Fujimoto, Keigo Mizuno, Tomoki Yokoyama, Akihiro Ohnishi, Masaharu Suzuki, Satoru Watanabe, Kenji Komatsu, Yoichi Sakata, Naohiro Kishida, Michihiro Akiba, and Satoko Matsukura.
Community analysis of picocyanobacteria in an oligotrophic lake by cloning 16S rRNA gene and 16S rRNA gene amplicon sequencing.
The Journal of General and Applied Microbiology, 61(5):171–176, 2015.

[37] Joachim Ruber, Franziska R. Bauer, Andrew D. Millard, Uta Raeder, Juergen Geist, and Katrin Zwirglmaier.
Synechococcus diversity along a trophic gradient in the Osterseen Lake District, Bavaria.

Microbiology (United Kingdom), 162(12):2053–2063, 2016.

[38] A.M.M. Batista and A. Giani.
Spatiotemporal variability of cyanobacterial community in a Brazilian oligomesotrophic reservoir: The picocyanobacterial dominance.
Ecohydrology & Hydrobiology, 19(4):566–576, 2019.

[39] Alexis Dufresne, Marcel Salanoubat, Frédéric Partensky, François Artiguenave, Ilka M. Axmann, Valérie Barbe, Simone Duprat, Michael Y. Galperin, Eugene V. Koonin, Florence Le Gall, Kira S. Makarova, Martin Ostrowski, Sophie Oztas, Catherine Robert, Igor B. Rogozin, David J. Scanlan, Nicole Tandeau De Marsac, Jean Weissenbach, Patrick Wincker, Yuri I. Wolf, and Wolfgang R. Hess.
Genome sequence of the cyanobacterium *Prochlorococcus marinus* SS120, a nearly minimal oxyphototrophic genome.
Proceedings of the National Academy of Sciences of the United States of America, 100(17):10020–10025, 2003.

[40] Paul W. MacKeigan, Rebecca E. Garner, Marie Ève Monchamp, David A. Walsh, Vera E. Onana, Susanne A. Kraemer, Frances R. Pick, Beatrix E. Beisner, Michael D. Agbeti, Naíla Barbosa da Costa, B. Jesse Shapiro, and Irene Gregory-Eaves.
Comparing microscopy and DNA metabarcoding techniques for identifying cyanobacteria assemblages across hundreds of lakes.
Harmful Algae, 113, 3 2022.

[41] Mia Bužančić, Živana Ninčević Gladan, Ivona Marasović, Grozdan Kušpilić, and Branka Grbec.
Eutrophication influence on phytoplankton community composition in three bays on the eastern Adriatic coast.
Oceanologia, 58(4):302–316, 2016.

[42] Maciej Karpowicz, Piotr Zieliński, Magdalena Grabowska, Jolanta Ejsmont-Karabin, Joanna Kozłowska, and Irina Feniova.

Effect of eutrophication and humification on nutrient cycles and transfer efficiency of matter in freshwater food webs.
Hydrobiologia, 847(11):2521–2540, 2020.

[43] Yuanfeng Cai and Fanxiang Kong.
Diversity and dynamics of picocyanobacteria and bloom-forming cyanobacteria in a large shallow eutrophic lake (lake Chaohu, China).
Journal of Limnology, 72(3):473–484, 2013.

[44] Congmin Zhu, Junyi Zhang, Muhammad Zohaib Nawaz, Shahid Mahboob, Khalid A. Al-Ghanim, Iqrar Ahmad Khan, Zuhong Lu, and Ting Chen.
Seasonal succession and spatial distribution of bacterial community structure in a eutrophic freshwater Lake, Lake Taihu.
Science of the Total Environment, 669:29–40, 2019.

[45] Théo Veaudor, Victoire Blanc-Garin, Célia Chenebault, Encarnación Diaz-Santos, Jean François Sassi, Corinne Cassier-Chauvat, and Franck Chauvat.
Recent advances in the photoautotrophic metabolism of cyanobacteria: Biotechnological implications.
Life, 10(5), 2020.

[46] Vincenzo Mascoli, Ahmad Farhan Bhatti, Luca Bersanini, Herbert van Amerongen, and Roberta Croce.
The antenna of far-red absorbing cyanobacteria increases both absorption and quantum efficiency of Photosystem II.
Nature Communications, 13(1), 2022.

[47] Shubha Sathyendranath, Luigi Lazzara, and Louis Prieur.
Variations in the spectral values of specific absorption of phytoplankton.
Limnology and Oceanography, 32(2):403–415, 1987.

[48] B. Palenik, B. Brahamsha, F. W. Larimer, M. Land, L. Hauser, P. Chain, J. Lamerdin, W. Regala, E. E. Allen, J. McCarren, I. Paulsen, A. Dufresne, F. Partensky, E. A. Webb, and J. Waterbury.

The genome of a motile marine *Synechococcus*.
Nature, 424(6952):1037–1042, 2003.

[49] Antonio Camacho, María R. Miracle, and Eduardo Vicente.
Which factors determine the abundance and distribution of picocyanobacteria in inland waters? A comparison among different types of lakes and ponds.
Archiv für Hydrobiologie, 157(3):321–338, 2003.

[50] Maayke Stomp, Jef Huisman, Lajos Vörös, Frances R. Pick, Maria Laamanen, Thomas Haverkamp, and Lucas J. Stal.
Colourful coexistence of red and green picocyanobacteria in lakes and seas.
Ecology Letters, 10(4):290–298, 2007.

[51] Cristiana Callieri.
Micro-players for macro-roles: Aquatic microbes in deep lakes.
Journal of Limnology, 75(1S):191–200, 2016.

[52] Sven Becker, Patricia Sánchez-Baracaldo, Arvind K. Singh, and Paul K. Hayes.
Spatio-temporal niche partitioning of closely related picocyanobacteria clades and phycocyanin pigment types in Lake Constance (Germany).
FEMS Microbiology Ecology, 80(2):488–500, 2012.

[53] Boglárka Somogyi, Tamás Felföldi, József Vanyovszki, Ákos Ágyi, Károly Márialigeti, and Lajos Vörös.
Winter bloom of picoeukaryotes in Hungarian shallow turbid soda pans and the role of light and temperature.
Aquatic Ecology, 43(3):735–744, 2009.

[54] John Beardall.
Blooms of emphSynechococcus: An analysis of the problem worldwide and possible causative factors in relation to nuisance blooms in the Gippsland Lakes.
Technical report, Monash University, Clayton, VIC, Australia, 2008.

[55] A. I Kopylov, A. V. Romanenko, E. A Zabotkina, N. M. Mineeva, I. N. Krylova, and T. S Maslennikova.
Picocyanobacteria in eutrophic reservoirs of the Middle Volga: abundance, production, viral infection.
Zh Obshch Biol, 75(3):234–44, 2014.

[56] María Cristina Casero, David Velázquez, Miguel Medina-Cobo, Antonio Quesada, and Samuel Cirés.
Unmasking the identity of toxigenic cyanobacteria driving a multi-toxin bloom by high-throughput sequencing of cyanotoxins genes and 16S rRNA metabarcoding.
Science of the Total Environment, 665:367–378, 2019.

[57] Iwona Bubak, Sylwia Sliwinka-Wilczewska, Paulina Głowacka, Agnieszka Szczerba, and Katarzyna Mozdzen.
The Importance of Allelopathic Picocyanobacterium *Synechococcus* sp. on the Abundance, Biomass Formation, and Structure of Phytoplankton Assemblages in Three Freshwater Lakes.
Toxins, 12(4):259, 2020.

[58] P. Yu Sorokin, Yu I. Sorokin, R. Boscolo, and O. Giovanardi.
Bloom of picocyanobacteria in the Venice lagoon during summer-autumn 2001: Ecological sequences.
Hydrobiologia, 523(1-3):71–85, 2004.

[59] Yuri I. Sorokin and Olga Yu Zakuskina.
Features of the Comacchio ecosystem transformed during persistent bloom of picocyanobacteria.
Journal of Oceanography, 66(3):373–387, 2010.

[60] Jiajun Li, Zuozhi Chen, Zhiyou Jing, Linbin Zhou, Gang Li, Zhixin Ke, Xin Jiang, Jiaxing Liu, Huaxue Liu, and Y. Tan.
Synechococcus bloom in the Pearl River Estuary and adjacent coastal area–With special focus on flooding during wet seasons.

Science of the Total Environment, 692:769–783, 2019.

[61] Michael C. Murrell and Emile M. Lores.
Phytoplankton and zooplankton seasonal dynamics in a subtropical estuary: Importance of cyanobacteria.
Journal of Plankton Research, 26(3):371–382, 2004.

[62] Edward J Philips, Susan Badylak, and Tammy C Lynch.
Blooms of the picoplanktonic in Florida Bay , a cyanobacterium *Synechococcus* subtropical.
Limnology and Oceanography, 44(4):1166–1175, 1999.

[63] Shu Harn Te, Jerome Wai Kit Kok, Rong Luo, Luhua You, Nur Hanisah Sukarji, Kwan Chien Goh, Zhi Yang Sim, Dong Zhang, Yiliang He, and Karina Yew Hoong Gin.
Coexistence of *Synechococcus* and *Microcystis* Blooms in a Tropical Urban Reservoir and Their Links with Microbiomes.
Environmental Science and Technology, 2022.

[64] Nathan S. Hall, R. Wayne Litaker, W. Judson Kenworthy, Mark W. Vandersea, William G. Sunda, James P. Reid, Daniel H. Slone, and Susan Butler.
Consortial brown tide - picocyanobacteria blooms in Guantánamo Bay, Cuba.
Harmful Algae, 73:30–43, 2018.

[65] Yasuo Nakamura, Kentaro Suzuki, Shin Ya Suzuki, and Juro Hiromi.
Production of *Oikopleura dioica* (Appendicularia) following a picoplankton 'bloom' in a eutrophic coastal area.
Journal of Plankton Research, 19(1):113–124, 1997.

[66] Harri Kuosa.
Picoplanktonic algae in the northern Baltic Sea: seasonal dynamics and flagellate grazing.
Mar. Ecol. Prog. Ser., 73:269–276, 1991.

[67] Zahit Uysal.
Pigments, size and distribution of *Synechococcus* spp. in the Black Sea.
Journal of Marine Systems, 24(3-4):313–326, 2000.

[68] Boris Wawrik and John H. Paul.
Phytoplankton community structure and productivity along the axis of the Mississippi River plume in oligotrophic Gulf of Mexico waters.
Aquatic Microbial Ecology, 35(2):185–196, 2004.

[69] Kaitlin L. Reinl, Justin D. Brookes, Cayelan C. Carey, Ted D. Harris, Bas W. Ibelings, Ana M. Morales-Williams, Lisette N. De Senerpont Domis, Karen S. Atkins, Peter D.F. Isles, Jorrit P. Mesman, Rebecca L. North, Lars G. Rudstam, Julio A.A. Stelzer, Jason J. Venkiteswaran, Kiyoko Yokota, and Qing Zhan.
Cyanobacterial blooms in oligotrophic lakes: Shifting the high-nutrient paradigm.
Freshwater Biology, 66(9):1846–1859, 2021.

[70] John Stockner, Cristiana Callieri, and Gertrud Cronberg.
Picoplankton and Other Non-Bloom-Forming Cyanobacteria in Lakes.
The Ecology of Cyanobacteria, pages 195–231, 2000.

[71] H. W. Paerl, R. S. Fulton, P. H. Moisander, and J. Dyble.
Harmful freshwater algal blooms, with an emphasis on cyanobacteria.
TheScientificWorldJournal, 1:76–113, 2001.

[72] Trevor A. Gearhart, Katie Ritchie, Evan Nathan, Jason D. Stockwell, and Jana Kraft.
Alteration of essential fatty acids in secondary consumers across a gradient of cyanobacteria.
Hydrobiologia, 784(1):155–170, 2017.

[73] Stefanie M. Hixson and Michael T. Arts.
Climate warming is predicted to reduce omega-3, long-chain, polyunsaturated fatty acid production in phytoplankton.
Global change biology, 22(8):2744–2755, 2016.

[74] Liqing Zeng, Changzhou Yan, Jianhua Guo, Zhuo Zhen, Yuan Zhao, and Dapeng Wang.
Influence of algal blooms decay on arsenic dynamics at the sediment-water interface of a shallow lake.
Chemosphere, 219:1014–1023, 2019.

[75] J. M. O'Neil, T. W. Davis, M. A. Burford, and C. J. Gobler.
The rise of harmful cyanobacteria blooms: The potential roles of eutrophication and climate change.
Harmful Algae, 14:313–334, 2012.

[76] Elke Dittmann, David P. Fewer, and Brett A. Neilan.
Cyanobacterial toxins: Biosynthetic routes and evolutionary roots.
FEMS Microbiology Reviews, 37(1):23–43, 2013.

[77] Thanh Luu Pham and Motoo Utsumi.
An overview of the accumulation of microcystins in aquatic ecosystems.
Journal of Environmental Management, 213:520–529, 2018.

[78] Daria Filatova, Massimo Picardo, Oscar Núñez, and Marinella Farré.
Analysis, levels and seasonal variation of cyanotoxins in freshwater ecosystems.
Trends in Environmental Analytical Chemistry, 26, 2020.

[79] Ana Luiza Fonseca Fortes Furtado, Maria Do Carmo Calijuri, Adriana Sturion Lorenzi, Ricardo Yukio Honda, Diego Bonaldo Genuário, and Marli Fátima Fiore.
Morphological and molecular characterization of cyanobacteria from a Brazilian facultative wastewater stabilization pond and evaluation of microcystin production.
Hydrobiologia, 627(1):195–209, 2009.

[80] Wayne W Carmichael and Renhui Li.
Cyanobacteria toxins in the Salton Sea.
Saline Systems, 13:1–13, 2006.

[81] Natalia Jakubowska and Elzbieta Szeląg-Wasielewska.
Toxic picoplanktonic cyanobacteria - Review.
Marine Drugs, 13(3):1497–1518, 2015.

[82] Lorraine C. Backer, Jan H. Landsberg, Melissa Miller, Kevin Keel, and Tegwin K. Taylor.
Canine cyanotoxin poisonings in the United States (1920s-2012): Review of suspected and confirmed cases from three data sources.

Toxins, 5(9):1597–1628, 9 2013.

[83] Apurva Lad, Joshua D. Breidenbach, Robin C. Su, Jordan Murray, Rebecca Kuang, Alison Mascarenhas, John Najjar, Shivani Patel, Prajwal Hegde, Mirella Youssef, Jason Breuler, Andrew L. Kleinhenz, Andrew P. Ault, Judy A. Westrick, Nikolai N. Modyanov, David J. Kennedy, and Steven T. Haller.
As We Drink and Breathe: Adverse Health Effects of Microcystins and Other Harmful Algal Bloom Toxins in the Liver, Gut, Lungs and Beyond.
Life, 12(3), 3 2022.

[84] Karina Yew Hoong Gin, Zhi Yang Sim, Kwan Chien Goh, Jerome Wai Kit Kok, Shu Harn Te, Ngoc Han Tran, Wenxuan Li, and Yiliang He.
Novel cyanotoxin-producing *Synechococcus* in tropical lakes.
Water Research, 192, 3 2021.

[85] Susan B. Watson, Paul Monis, Peter Baker, and Steven Giglio.
Biochemistry and genetics of taste- and odor-producing cyanobacteria.
Harmful Algae, 54:112–127, 2016.

[86] Henry L Ehrlich.
Geomicrobiology: its significance for geology.
Earth-Science Reviews, 45:45–60, 1998.

[87] Maria Dittrich, Philipp Kurz, and Bernhard Wehrli.
The role of autotrophic picocyanobacteria in calcite precipitation in an oligotrophic lake.
Geomicrobiology Journal, 21(1):45–53, 2004.

[88] I. B.M. Kosamu and M. Obst.
The influence of picocyanobacterial photosynthesis on calcite precipitation.
International Journal of Environmental Science and Technology, 6(4):557–562, 2009.

[89] Fei Xue Fu, Mark E. Warner, Yaohong Zhang, Yuanyuan Feng, and David A. Hutchins.
Effects of increased temperature and CO2 on photosynthesis, growth, and elemental ratios in marine *Synechococcus* and *Prochlorococcus* (Cyanobacteria).

Journal of Phycology, 43(3):485–496, 2007.

[90] Peter Hofmann, Antonis Chatzinotas, W. Stanley Harpole, and Susanne Dunker.
Temperature and stoichiometric dependence of phytoplankton traits.
Ecology, 100(12):1–14, 2019.

[91] Natalia Visintini, Adam C. Martiny, and Pedro Flombaum.
Prochlorococcus, *Synechococcus*, and picoeukaryotic phytoplankton abundances in the global ocean.
Limnology And Oceanography Letters, 6(4):207–215, 2021.

[92] Patricia Sánchez-baracaldo, Giorgio Bianchini, Jamie D Wilson, and Andrew H Knoll.
Cyanobacteria and biogeochemical cycles through Earth history.
Trends in Microbiology, 30(2):143–157, 2021.

[93] Catherine F. Demoulin, Yannick J. Lara, Luc Cornet, Camille François, Denis Baurain, Annick Wilmotte, and Emmanuelle J. Javaux.
Cyanobacteria evolution: Insight from the fossil record.
Free Radical Biology and Medicine, 140:206–223, 2019.

[94] Joanne S. Boden, Kurt O. Konhauser, Leslie J. Robbins, and Patricia Sánchez-Baracaldo.
Timing the evolution of antioxidant enzymes in cyanobacteria.
Nature Communications, 12(1), 2021.

[95] P. M. Shih, J. Hemp, L. M. Ward, N. J. Matzke, and W. W. Fischer.
Crown group Oxyphotobacteria postdate the rise of oxygen.
Geobiology, 15(1):19–29, 2017.

[96] Patricia Sánchez-Baracaldo, Giorgio Bianchini, Jamie D. Wilson, and Andrew H. Knoll.
Cyanobacteria and biogeochemical cycles through Earth history.
Trends in Microbiology, 30(2):143–157, 2022.

[97] C. E. Blank and P. Sanchez-Baracaldo.

Timing of morphological and ecological innovations in the cyanobacteria - A key to understanding the rise in atmospheric oxygen.
Geobiology, 8(1):1–23, 2010.

[98] Patricia Sánchez-Baracaldo.
Origin of marine planktonic cyanobacteria.
Scientific Reports, 5:14–17, 2015.

[99] Lynne Whitehead, Benedict M. Long, G. Dean Price, and Murray R. Badger.
Comparing the in vivo function of α-carboxysomes and β-carboxysomes in two model cyanobacteria.
Plant Physiology, 165(1):398–411, 2014.

[100] P. Sánchez-Baracaldo, P. K. Hayes, and C. E. Blank.
Morphological and habitat evolution in the Cyanobacteria using a compartmentalization approach.
Geobiology, 3(3):145–165, 2005.

[101] Sijun Huang, Steven W Wilhelm, H Rodger Harvey, Karen Taylor, Nianzhi Jiao, and Feng Chen.
Novel lineages of *Prochlorococcus* and *Synechococcus* in the global oceans.
The ISME Journal, pages 285–297, 2012.

[102] Ahmed A. Shibl, Luke R. Thompson, David K. Ngugi, and Ulrich Stingl.
Distribution and diversity of *Prochlorococcus* ecotypes in the Red Sea.
FEMS Microbiology Letters, 356(1):118–126, 2014.

[103] Pedro J. Cabello-Yeves, Cristiana Callieri, Antonio Picazo, Lena Schallenberg, Paula Huber, Juan J. Roda-Garcia, Maciej Bartosiewicz, Olga I. Belykh, Irina V. Tikhonova, Alberto Torcello-Requena, Paula Martin De Prado, Richard J. Puxty, Andrew D. Millard, Antonio Camacho, Francisco Rodriguez-Valera, and David J. Scanlan.
Elucidating the picocyanobacteria salinity divide through ecogenomics of new freshwater isolates.
BMC Biology, 20(1):1–24, 2022.

[104] Otakar Strunecký, Anna Pavlovna Ivanova, and Jan Mareš.
An updated classification of cyanobacterial orders and families based on phylogenomic and polyphasic analysis.
Journal of Phycology, 59(1):12–51, 2023.

[105] Carrine E. Blank.
Phylogenetic distribution of compatible solute synthesis genes support a freshwater origin for cyanobacteria.
Journal of Phycology, 49(5):880–895, 2013.

[106] Maria Sofia Costa, Margarida Costa, Vítor Ramos, Pedro N. Leão, Aldo Barreiro, Vítor Vasconcelos, and Rosário Martins.
Picocyanobacteria from a Clade of Marine Cyanobium Revealed Bioactive Potential Against Microalgae, Bacteria, and Marine Invertebrates.
Journal of Toxicology and Environmental Health - Part A: Current Issues, 78(7):432–442, 2015.

[107] Eric D. Becraft, Jason M. Wood, Frederick M. Cohan, and David M. Ward.
Biogeography of American Northwest Hot Spring A/B-Lineage *Synechococcus* Populations.
Frontiers in Microbiology, 11, 2020.

[108] Patricia Sánchez-Baracaldo, Giorgio Bianchini, Andrea Di Cesare, Cristiana Callieri, and Nathan A. M. Chrismas.
Insights Into the Evolution of Picocyanobacteria and Phycoerythrin Genes (*mpeBA* and *cpeBA*).
Frontiers in Microbiology, 10:1–17, 2019.

[109] Takao Kondo, Carl A Strayert, Resham D Kulkarnit, Walter Taylor, Masahiro Ishiura, Susan S Goldent, Carl Hirschie, and Johnson Ii.
Circadian rhythms in prokaryotes: Luciferase as a reporter of circadian gene expression in cyanobacteria.
Proc. Natl. Acad. Sci. USA, 90(12):5672–5676, 1993.

[110] Athina Zouni, Horst-Tobias Witt, Jan Kern, Petra Fromme, Norbert Krauss, Wolfram Saenger, and Peter Orth.

Crystal structure of photosystem II from *Synechococcus elongatus* at 3.8 A resolution.
Nature, 409:739–743, 2001.

[111] P Jordan, P Fromme, O Klukas, H T Witt, W Saenger, and N Krauß.
Three-dimensional structure of cyanobacterial photosystem I at 2.5 Å resolution.
Nature, 411:909–917, 2001.

[112] Dominik Martin-Creuzburg, Erik Sperfeld, and Alexander Wacker.
Colimitation of a freshwater herbivore by sterols and polyunsaturated fatty acids.
Proceedings of the Royal Society B: Biological Sciences, 276(1663):1805–1814, 5 2009.

[113] Nina Schlotz, Michael Pester, Heike M. Freese, and Dominik Martin-Creuzburg.
A dietary polyunsaturated fatty acid improves consumer performance during challenge with an opportunistic bacterial pathogen.
FEMS Microbiology Ecology, 90(2):467–477, 11 2014.

[114] Dominik Martin-Creuzburg, Alexander Wacker, and Timo Basena.
Interactions between limiting nutrients: Consequences for somatic and population growth of *Daphnia magna*.
Limnology and Oceanography, 55(6):2597–2607, 11 2010.

[115] Thomas Weisse and Michael Moser.
Light affects picocyanobacterial grazing and growth response of the mixotrophic flagellate *Poterioochromonas malhamensis*.
Journal of Microbiology, 58(4):268–278, 4 2020.

[116] Pia I. Scherer, Carolin Absmeier, Maria Urban, Uta Raeder, Juergen Geist, and Katrin Zwirglmaier.
Influence of cyanobacteria, mixotrophic flagellates, and virioplankton size fraction on transcription of microcystin synthesis genes in the toxic cyanobacterium *Microcystis aeruginosa*.
MicrobiologyOpen, 7(1), 2 2018.

[117] Samantha L. Evans, William T. Anderson, and Frank J. Jochem.
Spatial variability in Florida Bay particulate organic matter composition: Combining flow cytometry with stable isotope analyses.

Hydrobiologia, 569(1):151–165, 10 2006.

[118] Přemysl Mikula, Marie Mlnaříková, Enrico T. Nadres, Haruko Takahashi, Pavel Babica, Kenichi Kuroda, Luděk Bláha, and Iva Sovadinová.
Synthetic biomimetic polymethacrylates: Promising platform for the design of anti-cyanobacterial and anti-algal agents.
Polymers, 13(7), 4 2021.

[119] Meng Yun Chen, Wen Kai Teng, Liang Zhao, Chun Xiang Hu, Yang Kai Zhou, Bo Ping Han, Li Rong Song, and Wen Sheng Shu.
Comparative genomics reveals insights into cyanobacterial evolution and habitat adaptation.
ISME Journal, 2020.

[120] Lu Zhang, Feng Xin Chen, Zhan Zeng, Mengjiao Xu, Fangfang Sun, Liu Yang, Xiaoyue Bi, Yanjie Lin, Yuan Jiao Gao, Hong Xiao Hao, Wei Yi, Minghui Li, and Yao Xie.
Advances in Metagenomics and Its Application in Environmental Microorganisms.
Frontiers in Microbiology, 12:1–15, 2021.

[121] Tim M. Blackburn, Phillip Cassey, and Kevin J. Gaston.
Variations on a theme: Sources of heterogeneity in the form of the interspecific relationship between abundance and distribution.
Journal of Animal Ecology, 75(6):1426–1439, 2006.

[122] Patricia Sánchez-Baracaldo, Barbara A. Handley, and Paul K. Hayest.
Picocyanobacterial community structure of freshwater lakes and the Baltic Sea revealed by phylogenetic analyses and clade-specific quantitative PCR.
Microbiology, 154(11):3347–3357, 2008.

[123] Jonathan P. Zehr, Bethany D. Jenkins, Steven M. Short, and Grieg F. Steward.
Nitrogenase gene diversity and microbial community structure: A cross-system comparison.
Environmental Microbiology, 5(7):539–554, 2003.

[124] Lasse Riemann, Hanna Farnelid, and Grieg F. Steward.

Nitrogenase genes in non-cyanobacterial plankton: Prevalence, diversity and regulation in marine waters.
Aquatic Microbial Ecology, 61(3):235–247, 2010.

[125] Walter K. Dodds and Eugene B. Welch.
Establishing nutrient criteria in streams.
Journal of the North American Benthological Society, 19(1):186–196, 2000.

[126] G. J.C. Underwood and J. Kromkamp.
Primary Production by Phytoplankton and Microphytobenthos in Estuaries, volume 29.
Elsevier Masson SAS, 1999.

[127] Patrick Durand, Lutz Breuer, Penny J. Johnes, Gilles Billen, Andrea Butturini, Gilles Pinay, Hans van Grinsven, Josette Garnier, Michael Rivett, David S. Reay, Chris Curtis, Jan Siemens, Stephen Maberly, Øyvind Kaste, Christoph Humborg, Roos Loeb, Jeroen de Klein, Josef Hejzlar, Nikos Skoulikidis, Pirkko Kortelainen, Ahti Lepistö, and Richard Wright.
Nitrogen processes in aquatic ecosystems.
In *The European Nitrogen Assessment*, pages 126–146. Cambridge University Press, 2011.

[128] H. P. Jarvie, T. D. Jickells, R. A. Skeffington, and P. J.A. Withers.
Climate change and coupling of macronutrient cycles along the atmospheric, terrestrial, freshwater and estuarine continuum.
Science of the Total Environment, 434:252–258, 2012.

[129] Fabio Lepori and François Keck.
Effects of atmospheric nitrogen deposition on remote freshwater ecosystems.
Ambio, 41(3):235–246, 2012.

[130] Ann Kristin Bergström and Mats Jansson.
Atmospheric nitrogen deposition has caused nitrogen enrichment and eutrophication of lakes in the northern hemisphere.
Global Change Biology, 12(4):635–643, 2006.

[131] Tom Berman and Deborah A. Bronk.
Dissolved organic nitrogen: A dynamic participant in aquatic ecosystems.
Aquatic Microbial Ecology, 31(3):279–305, 2003.

[132] Camila Fernandez, Laura Farías, and Osvaldo Ulloa.
Nitrogen fixation in denitrified marine waters.
PLoS ONE, 6(6), 2011.

[133] Devaki Bhaya, Arthur R. Grossman, Anne Soisig Steunou, Natalia Khuri, Frederick M. Cohan, Natsuko Hamamura, Melanie C. Melendrez, Mary M. Bateson, David M. Ward, and John F. Heidelberg.
Population level functional diversity in a microbial community revealed by comparative genomic and metagenomic analyses.
ISME Journal, 1(8):703–713, 2007.

[134] Peter M Vitousek, John D Aber, Robert W Howarth, Gene E Likens, Pamela A Matson, David W Schindler, William H Schlesinger, and David G Tilman.
Human Alteration of the Global Nitrogen Cycle: Sources and Consequences.
Ecological Applications, 7(3):737–750, 1997.

[135] LJ Puckett.
Nonpoint and point sources of nitrogen in major watersheds of the United States.
Technical report, U.S. Geological Survey, 1994.

[136] W. H. Schlesinger.
On the fate of anthropogenic nitrogen.
Proceedings of the National Academy of Sciences, 106(1):203–208, 2009.

[137] W. S. Zhang, D. P. Swaney, X. Y. Li, B. Hong, R. W. Howarth, and S. H. Ding.
Anthropogenic point-source and non-point-source nitrogen inputs into Huai River basin and their impacts on riverine ammonia-nitrogen flux.
Biogeosciences, 12(14):4275–4289, 2015.

[138] Robert W. Howarth.
Coastal nitrogen pollution: A review of sources and trends globally and regionally.
Harmful Algae, 8(1):14–20, 2008.

[139] David Forman, Samim Al-Dabbagh, and Richard Doll.
Nitrates, nitrites and gastric cancer in Great Britain.
Nature, 313(6004):620–625, 1985.

[140] WHO.
Guidelines for Drinking-water Quality.
Technical report, WHO, Geneva, 2011.

[141] I.-N. Kim, T.-W. Kim, N. Gruber, J. L. Bullister, S. Yang, D. M. Karl, and K. Lee.
Increasing anthropogenic nitrogen in the North Pacific Ocean.
Science, 346(6213):1102–1106, 2014.

[142] Serina Ahlgren, Andras Baky, Sven Bernesson, Åke Nordberg, Olle Norén, and Per Anders Hansson.
Ammonium nitrate fertiliser production based on biomass - Environmental effects from a life cycle perspective.
Bioresource Technology, 99(17):8034–8041, 2008.

[143] DEFRA.
British Survey of Fertiliser Practice 2017.
Technical report, Department for Environment, Food, and Rural Affairs, London, 2018.

[144] Patricia M. Glibert, John Harrison, Cynthia Heil, and Sybil Seitzinger.
Escalating worldwide use of urea - A global change contributing to coastal eutrophication.
Biogeochemistry, 77(3):441–463, 2006.

[145] Jihua Li, Jibiao Zhang, Wei Huang, Fanlong Kong, Yue Li, Min Xi, and Zheng Zheng.
Comparative bioavailability of ammonium, nitrate, nitrite and urea to typically harmful cyanobacterium *Microcystis aeruginosa.*
Marine Pollution Bulletin, 110(1):93–98, 2016.

[146] Penny J. Johnes, Richard P. Evershed, Davey L. Jones, and Stephen C. Maberly.
Exploring the nature, origins and ecological significance of dissolved organic matter in freshwaters: state of the science and new directions.
Biogeochemistry, 2023.

[147] Heonseop Eom, Douglas Borgatti, Hans W. Paerl, and Chul Park.
Formation of Low-Molecular-Weight Dissolved Organic Nitrogen in Predenitrification Biological Nutrient Removal Systems and Its Impact on Eutrophication in Coastal Waters.
Environmental Science and Technology, 51(7):3776–3783, 2017.

[148] Weiying Feng, Shasha Liu, Cuicui Li, Xiaofeng Li, Fanhao Song, Beibei Wang, Haiyan Chen, and Fengchang Wu.
Algal uptake of hydrophilic and hydrophobic dissolved organic nitrogen in the eutrophic lakes.
Chemosphere, 214:295–302, 2019.

[149] Mary G. Lusk and Gurpal S. Toor.
Dissolved organic nitrogen in urban streams: Biodegradability and molecular composition studies.
Water Research, 96:225–235, 2016.

[150] Abubakari Said Mgelwa, Zabibu Kabalika, and Ya Lin Hu.
Increasing importance of nitrate-nitrogen and organic nitrogen concentrations in bulk and throughfall precipitation across urban forests in southern China.
Global Ecology and Conservation, 22:e00983, 2020.

[151] Brian A. Pellerin, Sujay S. Kaushal, and William H. McDowell.
Does anthropogenic nitrogen enrichment increase organic nitrogen concentrations in runoff from forested and human-dominated watersheds?
Ecosystems, 9(5):852–864, 2006.

[152] Björn Gücker, Ricky C.S. Silva, Daniel Graeber, José A.F. Monteiro, E. N.Jack Brookshire, Ronaldo C. Chaves, and Iola G. Boëchat.
Dissolved nutrient exports from natural and human-impacted Neotropical catchments.

Global Ecology and Biogeography, 25(4):378–390, 2016.

[153] Iola Gonçalves Boëchat, Aparecida Beatriz das Mercês Paiva-Magela, Helbert Rocha Reis, and Björn Gücker.
Dissolved organic N in shallow tropical lakes and reservoirs: Contribution to total dissolved N and relationships with eutrophication.
International Review of Hydrobiology, 104(5-6):106–115, 2019.

[154] E. B. Mackay, H. Feuchtmayr, M. M. De Ville, S. J. Thackeray, N. Callaghan, M. Marshall, G. Rhodes, C. A. Yates, P. J. Johnes, and S. C. Maberly.
Dissolved organic nutrient uptake by riverine phytoplankton varies along a gradient of nutrient enrichment.
Science of the Total Environment, 722:137837, 2020.

[155] Xiaojun Cui and Kwang Ho Choo.
Granular iron oxide adsorbents to control natural organic matter and membrane fouling in ultrafiltration water treatment.
Water Research, 47(13):4227–4237, 2013.

[156] David L. Jones, David Shannon, Daniel V. Murphy, and John Farrar.
Role of dissolved organic nitrogen (DON) in soil N cycling in grassland soils.
Soil Biology and Biochemistry, 36(5):749–756, 2004.

[157] Heonseop Eom and Chul Park.
Investigation of characteristics of effluent DON derived from conventional activated sludge (CAS) and predenitrification biological removal (BNR): In terms of proteins and humic substances.
Environmental Research, 196(September 2020):110912, 2021.

[158] Francesca L. Brailsford, Helen C. Glanville, Miles R. Marshall, Christopher A. Yates, Alun T. Owen, Peter N. Golyshin, Penny J. Johnes, and Davey L. Jones.
Land cover and nutrient enrichment regulates low-molecular weight dissolved organic matter turnover in freshwater ecosystems.

Limnology and Oceanography, 66(8):2979–2987, 2021.

[159] Chao Qin, Haizhou Liu, Lei Liu, Scott Smith, David L. Sedlak, and April Z. Gu.
Bioavailability and characterization of dissolved organic nitrogen and dissolved organic phosphorus in wastewater effluents.
Science of the Total Environment, 511:47–53, 2015.

[160] Dorothea Fiedler, Daniel Graeber, Maria Badrian, and Jan Köhler.
Growth response of four freshwater algal species to dissolved organic nitrogen of different concentration and complexity.
Freshwater Biology, 60(8):1613–1621, 2015.

[161] EEC.
Council Directive 91/676/EEC concerning the protection of waters against pollution caused by nitrates from agricultural sources.
Off J Eur Communities, 1991.

[162] EC.
Directive 2000/60/EC of the European Parliament and the Council establishing a framework for community action in the field of water policy.
Off J Eur Communities, 2000.

[163] EC.
Directive 2008/56/EC of the European Parliament and the Council establishing a framework for community action in the field of marine environmental policy.
Off J Eur Communities, 2008.

[164] R. Kapp, S. E. Stevens, and J. L. Fox.
A survey of available nitrogen sources for the growth of the blue-green alga, *Agmenellum quadruplicatum.*
Archives of Microbiology, 104(1):135–138, 1975.

[165] Aladair H. Neilson and Tommy Larsson.

The utilization of organic nitrogen for growth of algae: physiological aspects.
Physiologia Plantarum, 48(4):542–553, 1980.

[166] D. M. Rawson.
The effects of exogenous amino acids on growth and nitrogenase activity in the cyanobacterium *Anabaena cylindrica* PCC 7122.
Journal of General Microbiology, 131(10):2549–2554, 1985.

[167] A. Vaishampayan.
Amino Acid Nutrition in the Blue-Green Alga *Nostoc Muscorum*.
New Phytologist, 90(3):545–549, 1982.

[168] Jyotirmoy Bhattacharya, Arvind Kumar Singh, and Amar Nath Rai.
Nitrogen nutrition in the cyanobacterium *Nostoc ANTH*, a symbiotic isolate from *Anthoceros*: Uptake and assimilation of inorganic-N and amino acids.
Indian Journal of Biochemistry & Biophysics, 39:163–169, 2002.

[169] Ruihua Dai, Huijuan Liu, Jiuhui Qu, Xu Zhao, and Yining Hou.
Effects of amino acids on microcystin production of the *Microcystis aeruginosa*.
Journal of Hazardous Materials, 161(2-3):730–736, 2009.

[170] G Shanthi, M Premalatha, and N Anantharaman.
Effects of L-amino acids as organic nitrogen source on the growth rate, biochemical composition and polyphenol content of *Spirulina platensis*.
Algal Research, 35:471–478, 2018.

[171] Karel Horňák, Helen Schmidheiny, and Jakob Pernthaler.
High-throughput determination of dissolved free amino acids in unconcentrated freshwater by ion-pairing liquid chromatography and mass spectrometry.
Journal of Chromatography A, 1440:85–93, 2016.

[172] Yuan Shen, Francis H Chapelle, Eric W Strom, and Ronald Benner.
Origins and bioavailability of dissolved organic matter in groundwater.
Biogeochemistry, 122(1):61–78, 2015.

[173] J. D. Thomas and P. Eaton.

The spatio-temporal patterns and ecological significance of free amino acids and humic substances in contrasting oligotrophic ad eutrophic freshwater ecosystems.

Hydrobiologia, 332(3):183–211, 1996.

[174] Hakan Alyuruk and Aynur Kontas.

Dissolved free, total and particulate enantiomeric amino acid levels in eutrophic and oligotrophic parts of a semi-enclosed bay (İzmir, Aegean Sea).

Regional Studies in Marine Science, 44:101750, 2021.

[175] Ewen Silvester, Terry Karis, Anne Yusuf, John Pengelly, Samantha Grover, and Gavin N Rees.

Organic carbon and nitrogen dynamics during a peatland storm event : How dissolved combined amino acids reveal the spatial and temporal separation of organic molecules.

Journal of Hydrology, 597(October 2020):126191, 2021.

[176] Rubén J. Lara, Volker Rachold, Gerhard Kattner, Hans W. Hubberten, Georg Guggenberger, Annelie Skoog, and David N. Thomas.

Dissolved organic matter and nutrients in the Lena River, Siberian Arctic: Characteristics and distribution.

Marine Chemistry, 59(3-4):301–309, 1998.

[177] Rachel E Sipler and Deborah A Bronk.

Dynamics of Dissolved Organic Nitrogen.

Elsevier Inc., second edi edition, 2015.

[178] Liguang Li, Zhenli L He, Malak M Tfaily, Patrick Inglett, and Peter J Stoffella.

Spatial-temporal variations of dissolved organic nitrogen molecular composition in agricultural runoff water.

Water Research, 137:375–383, 2018.

[179] Hugo Sarmento, Cristina Romera-Castillo, Markus Lindh, Jarone Pinhassi, M. Montserrat Sala, Josep M. Gasol, Celia Marrasé, and Gordon T. Taylor.

Phytoplankton species-specific release of dissolved free amino acids and their selective consumption by bacteria.
Limnology and Oceanography, 58(3):1123–1135, 2013.

[180] M. Isabel Muro-Pastor, Jose C. Reyes, and Francisco J. Florencio.
Ammonium assimilation in cyanobacteria.
Photosynthesis Research, 83(2):135–150, 2005.

[181] Enrique Flores, José E. Frías, Luis M Rubio, and Antonia Herrero.
Photosynthetic nitrate assimilation in cyanobacteria.
Photosynthesis Research, 83(2):117–133, 2005.

[182] Hermann Bothe, Oliver Schmitz, M. Geoffrey Yates, and William E. Newton.
Nitrogen Fixation and Hydrogen Metabolism in Cyanobacteria.
Microbiology and Molecular Biology Reviews, 74(4):529–551, 2010.

[183] Shouqin Sun, Thomas H. DeLuca, Jun Zhang, Genxu Wang, Xiangyang Sun, Zhaoyong Hu, Wenzi Wang, and Wei Zhang.
Evidence of endophytic nitrogen fixation as a potential mechanism supporting colonization of non-nodulating pioneer plants on a glacial foreland.
Biology and Fertility of Soils, 58(5):527–539, 2022.

[184] Théo Veaudor, Corinne Cassier-Chauvat, and Franck Chauvat.
Genomics of Urea Transport and Catabolism in Cyanobacteria: Biotechnological Implications.
Frontiers in Microbiology, 10:1–14, 2019.

[185] Alberto A. Esteves-Ferreira, Masami Inaba, Antoine Fort, Wagner L. Araújo, and Ronan Sulpice.
Nitrogen metabolism in cyanobacteria: metabolic and molecular control, growth consequences and biotechnological applications.
Critical Reviews in Microbiology, 44(5):541–560, 2018.

[186] Marcus Ludwig and Donald A. Bryant.
Acclimation of the global transcriptome of the cyanobacterium *Synechococcus* sp. strain PCC 7002 to nutrient limitations and different nitrogen sources.

Frontiers in Microbiology, 3:1–15, 2012.

[187] Celia R. Carlini and Rodrigo Ligabue-Braun.
Ureases as multifunctional toxic proteins: A review.
Toxicon, 110:90–109, 2016.

[188] Mark C. Walker and Wilfred A. van der Donk.
The Many Roles of Glutamate in Metabolism.
J Ind Microbiol Biotechnol, 43(0):419–430, 2016.

[189] Lauren A. Mills, Alistair J. McCormick, and David J. Lea-Smith.
Current knowledge and recent advances in understanding metabolism of the model cyanobacterium Synechocystis sp. PCC 6803.
Bioscience Reports, 40(4):1–33, 2020.

[190] Lauren L. Wingard, Scott R. Miller, Jeanne M.L. Sellker, Erik Stenn, Mary M. Allen, and A. Michelle Wood.
Cyanophycin production in a phycoerythrin-containing marine Synechococcus strain of unusual phylogenetic affinity.
Applied and Environmental Microbiology, 68(4):1772–1777, 2002.

[191] Enrique Flores, Sergio Arévalo, and Mireia Burnat.
Cyanophycin and arginine metabolism in cyanobacteria.
Algal Research, 42:101577, 2019.

[192] Emily K. Cira, Hans W. Paerla, and Michael S. Wetza.
Effects of nitrogen availability and form on phytoplankton growth in a eutrophied estuary (Neuse River Estuary, NC, USA).
PLoS ONE, 11(8):1–15, 2016.

[193] Mengyuan Shen, Qi Li, Minglei Ren, Yan Lin, Juanping Wang, Li Chen, Tao Li, and Jindong Zhao.
Trophic Status Is Associated With Community Structure and Metabolic Potential of Planktonic Microbiota in Plateau Lakes.

Frontiers in Microbiology, 10:1–15, 2019.

[194] Karl E Havens, Gaohua Ji, John R Beaver, Rolland S Fulton, and Catherine E Teacher.
Dynamics of cyanobacteria blooms are linked to the hydrology of shallow Florida lakes and provide insight into possible impacts of climate change.
Hydrobiologia, 829(1):43–59, 2019.

[195] Kemal A Ger, Vanessa Becker, and Maria Gabriela.
Abiotic factors driving cyanobacterial biomass and composition under perennial bloom conditions in tropical latitudes.
Hydrobiologia, 848(4):943–960, 2021.

[196] Yves Collos, Béatrice Bec, Cécile Jauzein, Eric Abadie, Thierry Laugier, Jacques Lautier, Annie Pastoureaud, Philippe Souchu, and André Vaquer.
Oligotrophication and emergence of picocyanobacteria and a toxic dinoflagellate in Thau lagoon, southern France.
Journal of Sea Research, 61(1-2):68–75, 2009.

[197] Patricia M. Glibert, Frances P. Wilkerson, Richard C. Dugdale, John A. Raven, Christopher L. Dupont, Peter R. Leavitt, Alexander E. Parker, Joann M. Burkholder, and Todd M. Kana.
Pluses and minuses of ammonium and nitrate uptake and assimilation by phytoplankton and implications for productivity and community composition, with emphasis on nitrogen-enriched conditions.
Limnology and Oceanography, 61(1):165–197, 2016.

[198] Gry Mine Berg, Maija Balode, Ingrida Purina, Santa Bekere, Christian Béchemin, and Serge Y. Maestrini.
Plankton community composition in relation to availability and uptake of oxidized and reduced nitrogen.
Aquatic Microbial Ecology, 30(3):263–274, 2003.

[199] Cynthia A. Heil, Marta Revilla, Patricia M. Glibert, and Sue Murasko.

Nutrient quality drives differential phytoplankton community composition on the southwest Florida shelf.
Limnology and Oceanography, 52(3):1067–1078, 2007.

[200] Derek B. Donald, Matthew J. Bogard, Kerri Finlay, and Peter R. Leavitt.
Comparative effects of urea, ammonium, and nitrate on phytoplankton abundance, community composition, and toxicity in hypereutrophic freshwaters.
Limnology and Oceanography, 56(6):2161–2175, 2011.

[201] Dmitry Shvarev and Iris Maldener.
ATP-binding cassette transporters of the multicellular cyanobacterium *Anabaena* sp. PCC 7120: A wide variety for a complex lifestyle.
FEMS Microbiology Letters, 365(4):1–6, 2018.

[202] Rungaroon Waditee, Gazi Sakir Hossain, Yoshito Tanaka, Tatsunosuke Nakamura, Masamitsu Shikata, Jun Takano, Tetsuko Takabe, and Teruhiro Takabe.
Isolation and Functional Characterization of Ca^{2+}/H^+ Antiporters from Cyanobacteria.
Journal of Biological Chemistry, 279(6):4330–4338, 2004.

[203] A. Herrero, A. M. Muro-Pastor, and E. Flores.
Nitrogen control in cyanobacteria.
Journal of Bacteriology, 183(2):411–425, 2001.

[204] Javier Espinosa, Francisco Rodríguez-Mateos, Paloma Salinas, Val F. Lanza, Ray Dixon, Fernando De La Cruz, and Asuncion Contreras.
PipX, the coactivator of NtcA, is a global regulator in cyanobacteria.
Proceedings of the National Academy of Sciences of the United States of America, 111(23):2423–2430, 2014.

[205] Dirk Steinhauser, Alisdair R. Fernie, and Wagner L. Araújo.
Unusual cyanobacterial TCA cycles: Not broken just different.
Trends in Plant Science, 17(9):503–509, 2012.

[206] M. Isabel Muro-Pastor, José C. Reyes, and Francisco J. Florencio.
Cyanobacteria Perceive Nitrogen Status by Sensing Intracellular 2-Oxoglutarate Levels.
Journal of Biological Chemistry, 276(41):38320–38328, 2001.

[207] Karl Forchhammer and Andrea Hedler.
Phosphoprotein P(II) from cyanobacteria - Analysis of functional conservation with the P(II) signal-transduction protein from *Escherichia coli*.
European Journal of Biochemistry, 244(3):869–875, 1997.

[208] Javier Espinosa, Karl Forchhammer, Sergio Burillo, and Asunción Contreras.
Interaction network in cyanobacterial nitrogen regulation: PipX, a protein that interacts in a 2-oxoglutarate dependent manner with PII and NtcA.
Molecular Microbiology, 61(2):457–469, 2006.

[209] Oleksandra Fokina, Vasuki Ranjani Chellamuthu, Karl Forchhammer, and Kornelius Zeth.
Mechanism of 2-oxoglutarate signaling by the *Synechococcus elongatus* PII signal transduction protein.
Proceedings of the National Academy of Sciences of the United States of America, 107(46):19760–19765, 2010.

[210] Alicia Forcada-Nadal, Karl Forchhammer, and Vicente Rubio.
SPR analysis of promoter binding of *Synechocystis* PCC6803 transcription factors NtcA and CRP suggests cross-talk and sheds light on regulation by effector molecules.
FEBS Letters, 588(14):2270–2276, 2014.

[211] Ryohei Tanigawa, Masao Shirokane, Shin Ichi Maeda, Tatsuo Omata, Kan Tanaka, and Hideo Takahashi.
Transcriptional activation of NtcA-dependent promoters of *Synechococcus* sp. PCC 7942 by 2-oxoglutarate in vitro.
Proceedings of the National Academy of Sciences of the United States of America, 99(7):4251–4255, 2002.

[212] Elvira Olmedo-Verd, Ana Valladares, Enrique Flores, Antonia Herrero, and Alicia M. Muro-Pastor.

Role of two NtcA-binding sites in the complex *ntcA* gene promoter of the heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120.
Journal of Bacteriology, 190(22):7584–7590, 2008.

[213] Joaquín Giner-Lamia, Rocío Robles-Rengel, Miguel A. Hernández-Prieto, M. Isabel Muro-Pastor, Francisco J. Florencio, and Matthias E. Futschik.
Identification of the direct regulon of NtcA during early acclimation to nitrogen starvation in the cyanobacterium *Synechocystis* sp. PCC 6803.
Nucleic Acids Research, 45(20):11800–11820, 2017.

[214] Kornelius Zeth, Oleksandra Fokinas, and Karl Forchhammers.
Structural basis and target-specific modulation of ADP sensing by the *Synechococcus elongatus* PII signaling protein.
Journal of Biological Chemistry, 289(13):8960–8972, 2014.

[215] I. Suzuki, N. Horie, T. Sugiyama, and T. Omata.
Identification and characterization of two nitrogen-regulated genes of the cyanobacterium *Synechococcus* sp. strain PCC7942 required for maximum efficiency of nitrogen assimilation.
Journal of Bacteriology, 177(2):290–296, 1995.

[216] Makiko Aichi and Tatsuo Omata.
Involvement of NtcB, a LysR family transcription factor, in nitrite activation of the nitrate assimilation operon in the cyanobacterium *Synechococcus* sp. strain PCC 7942.
Journal of Bacteriology, 179(15):4671–4675, 1997.

[217] Yoshitake Ohashi, Wei Shi, Nobuyuki Takatani, Makiko Aichi, Shin Ichi Maeda, Satoru Watanabe, Hirofumi Yoshikawa, and Tatsuo Omata.
Regulation of nitrate assimilation in cyanobacteria.
Journal of Experimental Botany, 62(4):1411–1424, 2011.

[218] M. Aichi, N. Takatani, and T. Omata.
Role of NtcB in activation of nitrate assimilation genes in the cyanobacterium *Synechocystis* sp. strain PCC 6803.

Journal of Bacteriology, 183(20):5840–5847, 2001.

[219] Zhengchang Su, Victor Olman, Fenglou Mao, and Ying Xu.
Comparative genomics analysis of NtcA regulons in cyanobacteria: Regulation of nitrogen assimilation and its coupling to photosynthesis.
Nucleic Acids Research, 33(16):5156–5171, 2005.

[220] Marıa Luz Montesinos, Alicia Marıa Muro-Pastor, Antonia Herrero, and Enrique Flores.
Ammonium/Methylammonium Permeases of a Cyanobacterium.
Journal of Biological Chemistry, 273(47):31463–31470, 1998.

[221] Javier Paz-Yepes, Antonia Herrero, and Enrique Flores.
The NtcA-regulated *amtB* gene is necessary for full methylammonium uptake activity in the cyanobacterium *Synechococcus elongatus*.
Journal of Bacteriology, 189(21):7791–7798, 2007.

[222] Javier Paz-Yepes, Victoria Merino-Puerto, Antonia Herrero, and Enrique Flores.
The *amt* gene cluster of the heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120.
Journal of Bacteriology, 190(19):6535–6539, 2008.

[223] María F. Vázquez-Bermüdez, Javier Paz-Yepes, Antonia Herrero, and Enrique Flores.
The NtcA-activated *amt1* gene encodes a permease required for uptake of low concentrations of ammonium in the cyanobacterium *Synechococcus* sp. PCC 7942.
Microbiology, 148(3):861–869, 2002.

[224] Raymond J. Ritchie.
The Ammonia Transport, Retention and Futile Cycling Problem in Cyanobacteria.
Microbial Ecology, 65(1):180–196, 2013.

[225] J. R. Brown, Y. Masuchi, F. T. Robb, and W. F. Doolittlel.
Evolutionary relationships of bacterial and archaeal glutamine synthetase genes.
Journal of Molecular Evolution, 38(6):566–576, 1994.

[226] J. C. Reyes, M. I. Muro-Pastor, and F. J. Florencio.

Transcription of glutamine synthetase genes (*glnA* and *glnN*) from the cyanobacterium *Synechocystis* sp. strain PCC 6803 is differently regulated in response to nitrogen availability.
Journal of Bacteriology, 179(8):2678–2689, 1997.

[227] J. C. Reyes and F. J. Florencio.
A mutant lacking the glutamine synthetase gene (*glnA*) is impaired in the regulation of the nitrate assimilation system in the cyanobacterium *Synechocystis* sp. strain PCC 6803.
Journal of Bacteriology, 176(24):7516–7523, 1994.

[228] Paul Bolay, M. Isabel Muro-Pastor, Francisco J. Florencio, and Stephan Klähn.
The distinctive regulation of cyanobacterial glutamine synthetase.
Life, 8(4):1–21, 2018.

[229] Tatsuo Omata, Xanja Andriesse, and Atsushi Hirano.
Identification and characterization of a gene cluster involved in nitrate transport in the cyanobacterium *Synechococcus* sp. PCC7942.
MGG Molecular & General Genetics, 236(2-3):193–202, 1993.

[230] Tatsuo Omata.
Structure, function and regulation of the nitrate transport system of the cyanobacterium *Synechococcus* sp. PCC7942.
Plant and Cell Physiology, 36(2):207–213, 1995.

[231] Makiko Aichi, Saori Yoshihara, Madoka Yamashita, Shin Ichi Maeda, Kazuo Nagai, and Tatsuo Omata.
Characterization of the nitrate-nitrite transporter of the major facilitator superfamily (the *nrtP* gene product) from the cyanobacterium *Nostoc punctiforme* strain ATCC 29133.
Bioscience, Biotechnology and Biochemistry, 70(11):2682–2689, 2006.

[232] Gabrielle Rocap, Frank W. Larimer, Jane Lamerdin, Stephanie Malfatti, Patrick Chain, Nathan A. Ahlgren, Andrae Arellano, Maureen Coleman, Loren Hauser, Wolfgang R. Hess, Zackary I. Johnson, Miriam Land, Debbie Lindell, Anton F. Post, Warren Regala, Manesh Shah,

Stephanie L. Shaw, Claudia Steglich, Matthew B. Sullivan, Claire S. Ting, Andrew Tolonen, Eric A. Webb, Erik R. Zinser, and Sallie W. Chisholm.
Genome divergence in two Prochlorococcus ecotypes reflects oceanic niche differentiation.
Nature, 424(6952):1042–1047, 2003.

[233] Shin-ichi Maeda, Risa Aoba, Yuma Nishino, and Tatsuo Omata.
A Novel Bacterial Nitrate Transporter Composed of Small Transmembrane Proteins.
Plant and Cell Physiology, 0:1–13, 2019.

[234] Shin Ichi Maeda and Tatsuo Omata.
Nitrite transport activity of the ABC-type cyanate transporter of the cyanobacterium *Synechococcus elongatus*.
Journal of Bacteriology, 191(10):3265–3272, 2009.

[235] Nina A. Kamennaya and Anton F. Post.
Characterization of cyanate metabolism in marine *Synechococcus* and *Prochlorococcus* spp.
Applied and Environmental Microbiology, 77(1):291–301, 2011.

[236] Alberto A. Esteves-Ferreira, João Henrique Frota Cavalcanti, Marcelo Gomes Marçal Vieira Vaz, Luna V. Alvarenga, Adriano Nunes-Nesi, and Wagner L. Araújo.
Cyanobacterial nitrogenases: Phylogenetic diversity, regulation and functional predictions.
Genetics and Molecular Biology, 40(1):261–275, 2017.

[237] Jose A. Hernandez, Robert Y. Igarashi, Basem Soboh, Leonardo Curatti, Dennis R. Dean, Paul W. Ludden, and Luis M. Rubio.
NifX and NifEN exchange NifB cofactor and the VK-cluster, a newly isolated intermediate of the iron-molybdenum cofactor biosynthetic pathway.
Molecular Microbiology, 63(1):177–192, 2007.

[238] Luis M. Rubio and Paul W. Ludden.
Biosynthesis of the iron-molybdenum cofactor of nitrogenase.
Annual Review of Microbiology, 62:93–111, 2008.

[239] Wiebke Mohr, Maria Paola Intermaggio, and Julie LaRoche.
Diel rhythm of nitrogen and carbon metabolism in the unicellular, diazotrophic cyanobacterium *Crocosphaera watsonii* WH8501.
Environmental Microbiology, 12(2):412–421, 2010.

[240] Anthony Dron, Sophie Rabouille, Pascal Claquin, Bertrand Le Roy, Amélie Talec, and Antoine Sciandra.
Light-dark (12:12) cycle of carbon and nitrogen metabolism in *Crocosphaera watsonii* WH8501: Relation to the cell cycle.
Environmental Microbiology, 14(4):967–981, 2012.

[241] David G. Adams.
Heterocyst formation in cyanobacteria.
Current Opinion in Microbiology, 3(6):618–624, 2000.

[242] Birgitta Bergman, Gustaf Sandh, Senjie Lin, John Larsson, and Edward J. Carpenter.
Trichodesmium - a widespread marine cyanobacterium with unusual nitrogen fixation properties.
FEMS Microbiology Reviews, 37(3):286–302, 2013.

[243] P. Fay.
Oxygen relations of nitrogen fixation in cyanobacteria.
Microbiological Reviews, 56(2):340–373, 1992.

[244] Noelle A. Held, John B. Waterbury, Eric A. Webb, Riss M. Kellogg, Matthew R. McIlvin, Michael Jakuba, Frederica W. Valois, Dawn M. Moran, Kevin M. Sutherland, and Mak A. Saito.
Dynamic diel proteome and daytime nitrogenase activity supports buoyancy in the cyanobacterium *Trichodesmium*.
Nature Microbiology, 7(2):300–311, 2022.

[245] Eugenio Martín-Figueroa, Francisco Navarro, and Francisco J. Florencio.
The GS-GOGAT pathway is not operative in the heterocysts. Cloning and expression of *glsF* gene from the cyanobacterium *Anabaena* sp. PCC 7120.
FEBS Letters, 476(3):282–286, 2000.

[246] Mireia Burnat, Antonia Herrero, and Enrique Flores.
Compartmentalized cyanophycin metabolism in the diazotrophic filaments of a heterocystforming cyanobacterium.
Proceedings of the National Academy of Sciences of the United States of America, 111(10):3823–3828, 2014.

[247] Ana Valladares, María Luz Montesinos, Antonia Herrero, and Enrique Flores.
An ABC-type, high-affinity urea permease identified in cyanobacteria.
Molecular Microbiology, 43(3):703–715, 2002.

[248] Irina N. Shilova, Jonathan D. Magasin, Matthew M. Mills, Julie C. Robidart, Kendra A. Turk-Kubo, and Jonathan P. Zehr.
Phytoplankton transcriptomic and physiological responses to fixed nitrogen in the California current system.
PLoS ONE, 15(4):1–24, 2020.

[249] Jackie L. Collier, Bianca Brahamsha, and Brian Palenik.
The marine cyanobacterium *Synechococcus* sp. WH7805 requires urease (urea amidohydrolase, EC 3.5.1.5) to utilize urea as a nitrogen source: Molecular-genetic and biochemical analysis of the enzyme.
Microbiology, 145(2):447–459, 1999.

[250] Jong Kyong Kim, Scott B. Mulrooney, and Robert P. Hausinger.
The UreEF fusion protein provides a soluble and functional form of the UreF urease accessory protein.
Journal of Bacteriology, 188(24):8413–8420, 2006.

[251] Marta Palombo, Alessio Bonucci, Emilien Etienne, Stefano Ciurli, Vladimir N. Uversky, Bruno Guigliarelli, Valérie Belle, Elisabetta Mileo, and Barbara Zambelli.
The relationship between folding and activity in UreG, an intrinsically disordered enzyme.
Scientific Reports, 7(1):1–10, 2017.

[252] Ronald Stebegg, Georg Schmetterer, and Annette Rompel.

Transport of organic substances through the cytoplasmic membrane of cyanobacteria.
Phytochemistry, 157(August 2018):206–218, 2019.

[253] María Luz Montesinos, Antonia Herrero, and Enrique Flores.
Amino acid transport in taxonomically diverse cyanobacteria and identification of two genes encoding elements of a neutral amino acid permease putatively involved in recapture of leaked hydrophobic amino acids.
Journal of Bacteriology, 179(3):853–862, 1997.

[254] Rafael Pernil, Silvia Picossi, Vicente Mariscal, Antonia Herrero, and Enrique Flores.
ABC-type amino acid uptake transporters Bgt and N-II of *Anabaena* sp. strain PCC 7120 share an ATPase subunit and are expressed in vegetative cells and heterocysts.
Molecular Microbiology, 67(5):1067–1080, 2008.

[255] Rafael Pernil, Silvia Picossi, Antonia Herrero, Enrique Flores, and Vicente Mariscal.
Amino acid transporters and release of hydrophobic amino acids in the heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120.
Life, 5(2):1282–1300, 2015.

[256] M. L. Montesinos, A. Herrero, and E. Flores.
Amino acid transport systems required for diazotrophic growth in the cyanobacterium *Anabaena* sp. strain PCC 7120.
Journal of Bacteriology, 177(11):3150–3157, 1995.

[257] Silvia Picossi, María Luz Montesinos, Rafael Pernil, Christiane Lichtlé, Antonia Herrero, and Enrique Flores.
ABC-type neutral amino acid permease N-I is required for optimal diazotrophic growth and is repressed in the heterocysts of *Anabaena* sp. strain PCC 7120.
Molecular Microbiology, 57(6):1582–1592, 2005.

[258] Leticia Escudero, Vicente Mariscal, and Enrique Flores.
Functional dependence between septal protein SepJ from *Anabaena* sp. Strain PCC 7120 and an Amino Acid ABC-Type Uptake Transporter.

Journal of Bacteriology, 197(16):2721–2730, 2015.

[259] M. J. Quintero, M. L. Montesinos, A. Herrero, and E. Flores.
Identification of genes encoding amino acid permeases by inactivation of selected ORFs from the *Synechocystis* genomic sequence.
Genome Research, 11(12):2034–2040, 2001.

[260] Aporn Bualuang, Hakuto Kageyama, Yoshito Tanaka, Aran Incharoensakdi, and Teruhiro Takabe.
Functional characterization of a member of alanine or glycine: Cation symporter family in halotolerant cyanobacterium Aphanothece halophytica.
Bioscience, Biotechnology and Biochemistry, 79(2):230–235, 2015.

[261] Tsuneo Fujishiro, Takakazu Kaneko, Masahiro Sugiura, and Mamoru Sugita.
Organization and transcription of a putative gene cluster encoding ribosomal protein S14 and an oligopeptide permease-like protein in the cyanobacterium *Synechococcus* sp. strain PCC 6301.
DNA Research, 3(3):165–169, 1996.

[262] Daniel Pletzer, Corinne Lafon, Yvonne Braun, Thilo Köhler, Malcolm G.P. Page, Michael Mourez, and Helge Weingart.
High-throughput screening of dipeptide utilization mediated by the ABC transporter DppBCDF and its substrate-binding proteins DppA1-A5 in *Pseudomonas aeruginosa*.
PLoS ONE, 9(10), 2014.

[263] Francisco Navarro, Eugenio Martín-Figueroa, Pedro Candau, and Francisco J. Florencio.
Ferredoxin-dependent iron-sulfur flavoprotein glutamate synthase (GlsF) from the cyanobacterium *Synechocystis* sp. PCC 6803: Expression and assembly in *Escherichia coli*.
Archives of Biochemistry and Biophysics, 379(2):267–276, 2000.

[264] Sebastián Chávez, José Carlos Reyes, Franck Chauvat, Francisco J. Florencio, and Pedro Candau.
The NADP-glutamate dehydrogenase of the cyanobacterium *Synechocystis* 6803: cloning, transcriptional analysis and disruption of the *gdhA* gene.
Plant Molecular Biology, 28(1):173–188, 1995.

[265] Li-Kai Liu, Donald F. Becker, and John Tanner.
Structure, function, and mechanism of proline utilization A.
Arch Biochem Biophys, 632:142–157, 2017.

[266] Sarah Schriek, Christian Rückert, Dorothee Staiger, Elfriede K. Pistorius, and Klaus Peter Michel.
Bioinformatic evaluation of L-arginine catabolic pathways in 24 cyanobacteria and transcriptional analysis of genes encoding enzymes of L-arginine catabolism in the cyanobacterium *Synechocystis* sp. PCC 6803.
BMC Genomics, 8:1–28, 2007.

[267] Natalia Correa-Aragunde, Noelia Foresi, Fiorella Del Castello, and Lorenzo Lamattina.
A singular nitric oxide synthase with a globin domain found in *Synechococcus* PCC 7335 mobilizes N from arginine to nitrate.
Scientific Reports, 8(1):1–11, 2018.

[268] Xinyi Xu, Liping Gu, Ping He, and Ruanbao Zhou.
Characterization of five putative aspartate aminotransferase genes in the N2-fixing heterocystous cyanobacterium *Anabaena* sp. Strain PCC 7120.
Microbiology (United Kingdom), 161(6):1219–1230, 2015.

[269] G. Van Heeke and S. M. Schuster.
Expression of human asparagine synthetase in *Escherichia coli.*
Journal of Biological Chemistry, 264(10):5503–5509, 1989.

[270] Xunyan Dong, Peter J. Quinn, and Xiaoyuan Wang.
Metabolic engineering of *Escherichia coli* and *Corynebacterium glutamicum* for the production of l-threonine.
Biotechnology Advances, 29(1):11–23, 2011.

[271] André O. Hudson, Charles Gilvarg, and Thomas Leustek.
Biochemical and phylogenetic characterization of a novel diaminopimelate biosynthesis pathway in prokaryotes identifies a diverged form of LL-diaminopimelate aminotransferase.
Journal of Bacteriology, 190(9):3256–3263, 2008.

[272] Gary Sawers.
The anaerobic degradation of L-serine and L-threonine in enterobacteria: Networks of pathways and regulatory signals.
Archives of Microbiology, 171(1):1–5, 1998.

[273] Guilherme Coutinho De Mello Serrano, Thaís Rezende E Silva Figueira, Eduardo Kiyota, Natalia Zanata, and Paulo Arruda.
Lysine degradation through the saccharopine pathway in bacteria: LKR and SDH in bacteria and its relationship to the plant and animal enzymes.
FEBS Letters, 586(6):905–911, 2012.

[274] Sebastian Knorr, Malte Sinn, Dmitry Galetskiy, Rhys M. Williams, Changhao Wang, Nicolai Müller, Olga Mayans, David Schleheck, and Jörg S. Hartig.
Widespread bacterial lysine degradation proceeding via glutarate and L-2-hydroxyglutarate.
Nature Communications, 9(1):2–11, 2018.

[275] Wonjae Kim, Minkyung Kim, and Woojun Park.
Unlocking the mystery of lysine toxicity on *Microcystis aeruginosa*.
Journal of Hazardous Materials, 448:130932, 2023.

[276] Matteo P. Ferla and Wayne M. Patrick.
Bacterial methionine biosynthesis.
Microbiology (United Kingdom), 160(PART 8):1571–1584, 2014.

[277] M. Pettine, L. Patrolecco, M. Manganelli, S. Capri, and M. G. Farrace.
Seasonal variations of dissolved organic matter in the northern Adriatic Sea.
Marine Chemistry, 64(3):153–169, 1999.

[278] Chong Xiao Ji, Yan Chen, and Gui Peng Yang.
Seasonal Variation, Degradation, and Bioavailability of Dissolved Organic Matter in the Changjiang Estuary and its Adjacent East China Sea.
Journal of Geophysical Research: Oceans, 126(3):1–20, 2021.

[279] L. K. Massey, J. R. Sokatch, and R. S. Conrad.
Branched chain amino acid catabolism in bacteria.
Bacteriological Reviews, 40(1):42–54, 1976.

[280] Irene Ylla, Isis Sanpera-Calbet, Isabel Muñoz, Anna M. Romaní, and Sergi Sabater.
Organic matter characteristics in a Mediterranean stream through amino acid composition: Changes driven by intermittency.
Aquatic Sciences, 73(4):523–535, 2011.

[281] John A. O'Connor, Kaijun Lu, Laodong Guo, Brad E. Rosenheim, and Zhanfei Liu.
Composition and lability of riverine dissolved organic matter: Insights from thermal slicing ramped pyrolysis GC–MS, amino acid, and stable isotope analyses.
Organic Geochemistry, 149:104100, 2020.

[282] Hiroshi Maeda and Natalia Dudareva.
The shikimate pathway and aromatic amino acid biosynthesis in plants.
Annual Review of Plant Biology, 63:73–105, 2012.

[283] Pierre Legrand, Renaud Dumas, Marlene Seux, Pascal Rippert, Raimond Ravelli, Jean Luc Ferrer, and Michel Matringe.
Biochemical Characterization and Crystal Structure of *Synechocystis* Arogenate Dehydrogenase Provide Insights into Catalytic Reaction.
Structure, 14(4):767–776, 2006.

[284] Keri L. Colabroy and Tadhg P. Begley.
Tryptophan catabolism: Identification and characterization of a new degradative pathway.
Journal of Bacteriology, 187(22):7866–7869, 2005.

[285] Michelle C. Moffitt, Gordon V. Louie, Marianne E. Bowman, Janelle Pence, Joseph P. Noel, and Bradley S. Moore.
Discovery of Two Cyanobacterial Phenylalanine Ammonia Lyases: Kinetic and Structural Characterization.
Biochemistry, 46(4):1004–1012, 2007.

[286] Margaret V. McCaul, David Sutton, André J. Simpson, Adrian Spence, David J. McNally, Brian W. Moran, Alok Goel, Brendan O'Connor, Kris Hart, and Brian P. Kelleher.
Composition of dissolved organic matter within a lacustrine environment.
Environmental Chemistry, 8(2):146–154, 2011.

[287] M. Hagemann, J. Vinnemeier, I. Oberpichler, R. Boldt, and H. Bauwe.
The glycine decarboxylase complex is not essential for the cyanobacterium *Synechocystis* sp. strain PCC 6803.
Plant Biology, 7(1):15–22, 2005.

[288] Friederike Klemke, Antje Baier, Henning Knoop, Ramona Kern, Jiri Jablonsky, Gabriele Beyer, Thomas Volkmer, Ralf Steuer, Wolfgang Lockau, and Martin Hagemann.
Identification of the light-independent phosphoserine pathway as an additional source of serine in the cyanobacterium *Synechocystis* sp. PCC 6803.
Microbiology, 161:1050–1060, 2015.

[289] Yan Chen, Gui Peng Yang, Guan Wei Wu, Xian Chi Gao, and Qing Yan Xia.
Concentration and characterization of dissolved organic matter in the surface microlayer and subsurface water of the Bohai Sea, China.
Continental Shelf Research, 52:97–107, 2013.

[290] Karl Kaiser, Maria Canedo-Oropeza, Rachel McMahon, and Rainer M.W. Amon.
Origins and transformations of dissolved organic matter in large Arctic rivers.
Scientific Reports, 7(1):1–11, 2017.

[291] Yi Kai Liu, Hung Chih Kuo, Chih Ho Lai, and Chi Chung Chou.
Single amino acid utilization for bacterial categorization.
Scientific Reports, 10(1):1–12, 2020.

[292] Mary Louise Nicholson, Michelle Gaasenbeek, and David E. Laudenbach.
Two enzymes together capable of cysteine biosynthesis are encoded on a cyanobacterial plasmid.
MGG Molecular & General Genetics, 247(5):623–632, 1995.

[293] Bhramara Tirupati, Jessica Lynn Vey, Catherine L. Drennan, and J. Martin Bollinger.
Kinetic and structural characterization of Slr0077/SufS, the essential cysteine desulfurase from *Synechocystis* sp. PCC 6803.
Biochemistry, 43(38):12210–12219, 2004.

[294] Rafael Pernil, Antonia Herrero, and Enrique Flores.
Catabolic function of compartmentalized alanine dehydrogenase in the heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120.
Journal of Bacteriology, 192(19):5165–5172, 2010.

[295] Robert A. Bender.
Regulation of the Histidine Utilization (Hut) System in Bacteria.
Microbiology and Molecular Biology Reviews, 76(3):565–584, 2012.

[296] A Steitz and B Velimirov.
Contribution of Picocyanobacteria to total primary production and community respiratory losses in a backwater system.
Journal of Plankton Research, 21(12):2341–2360, 1999.

[297] Lidia Romero-Viana, Brendan J. Keely, Antonio Camacho, Eduardo Vicente, and M. Rosa Miracle.
Primary production in Lake La Cruz (Spain) over the last four centuries: Reconstruction based on sedimentary signal of photosynthetic pigments.
Journal of Paleolimnology, 43(4):771–786, 2010.

[298] Małgorzata Waleron, Krzysztof Waleron, Warwick F. Vincent, and Annick Wilmotte.
Allochthonous inputs of riverine picocyanobacteria to coastal waters in the Arctic Ocean.
FEMS Microbiology Ecology, 59(2):356–365, 2007.

[299] Pedro Flombaum, José L. Gallegos, Rodolfo A. Gordillo, José Rincón, Lina L. Zabala, Nianzhi Jiao, David M. Karl, William K.W. Li, Michael W. Lomas, Daniele Veneziano, Carolina S. Vera, Jasper A. Vrugt, and Adam C. Martiny.
Present and future global distributions of the marine Cyanobacteria *Prochlorococcus* and *Synechococcus*.

Proceedings of the National Academy of Sciences of the United States of America, 110(24):9824–9829, 2013.

[300] Sijun Huang, Yongqin Liu, Anyi Hu, Xiaobo Liu, Feng Chen, Tandong Yao, and Nianzhi Jiao.
Genetic diversity of picocyanobacteria in Tibetan lakes: Assessing the endemic and universal distributions.
Applied and Environmental Microbiology, 80(24):7640–7650, 2014.

[301] Pedro J. Cabello-Yeves, Antonio Picazo, Antonio Camacho, Cristiana Callieri, Riccardo Rosselli, Juan J. Roda-Garcia, Felipe H. Coutinho, and Francisco Rodriguez-Valera.
Ecological and genomic features of two widespread freshwater picocyanobacteria.
Environmental Microbiology, 20(10):3757–3771, 2018.

[302] S. V. Shestakov and Nguyen Than Khyen.
Evidence for genetic transformation in blue-green alga *Anacystis nidulans*.
MGG Molecular & General Genetics, 107(4):372–375, 1970.

[303] Takakazu Kaneko, Tohru Matsubayashi, Mamoru Sugita, and Masahiro Sugiura.
Physical and gene maps of the unicellular cyanobacterium *Synechococcus* sp. strain PCC6301 genome.
Plant Molecular Biology, 31(1):193–201, 1996.

[304] A. N. Glazer, D. J. Lundell, G. Yamanaka, and R. C. Williams.
The structure of a simple phycobilisome.
Annales de l'Institut Pasteur Microbiology, 134(1):159–180, 1983.

[305] David J. Bonfil, Michal Ronen-Tarazi, Dieter Sültemeyer, Judy Lieman-Hurwitz, Daniella Schatz, and Aaron Kaplan.
A putative HCO_3^- transporter in the cyanobacterium *Synechococcus* sp. strain PCC 7942.
FEBS Letters, 430(3):236–240, 1998.

[306] Lorraine G. Van Waasbergen, Nadia Dolganov, and Arthur R. Grossman.
nblS, a gene involved in controlling photosynthesis-related gene expression during high light and nutrient stress in *Synechococcus elongatus* PCC 7942.

Journal of Bacteriology, 184(9):2481–2490, 2002.

[307] Kathryn A. Durham, David Porta, R. Michael L. McKay, and George S. Bullerjahn.
Expression of the iron-responsive *irpA* gene from the cyanobacterium *Synechococcus* sp. strain PCC 7942.
Archives of Microbiology, 179(2):131–134, 2003.

[308] Jared M. Fraser, Sarah E. Tulk, Jennifer A. Jeans, Douglas A. Campbell, Thomas S. Bibby, and Amanda M. Cockshutt.
Photophysiological and Photosynthetic Complex Changes during Iron Starvation in *Synechocystis* sp. PCC 6803 and *Synechococcus elongatus* PCC 7942.
PLoS ONE, 8(3), 2013.

[309] Jason W. Hickman, Kimberly M. Kotovic, Cameron Miller, Paul Warrener, Brett Kaiser, Tracey Jurista, Mark Budde, Fred Cross, James M. Roberts, and Michael Carleton.
Glycogen synthesis is a required component of the nitrogen stress response in *Synechococcus elongatus* PCC 7942.
Algal Research, 2(2):98–106, 2013.

[310] P. V. Sane, Alexander G. Ivanov, Dmitry Sveshnikov, Norman P.A. Huner, and Gunnar Öquist.
A transient exchange of the photosystem II reaction center protein D1:1 with D1:2 during low temperature stress of *Synechococcus* sp. PCC 7942 in the light lowers the redox potential of QB.
Journal of Biological Chemistry, 277(36):32739–32745, 2002.

[311] M. R. Schaefer and S. S. Golden.
Light availability influences the ratio of two forms of D1 in cyanobacterial thylakoids.
The Journal of biological chemistry, 264(13):7412–7417, 1989.

[312] Motarab Hossain and Hitoshi Nakamoto.
HtpG plays a role in cold acclimation in cyanobacteria.
Current Microbiology, 44(4):291–296, 2002.

[313] Rachel D. Hood, Sean A. Higgins, Avi Flamholz, Robert J. Nichols, and David F. Savage.
The stringent response regulates adaptation to darkness in the cyanobacterium *Synechococcus elongatus*.
Proceedings of the National Academy of Sciences of the United States of America, 113(33):E4867–E4876, 2016.

[314] Damini Jaiswal, Annesha Sengupta, Shinjinee Sengupta, Swati Madhu, Himadri B. Pakrasi, and Pramod P. Wangikar.
A Novel Cyanobacterium *Synechococcus elongatus* PCC 11802 has Distinct Genomic and Metabolomic Characteristics Compared to its Neighbor PCC 11801.
Scientific Reports, 10(1):1–15, 2020.

[315] O. A. Gorelova, O. I. Baulina, U. Rasmussen, and O. A. Koksharova.
The pleiotropic effects of *ftn2* and *ftn6* mutations in cyanobacterium *Synechococcus* sp. PCC 7942: An ultrastructural study.
Protoplasma, 250(4):931–942, 2013.

[316] B. R. Robertson, N. Tezuka, and M. M. Watanabe.
Phylogenetic analyses of *Synechococcus* strains (cyanobacteria) using sequences of 16S rDNA and part of the phycocyanin operon reveal multiple evolutionary lines and reflect phycobilin content.
International Journal of Systematic and Evolutionary Microbiology, 51(3):861–871, 2001.

[317] Bettina E. Schirrmeister, Muriel Gugger, and Philip C.J. Donoghue.
Cyanobacteria and the Great Oxidation Event: Evidence from genes and fossils.
Palaeontology, 58(5):769–785, 2015.

[318] Robert E. Blankenship.
Early Evolution of Photosynthesis.
Plant Physiology, 154(2):434–438, 2010.

[319] Patricia Sánchez-Baracaldo and Tanai Cardona.
On the origin of oxygenic photosynthesis and Cyanobacteria.

New Phytologist, 225(4):1440–1446, 2020.

[320] Sun Young Choi, Byeonghyeok Park, In Geol Choi, Sang Jun Sim, Sun Mi Lee, Youngsoon Um, and Han Min Woo.
Transcriptome landscape of *Synechococcus elongatus* PCC 7942 for nitrogen starvation responses using RNA-seq.
Scientific Reports, 6:1–10, 2016.

[321] Brian Palenik, Qinghu Ren, Chris L. Dupont, Garry S. Myers, John F. Heidelberg, Jonathan H. Badger, Ramana Madupu, William C. Nelson, Lauren M. Brinkac, Robert J. Dodson, A. Scott Durkin, Sean C. Daugherty, Stephen A. Sullivan, Hoda Khouri, Yasmin Mohamoud, Rebecca Halpin, and Ian T. Paulsen.
Genome sequence of *Synechococcus* CC9311: Insights into adaptation to a coastal environment.
Proceedings of the National Academy of Sciences of the United States of America, 103(36):13555–13559, 2006.

[322] Alexis P. Yelton, Silvia G. Acinas, Shinichi Sunagawa, Peer Bork, Carlos Pedrós-Alió, and Sallie W. Chisholm.
Global genetic capacity for mixotrophy in marine picocyanobacteria.
ISME Journal, 10(12):2946–2957, 2016.

[323] Hugo Doré, Gregory K. Farrant, Ulysse Guyet, Julie Haguait, Florian Humily, Morgane Ratin, Frances D. Pitt, Martin Ostrowski, Christophe Six, Loraine Brillet-Guéguen, Mark Hoebeke, Antoine Bisch, Gildas Le Corguillé, Erwan Corre, Karine Labadie, Jean Marc Aury, Patrick Wincker, Dong Han Choi, Jae Hoon Noh, Damien Eveillard, David J. Scanlan, Frédéric Partensky, and Laurence Garczarek.
Evolutionary Mechanisms of Long-Term Genome Diversification Associated With Niche Partitioning in Marine Picocyanobacteria.
Frontiers in Microbiology, 11:1–23, 2020.

[324] Christophe Six, Jean Claude Thomas, Laurence Garczarek, Martin Ostrowski, Alexis Dufresne, Nicolas Blot, David J. Scanlan, and Frédéric Partensky.

Diversity and evolution of phycobilisomes in marine *Synechococcus* spp.: A comparative genomics study.
Genome Biology, 8(12), 2007.

[325] Liguang Li, Zhenli He, Zhigang Li, Songhe Zhang, Suli Li, Yongshan Wan, and Peter J. Stoffella.
Spatial and temporal variation of nitrogen concentration and speciation in runoff and storm water in the Indian River watershed, South Florida.
Environmental Science and Pollution Research, 23(19):19561–19569, 2016.

[326] Miriam Land, Loren Hauser, Se Ran Jun, Intawat Nookaew, Michael R. Leuze, Tae Hyuk Ahn, Tatiana Karpinets, Ole Lund, Guruprased Kora, Trudy Wassenaar, Suresh Poudel, and David W. Ussery.
Insights from 20 years of bacterial genome sequencing.
Functional and Integrative Genomics, 15(2):141–161, 2015.

[327] Danillo O. Alvarenga, Marli F. Fiore, and Alessandro M. Varani.
A metagenomic approach to cyanobacterial genomics.
Frontiers in Microbiology, 8:1–16, 2017.

[328] Chau Hai Thai Vu, Hyung Gwan Lee, Yong Keun Chang, and Hee Mock Oh.
Axenic cultures for microalgal biotechnology: Establishment, assessment, maintenance, and applications.
Biotechnology Advances, 36(2):380–396, 2018.

[329] Nathan A.M. Chrismas, Gary Barker, Alexandre M. Anesio, and Patricia Sánchez-Baracaldo.
Genomic mechanisms for cold tolerance and production of exopolysaccharides in the Arctic cyanobacterium *Phormidesmis priestleyi* BC1401.
BMC Genomics, 17(1):1–14, 2016.

[330] Anton Bankevich, Sergey Nurk, Dmitry Antipov, Alexey A. Gurevich, Mikhail Dvorkin, Alexander S. Kulikov, Valery M. Lesin, Sergey I. Nikolenko, Son Pham, Andrey D. Prjibelski, Alexey V. Pyshkin, Alexander V. Sirotkin, Nikolay Vyahhi, Glenn Tesler, Max A. Alekseyev, and Pavel A. Pevzner.

SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing.
Journal of Computational Biology, 19(5):455–477, 2012.

[331] Phillip E.C. Compeau, Pavel A. Pevzner, and Glenn Tesler.
How to apply de Bruijn graphs to genome assembly.
Nature Biotechnology, 29(11):987–991, 2011.

[332] R. Y. Stanier, R. Kunisawa, M. Mandel, and G. Cohen-Bazire.
Purification and properties of unicellular blue-green algae (order Chroococcales).
Bacteriological reviews, 35(2):171–205, 1971.

[333] Anthony M. Bolger, Marc Lohse, and Bjoern Usadel.
Trimmomatic: A flexible trimmer for Illumina sequence data.
Bioinformatics, 30(15):2114–2120, 2014.

[334] Stephen F. Altschul, Warren Gish, Webb Miller, Eugene W. Myers, and David J. Lipman.
Basic local alignment search tool.
Journal of Molecular Biology, 215(3):403–410, 1990.

[335] Armen Y. Mulkidjanian, Eugene V. Koonin, Kira S. Makarova, Sergey L. Mekhedov, Alexander Sorokin, Yuri I. Wolf, Alexis Dufresne, Frédéric Partensky, Henry Burd, Denis Kaznadzey, Robert Haselkorn, and Michael Y. Galperin.
The cyanobacterial genome core and the origin of photosynthesis.
Proceedings of the National Academy of Sciences of the United States of America, 103(35):13126–13131, 2006.

[336] Ryan R. Wick, Mark B. Schultz, Justin Zobel, and Kathryn E. Holt.
Bandage: Interactive visualization of de novo genome assemblies.
Bioinformatics, 31(20):3350–3352, 2015.

[337] John Besemer and Mark Borodovsky.
GeneMark: Web software for gene finding in prokaryotes, eukaryotes and viruses.
Nucleic Acids Research, 33(SUPPL. 2):451–454, 2005.

[338] Doug Hyatt, Gwo-Liang Chen, Philip F LoCascio, Miriam L Land, Frank W Larimer, and Loren J Hauser.
Prodigal: prokaryotic gene recognition and translation initiation site identification.
BMC Bioinformatics, 11(119):1–8, 2010.

[339] Eric P. Nawrocki and Sean R. Eddy.
Infernal 1.1: 100-fold faster RNA homology searches.
Bioinformatics, 29(22):2933–2935, 2013.

[340] Patricia P. Chan and Todd M. Lowe.
tRNAscan-SE: Searching for tRNA genes in genomic sequences.
Methods in Molecular Biology, 1962(5):1–14, 2019.

[341] Felipe A. Simão, Robert M. Waterhouse, Panagiotis Ioannidis, Evgenia V. Kriventseva, and Evgeny M. Zdobnov.
BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs.
Bioinformatics, 31(19):3210–3212, 2015.

[342] I. Min A. Chen, Ken Chu, Krishna Palaniappan, Manoj Pillay, Anna Ratner, Jinghua Huang, Marcel Huntemann, Neha Varghese, James R. White, Rekha Seshadri, Tatyana Smirnova, Edward Kirton, Sean P. Jungbluth, Tanja Woyke, Emiley A. Eloe-Fadrosh, Natalia N. Ivanova, and Nikos C. Kyrpides.
IMG/M v.5.0: An integrated data management and comparative analysis system for microbial genomes and microbiomes.
Nucleic Acids Research, 47(D1):D666–D677, 2019.

[343] Jaime Huerta-Cepas, Kristoffer Forslund, Luis Pedro Coelho, Damian Szklarczyk, Lars Juhl Jensen, Christian Von Mering, and Peer Bork.
Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper.
Molecular Biology and Evolution, 34(8):2115–2122, 2017.

[344] Minoru Kanehisa and Susumu Goto.
KEGG: Kyoto Encyclopedia of Genes and Genomes.

Nucleic Acids Research, 28(1):27–30, 2000.

[345] Christiam Camacho, George Coulouris, Vahram Avagyan, Ning Ma, Jason Papadopoulos, Kevin Bealer, and Thomas L. Madden.
BLAST+: Architecture and applications.
BMC Bioinformatics, 10:1–9, 2009.

[346] Kazutaka Katoh and Daron M Standley.
MAFFT multiple sequence alignment software version 7: improvements in performance and usability.
Molecular biology and evolution, 30(4):772–80, 2013.

[347] Bui Quang Minh, Heiko A. Schmidt, Olga Chernomor, Dominik Schrempf, Michael D. Woodhams, Arndt Von Haeseler, Robert Lanfear, and Emma Teeling.
IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era.
Molecular Biology and Evolution, 37(5):1530–1534, 2020.

[348] Subha Kalyaanamoorthy, Bui Quang Minh, Thomas K.F. Wong, Arndt Von Haeseler, and Lars S. Jermiin.
ModelFinder: Fast model selection for accurate phylogenetic estimates.
Nature Methods, 14(6):587–589, 2017.

[349] Olga Chernomor, Arndt Von Haeseler, and Bui Quang Minh.
Terrace Aware Data Structure for Phylogenomic Inference from Supermatrices.
Systematic Biology, 65(6):997–1008, 2016.

[350] Diep Thi Hoang, Olga Chernomor, Arndt von Haeseler, Bui Quang Minh, and Le Sy Vinh.
UFBoot2: Improving the Ultrafast Bootstrap Approximation. Molecular biology and evolution.
Molecular Biology and Evolution, 35(2):518–522, 2017.

[351] Stephen J. Giovannoni, J. Cameron Thrash, and Ben Temperton.
Implications of streamlining theory for microbial ecology.
ISME Journal, 8(8):1553–1565, 2014.

[352] Scott Mann and Yi Ping Phoebe Chen.
Bacterial genomic G + C composition-eliciting environmental adaptation.
Genomics, 95(1):7–15, 2010.

[353] Jake L. Weissman, William F. Fagan, and Philip L.F. Johnson.
Linking high GC content to the repair of double strand breaks in prokaryotic genomes.
PLoS Genetics, 15(11):1–19, 2019.

[354] Paul M Berube, Anna Rasmussen, Rogier Braakman, Ramunas Stepanauskas, and Sallie W Chisholm.
Emergence of trait variability through the lens of nitrogen assimilation in *Prochlorococcus*.
eLife, 8:1–28, 2019.

[355] Shakti N. Menon, P. Varuni, Freddy Bunbury, Devaki Bhaya, and Gautam I. Menon.
Phototaxis in Cyanobacteria: From Mutants to Models of Collective Behavior.
mBio, 12(6), 2021.

[356] J. McCarren and B. Brahamsha.
Swimming motility mutants of marine *Synechococcus* affected in production and localization of the S-layer protein SwmA.
Journal of Bacteriology, 191(3):1111–1114, 2009.

[357] Cristiana Callieri, Pedro J. Cabello-Yeves, and Filippo Bertoni.
The "Dark Side" of Picocyanobacteria: Life as We Do Not Know It (Yet).
Microorganisms, 10(3):1–18, 2022.

[358] Yiling Yang, Vinson Lam, Marie Adomako, Ryan Simkovsky, Annik Jakob, Nathan C. Rockwell, Susan E. Cohen, Arnaud Taton, Jingtong Wang, J. Clark Lagarias, Annegret Wilde, David R. Nobles, Jerry J. Brand, and Susan S. Golden.
Phototaxis in a wild isolate of the cyanobacterium *Synechococcus elongatus*.
Proceedings of the National Academy of Sciences of the United States of America, 115(52):E12378–E12387, 2018.

[359] Éva Kiss, Péter B. Kós, Min Chen, and Imre Vass.
A unique regulation of the expression of the *psbA*, *psbD*, and *psbE* genes, encoding the D1, D2 and cytochrome b559 subunits of the Photosystem II complex in the chlorophyll d containing cyanobacterium *Acaryochloris marina*.
Biochimica et Biophysica Acta - Bioenergetics, 1817(7):1083–1094, 2012.

[360] Cécile Breyton, Christophe Tribet, Jacqueline Olive, Jean Paul Dubacq, and Jean Luc Popott.
Dimer to monomer conversion of the cytochrome b6f complex: Causes and consequences.
Journal of Biological Chemistry, 272(35):21892–21900, 1997.

[361] Dirk Schneider, Thomas Volkmer, and Matthias Rögner.
PetG and PetN, but not PetL, are essential subunits of the cytochrome b6f complex from *Synechocystis* PCC 6803.
Research in Microbiology, 158(1):45–50, 2007.

[362] Raquel García-Cañas, Joaquín Giner-Lamia, Francisco J. Florencio, and Luis López-Maury.
A protease-mediated mechanism regulates the cytochrome c6/plastocyanin switch in *Synechocystis* sp. PCC 6803.
Proceedings of the National Academy of Sciences of the United States of America, 118(5), 2021.

[363] Alejandro Torrado, Carmen Ramírez-Moncayo, José A. Navarro, Vicente Mariscal, and Fernando P. Molina-Heredia.
Cytochrome c6 is the main respiratory and photosynthetic soluble electron donor in heterocysts of the cyanobacterium *Anabaena* sp. PCC 7120.
Biochimica et Biophysica Acta - Bioenergetics, 1860(1):60–68, 2019.

[364] Martina Bečková, Zdenko Gardian, Jianfeng Yu, Peter Konik, Peter J. Nixon, and Josef Komenda.
Association of Psb28 and Psb27 Proteins with PSII-PSI Supercomplexes upon Exposure of *Synechocystis* sp. PCC 6803 to High Light.
Molecular Plant, 10(1):62–72, 2017.

[365] Pablo I. Calzadilla, Fernando Muzzopappa, Pierre Sétif, and Diana Kirilovsky.

Different roles for ApcD and ApcF in *Synechococcus elongatus* and *Synechocystis* sp. PCC 6803 phycobilisomes.

Biochimica et Biophysica Acta - Bioenergetics, 1860(6):488–498, 2019.

[366] Qian Zhao Xu, Jia Xin Han, Qi Ying Tang, Wen Long Ding, Dan Miao, Ming Zhou, Hugo Scheer, and Kai Hong Zhao.

Far-red light photoacclimation: Chromophorylation of FR induced α- and β-subunits of allophycocyanin from *Chroococcidiopsis thermalis* sp. PCC7203.

Biochimica et Biophysica Acta - Bioenergetics, 1857(9):1607–1616, 2016.

[367] Fei Gan and Donald A. Bryant.

Adaptive and acclimative responses of cyanobacteria to far-red light.

Environmental Microbiology, 17(10):3450–3465, 2015.

[368] Robert de Lorimier, Donald A. Bryant, and S. Edward Stevens.

Genetic analysis of a 9 kDa phycocyanin-associated linker polypeptide.

BBA - Bioenergetics, 1019(1):29–41, 1990.

[369] Kumiko Kondo, Xing Geng Xiao, Mitsunori Katayama, and Masahiko Ikeuchi.

Distinct roles of CpcG1 and CpcG2 in phycobilisome assembly in the cyanobacterium *Synechocystis* sp. PCC 6803.

Photosynthesis Research, 84(1-3):269–273, 2005.

[370] Jessica Wiethaus, Andrea W.U. Busch, Klaus Kock, Lars I. Leichert, Christian Herrmann, and Nicole Frankenberg-Dinkel.

CpeS is a lyase specific for attachment of 3Z-PEB to Cys82 of β-phycoerythrin from *Prochlorococcus marinus* MED4.

Journal of Biological Chemistry, 285(48):37561–37569, 2010.

[371] Ryan P. Bezy, Lisa Wiltbank, and David M. Kehoe.

Light-dependent attenuation of phycoerythrin gene expression reveals convergent evolution of green light sensing in cyanobacteria.

Proceedings of the National Academy of Sciences of the United States of America, 108(45):18542–18547, 2011.

[372] R. F. Follett and J. L. Hatfield.
Nitrogen in the environment: sources, problems, and management.
TheScientificWorldJournal, 1 Suppl 2:920–926, 2001.

[373] José Enrique Frías and Enrique Flores.
Negative regulation of expression of the nitrate assimilation *nirA* operon in the heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120.
Journal of Bacteriology, 192(11):2769–2778, 2010.

[374] E. Flores and A. Herrero.
Nitrogen assimilation and nitrogen control in cyanobacteria.
Biochemistry Society Transactions, 33(1):164–167, 2005.

[375] Taishan Hu, Nilesh Chitnis, Dimitri Monos, and Anh Dinh.
Next-generation sequencing technologies: An overview.
Human Immunology, 82(11):801–811, 2021.

[376] Eduard Fadeev, Fabio De Pascale, Alessandro Vezzi, Sariel Hübner, Dikla Aharonovich, and Daniel Sher.
Why close a bacterial genome? The plasmid of *Alteromonas macleodii* HOT1A3 is a vector for inter-specific transfer of a flexible genomic Island.
Frontiers in Microbiology, 7(MAR):1–13, 2016.

[377] Adam Thrash, Federico Hoffmann, and Andy Perkins.
Toward a more holistic method of genome assembly assessment.
BMC Bioinformatics, 21(Suppl 4):1–8, 2020.

[378] Bradnam K.R., Fass J.N., Alexandrov A., Baranay P., Bechner M., Birol I., Boisvert S., Chapman J.A., Chapuis G., Chikhi R., Chitsaz H., Chou W.-C., Corbeil J., Fabbro C.D., Docking T.R., Durbin R., Earl D., Emrich S., Fedotov P., Fonseca N.A., Ganapathy G., Gibbs R.A., Gnerre S.,

Godzaridis E., Goldstein S., Haimel M., Hall G., Haussler D., Hiatt J.B., Ho I.Y., Howard J., Hunt M., Jackman S.D., Jaffe D.B., Jarvis E.D., Jiang H., Kazakov S., Kersey P.J., Kitzman J.O., Knight J.R., Koren S., Lam T.-W., Lavenier D., Laviolette F., Li Y., Li Z., Liu B., Liu Y., Luo R., MacCallum I., MacManes M.D., Maillet N., Melnikov S., Naquin D., Ning Z., Otto T.D., Paten B., Paulo O.S., Phillippy A.M., Pina-Martins F., Place M., Przybylski D., Qin X., Qu C., Ribeiro F.J., Richards S., Rokhsar D.S., Ruby J.G., Scalabrin S., Schatz M.C., Schwartz D.C., Sergushichev A., Sharpe T., Shaw T.I., Shendure J., Shi Y., Simpson J.T., Song H., Tsarev F., Vezzi F., Vicedomini R., Vieira B.M., Wang J., Worley K.C., Yin S., Yiu S.-M., Yuan J., Zhang G., Zhang H., Zhou S., and Korf I.F.
Assemblathon 2: Evaluating de novo methods of genome assembly in three vertebrate species.
GigaScience, 2(1):no pagination, 2013.

[379] April A. Jauhal and Richard D. Newcomb.
Assessing genome assembly quality prior to downstream analysis: N50 versus BUSCO.
Molecular Ecology Resources, 21(5):1416–1421, 2021.

[380] Hana Kocour Kroupová, Olga Valentová, Zdeňka Svobodová, Pavel Šauer, and Jana Máchová.
Toxic effects of nitrite on freshwater organisms: a review.
Reviews in Aquaculture, 10(3):525–542, 2018.

[381] Christopher B. Field, Michael J. Behrenfeld, James T. Randerson, and Paul Falkowski.
Primary production of the biosphere: Integrating terrestrial and oceanic components.
Science, 281(5374):237–240, 1998.

[382] Yantao Liang, Yongyu Zhang, Nannan Wang, Tingwei Luo, Yao Zhang, and Richard B. Rivkin.
Estimating primary production of picophytoplankton using the carbon-based ocean productivity model: A preliminary study.
Frontiers in Microbiology, 8(OCT):1–12, 2017.

[383] Ferran Garcia-Pichel, Jayne Belnap, Susanne Neuer, and Ferdinand Schanz.
Estimates of global cyanobacterial biomass and its distribution.
Algological Studies, 109(1):213–227, 2009.

[384] Nathan A. Ahlgren, Bernard Shafer Belisle, and Michael D. Lee.

Genomic mosaicism underlies the adaptation of marine *Synechococcus* ecotypes to distinct oceanic iron niches.

Environmental Microbiology, 22(5):1801–1815, 2020.

[385] M. C. Muñoz-Marín, G. Gómez-Baena, A. López-Lozano, J. A. Moreno-Cabezuelo, J. Díez, and J. M. García-Fernández.

Mixotrophy in marine picocyanobacteria: use of organic compounds by *Prochlorococcus* and *Synechococcus*.

ISME Journal, 14(5):1065–1073, 2020.

[386] John G. Stockner.

An overview from marine and freshwater ecosystems.

Limnol. Oceanogr., 33(4):765–775, 1988.

[387] Cristiana Callieri, Manuela Coci, Gianluca Corno, Miroslav Macek, Beatriz Modenutti, Esteban Balseiro, and Roberto Bertoni.

Phylogenetic diversity of nonmarine picocyanobacteria.

FEMS Microbiology Ecology, 85(2):293–301, 2013.

[388] S. C. Maberly, L. King, M. M. Dent, R. I. Jones, and C. E. Gibson.

Nutrient limitation of phytoplankton and periphyton growth in upland lakes.

Freshwater Biology, 47(11):2136–2152, 2002.

[389] Carrine E. Blank and Nancy W. Hinman.

Cyanobacterial and algal growth on chitin as a source of nitrogen; ecological, evolutionary, and biotechnological implications.

Algal Research, 15:152–163, 2016.

[390] Björn Watzer, Philipp Spät, Niels Neumann, Moritz Koch, Roman Sobotka, Boris MacEk, Oliver Hennrich, and Karl Forchhammer.

The signal transduction protein PII controls ammonium, nitrate and urea uptake in cyanobacteria.

Frontiers in Microbiology, 10(JUN):1–20, 2019.

[391] Keisuke Kurita.
Chitin and chitosan: Functional biopolymers from marine crustaceans.
Marine Biotechnology, 8(3):203–226, 2006.

[392] Mikhail V. Zubkov, Glen A. Tarran, Isabelle Mary, and Bernhard M. Fuchs.
Differential microbial uptake of dissolved amino acids and amino sugars in surface waters of the Atlantic Ocean.
Journal of Plankton Research, 30(2):211–220, 2008.

[393] D. Somashekar and Richard Joseph.
Chitosanases - Properties and applications: A review.
Bioresource Technology, 55(1):35–45, 1996.

[394] Wei Lin, Zhenyan Zhang, Yiling Chen, Qi Zhang, Mingjing Ke, Tao Lu, and Haifeng Qian.
The mechanism of different cyanobacterial responses to glyphosate.
Journal of Environmental Sciences (China), 125:258–265, 2023.

[395] Manuel Castro Berman, María E. Llames, Priscilla Minotti, Paulina Fermani, María V. Quiroga, Marcela A. Ferraro, Sebastián Metz, and Horacio E. Zagarese.
Field evidence supports former experimental claims on the stimulatory effect of glyphosate on picocyanobacteria communities.
Science of the Total Environment, 701:134601, 2020.

[396] Wenjing Wang, Ming Jiang, and Yanqing Sheng.
Glyphosate Accelerates the Proliferation of *Microcystis aeruginosa*, a Dominant Species in Cyanobacterial Blooms.
Environmental Toxicology and Chemistry, 40(2):342–351, 2021.

[397] Giliane Zanchett and Eduardo C. Oliveira-Filho.
Cyanobacteria and cyanotoxins: From impacts on aquatic ecosystems and human health to anticarcinogenic effects.
Toxins, 5(10):1896–1917, 2013.

[398] Aleksandra Bukowska, Tomasz Kaliński, Michał Koper, Iwona Kostrzewska-Szlakowska, Jan Kwiatowski, Hanna Mazur-Marzec, and Iwona Jasser.
Predicting blooms of toxic cyanobacteria in eutrophic lakes with diverse cyanobacterial communities.
Scientific Reports, 7(1):1–12, 2017.

[399] Justin D. Chaffin, Sachidananda Mishra, Douglas D. Kane, Darren L. Bade, Keara Stanislawczyk, Kristen N. Slodysko, Kevin W. Jones, Eric M. Parker, and Erica L. Fox.
Cyanobacterial blooms in the central basin of Lake Erie: Potentials for cyanotoxins and environmental drivers.
Journal of Great Lakes Research, 45(2):277–289, 2019.

[400] Nadav Kashtan, Sara E. Roggensack, Sébastien Rodrigue, Jessie W. Thompson, Steven J. Biller, Allison Coe, Huiming Ding, Pekka Marttinen, Rex R. Malmstrom, Roman Stocker, Michael J. Follows, Ramunas Stepanauskas, and Sallie W. Chisholm.
Single-cell genomics reveals hundreds of coexisting subpopulations in wild *Prochlorococcus*.
Science, 344(6182):416–420, 2014.

[401] Cuncun Qiao, Mingyi Zhang, Quan Luo, and Xuefeng Lu.
Identification of two two-component signal transduction mutants with enhanced sucrose biosynthesis in *Synechococcus elongatus* PCC 7942.
Journal of Basic Microbiology, 59(5):465–476, 2019.

[402] Khaled A. Selim, Michael Haffner, Björn Watzer, and Karl Forchhammer.
Tuning the in vitro sensing and signaling properties of cyanobacterial PII protein by mutation of key residues.
Scientific reports, 9(1):18985, 2019.

[403] Deepa Varkey, Sophie Mazard, Martin Ostrowski, Sasha G. Tetu, Paul Haynes, and Ian T. Paulsen.
Effects of low temperature on tropical and temperate isolates of marine *Synechococcus*.
ISME Journal, 10(5):1252–1263, 2016.

[404] Nuala A. O'Leary, Mathew W. Wright, J. Rodney Brister, Stacy Ciufo, Diana Haddad, Rich McVeigh, Bhanu Rajput, Barbara Robbertse, Brian Smith-White, Danso Ako-Adjei, Alexander Astashyn, Azat Badretdin, Yiming Bao, Olga Blinkova, Vyacheslav Brover, Vyacheslav Chetvernin, Jinna Choi, Eric Cox, Olga Ermolaeva, Catherine M. Farrell, Tamara Goldfarb, Tripti Gupta, Daniel Haft, Eneida Hatcher, Wratko Hlavina, Vinita S. Joardar, Vamsi K. Kodali, Wenjun Li, Donna Maglott, Patrick Masterson, Kelly M. McGarvey, Michael R. Murphy, Kathleen O'Neill, Shashikant Pujar, Sanjida H. Rangwala, Daniel Rausch, Lillian D. Riddick, Conrad Schoch, Andrei Shkeda, Susan S. Storz, Hanzhen Sun, Francoise Thibaud-Nissen, Igor Tolstoy, Raymond E. Tully, Anjana R. Vatsan, Craig Wallin, David Webb, Wendy Wu, Melissa J. Landrum, Avi Kimchi, Tatiana Tatusova, Michael DiCuccio, Paul Kitts, Terence D. Murphy, and Kim D. Pruitt.
Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation.
Nucleic Acids Research, 44(D1):D733–D745, 2016.

[405] Frank J. Millero, Rainer Feistel, Daniel G. Wright, and Trevor J. McDougall.
The composition of Standard Seawater and the definition of the Reference-Composition Salinity Scale.
Deep-Sea Research Part I: Oceanographic Research Papers, 55(1):50–72, 2008.

[406] Carles Ibáñez, Peter James Sharpe, John W. Day, Jason N. Day, and Narcís Prat.
Vertical accretion and relative sea level rise in the Ebro Delta wetlands (Catalonia, Spain).
Wetlands, 30(5):979–988, 2010.

[407] Robert W. Sterner, Bonnie Keeler, Stephen Polasky, Rajendra Poudel, Kirsten Rhude, and Maggie Rogers.
Ecosystem services of Earth's largest freshwater lakes.
Ecosystem Services, 41(January 2019):101046, 2020.

[408] Jiří Komárek, Jeffrey R. Johansen, Jan Šmarda, and Otakar Strunecký.
Phylogeny and taxonomy of *Synechococcus*–like cyanobacteria.
Fottea, 20(2):171–191, 2020.

[409] Mary M. Allen and R. Y. Stanier.
Growth and Division of Some Unicellular Blue-green Algae.
J. gen. Microbiol., 51:199–202, 1968.

[410] Connor B. Driscoll, Kevin A. Meyer, Sigitas Šulčius, Nathan M. Brown, Gregory J. Dick, Huansheng Cao, Giedrius Gasiūnas, Albertas Timinskas, Yanbin Yin, Zachary C. Landry, Timothy G. Otten, Timothy W. Davis, Susan B. Watson, and Theo W. Dreher.
A closely-related clade of globally distributed bloom-forming cyanobacteria within the Nostocales.
Harmful Algae, 77:93–107, 2018.

[411] Xiaochuang Li, Theo W. Dreher, and Renhui Li.
An overview of diversity, occurrence, genetics and toxin production of bloom-forming *Dolichospermum* (*Anabaena*) species.
Harmful Algae, 54:54–68, 2016.

[412] Lam Tung Nguyen, Heiko A. Schmidt, Arndt Von Haeseler, and Bui Quang Minh.
IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies.
Molecular Biology and Evolution, 32(1):268–274, 2015.

[413] R Core Team.
R: A language and environment for statistical computing, 2022.

[414] Jiandong Zhang, Furun Li, Lijuan Long, and Sijun Huang.
Quantification of Marine Picocyanobacteria on Water Column Particles and in Sediments Using Real-Time PCR Reveals Their Role in Carbon Export.
mSphere, 7(6), 2022.

[415] Nicholas J Fuller, Dominique Marie, Daniel Vaulot, Anton F Post, and David J Scanlan.
Clade-Specific 16S Ribosomal DNA Oligonucleotides Reveal the Predominance of a Single Marine *Synechococcus* Clade throughout a Stratified Water Column in the Red Sea.
Applied and Environmental Microbiology, 69(5):2430–2443, 2003.

[416] Shane Nowack, Millie T. Olsen, George A. Schaible, Eric D. Becraft, Gaozhong Shen, Isaac Klapper, Donald A. Bryant, and David M. Ward.
The molecular dimension of microbial species: 2. *Synechococcus* strains representative of putative ecotypes inhabiting different depths in the Mushroom Spring microbial mat exhibit different adaptive and acclimative responses to light.
Frontiers in Microbiology, 6(JUN):1–13, 2015.

[417] Peeter Nõges, Tiina Nõges, Michela Ghiani, Fabrizio Sena, Roswitha Fresner, Maria Friedl, and Johanna Mildner.
Increased nutrient loading and rapid changes in phytoplankton expected with climate change in stratified South European lakes: Sensitivity of lakes with different trophic state and catchment properties.
Hydrobiologia, 667(1):255–270, 2011.

[418] Yu Cao, Saara Olsen, María Florencia Gutierrez, Sandra Brucet, Thomas A. Davidson, Wei Li, Torben L. Lauridsen, Martin Søndergaard, and Erik Jeppesen.
Temperature effects on periphyton, epiphyton and epipelon under a nitrogen pulse in low-nutrient experimental freshwater lakes.
Hydrobiologia, 795(1):267–279, 2017.

[419] Jesús M. Mercado, Dolores Cortés, Francisco Gómez-Jakobsen, Candela García-Gómez, Sophia Ouaissa, Lidia Yebra, Isabel Ferrera, Nerea Valcárcel-Pérez, María López, Rocío García-Muñoz, Aranzazu Ramos, Jaime Bernardeau, María Dolores Belando, Eugenio Fraile-Nuez, and Juan M. Ruíz.
Role of small-sized phytoplankton in triggering an ecosystem disruptive algal bloom in a Mediterranean hypersaline coastal lagoon.
Marine Pollution Bulletin, 164(December 2020), 2021.

[420] Nagamitsu Maie, Kathleen J. Parish, Akira Watanabe, Heike Knicker, Ronald Benner, Tomonori Abe, Karl Kaiser, and Rudolf Jaffé.
Chemical characteristics of dissolved organic nitrogen in an oligotrophic subtropical coastal ecosystem.

Geochimica et Cosmochimica Acta, 70(17):4491–4506, 2006.

[421] Xin Yao, Guangwei Zhu, Linlin Cai, Mengyuan Zhu, Linlin Zhao, Guang Gao, and Boqiang Qin.
Geochemical Characteristics of Amino Acids in Sediments of Lake Taihu, A Large, Shallow, Eutrophic Freshwater Lake of China.
Aquatic Geochemistry, 18(3):263–280, 2012.

[422] Yan Chen, Gui Peng Yang, Li Liu, Peng Yan Zhang, and Wei Song Leng.
Sources, behaviors and degradation of dissolved organic matter in the East China Sea.
Journal of Marine Systems, 155:84–97, 2016.

[423] Xin Yao, Tuantuan Fan, Guang Gao, Li Liu, Jianying Chao, and Hao Liu.
Spatiotemporal pattern and biodegradation process of amino acids in the large shallow eutrophic lake Taihu, China.
Environmental Science and Pollution Research, 30(5):12584–12595, 2023.

[424] Abdul Jalil, Yiping Li, Ke Zhang, Xiaomeng Gao, Wencai Wang, Hafiz Osama Sarwar Khan, Baozhu Pan, Salar Ali, and Kumud Acharya.
Wind-induced hydrodynamic changes impact on sediment resuspension for large, shallow Lake Taihu, China.
International Journal of Sediment Research, 34(3):205–215, 2019.

[425] Daiske Honda, Akira Yokota, and Junta Sugiyama.
Detection of seven major evolutionary lineages in cyanobacteria based on the 16S rRNA gene sequence analysis with new sequences of five marine *Synechococcus* strains.
Journal of Molecular Evolution, 48(6):723–739, 1999.

[426] S. P. Seitzinger, J. A. Harrison, Egon Dumont, Arthur H.W. Beusen, and A. F. Bouwman.
Sources and delivery of carbon, nitrogen, and phosphorus to the coastal zone: An overview of Global Nutrient Export from Watersheds (NEWS) models and their application.
Global Biogeochemical Cycles, 19(4):1–11, 2005.

[427] Mikhail V. Zubkov, Bernhard M. Fuchs, Glen A. Tarran, Peter H. Burkill, and Rudolf Amann.

High rate of uptake of organic nitrogen compounds by *Prochlorococcus* cyanobacteria as a key to their dominance in oligotrophic oceanic waters.
Applied and Environmental Microbiology, 69(2):1299–1304, 2003.

[428] Chen Gong, Ruyuan Jiao, Weijin Yan, Qibiao Yu, Qingqian Li, Peipei Zhang, Yanqiang Li, and Dongsheng Wang.
Enhanced chemodiversity, distinctive molecular signature and diurnal dynamics of dissolved organic matter in streams of two headwater catchments, Southeastern China.
Water Research, 211:118052, 2022.

[429] Saara Suominen, Gonzalo V. Gomez-Saez, Thorsten Dittmar, Jaap S. Sinninghe Damsté, and Laura Villanueva.
Interplay between microbial community composition and chemodiversity of dissolved organic matter throughout the Black Sea water column redox gradient.
Limnology and Oceanography, 67(2):329–347, 2022.

[430] Kent K. Cavender-Bares, David M. Karl, and Sallie W. Chisholm.
Nutrient gradients in the western North Atlantic Ocean: Relationship to microbial community structure and comparison to patterns in the Pacific Ocean.
Deep-Sea Research Part I: Oceanographic Research Papers, 48(11):2373–2395, 2001.

[431] Isabelle Mary, Laurence Garczarek, Glen A. Tarran, Christian Kolowrat, Matthew J. Terry, David J. Scanlan, Peter H. Burkill, and Mikhail V. Zubkov.
Diel rhythmicity in amino acid uptake by *Prochlorococcus*.
Environmental Microbiology, 10(8):2124–2131, 2008.

[432] David M. Karl, Robert R. Bidigare, Matthew J. Church, John E. Dore, Ricardo M. Letelier, Claire Mahaffey, and Jonathan P. Zehr.
The Nitrogen Cycle in the North Pacific Trades Biome: An Evolving Paradigm.
In *Nitrogen in the Marine Environment*, pages 705–769. Academic Press, 2008.

[433] Jariani Jani and Gurpal S. Toor.

Composition, sources, and bioavailability of nitrogen in a longitudinal gradient from freshwater to estuarine waters.

Water Research, 137:344–354, 2018.

[434] Clayton W. Harris, Gavin N. Rees, Rick J. Stoffels, John Pengelly, Kirsten Barlow, and Ewen Silvester.

Longitudinal trends in concentration and composition of dissolved organic nitrogen (DON) in a largely unregulated river system.

Biogeochemistry, 139(2):139–153, 2018.

[435] N.O.G. Jørgensen.

Organic Nitrogen.

In *Encyclopedia of Inland Waters*, pages 832–851. Academic Press, Oxford, 2009.

[436] Giovanna Capovilla, Rogier Braakman, Gregory P. Fournier, Thomas Hackl, Julia Schwartzman, Xinda Lu, Alexis Yelton, Krista Longnecker, Melissa C. Kido Soule, Elaina Thomas, Gretchen Swarr, Alessandro Mongera, Jack G. Payette, Kurt G. Castro, Jacob R. Waldbauer, Elizabeth B. Kujawinski, Otto X. Cordero, and Sallie W. Chisholm.

Chitin utilization by marine picocyanobacteria and the evolution of a planktonic lifestyle.

PNAS, 120(20):e2213271120, 2023.

[437] Henry Michel Cauchie.

Chitin production by arthropods in the hydrosphere.

Hydrobiologia, 470:63–95, 2002.

[438] Krista E. Köllner, Dörte Carstens, Esther Keller, Francisco Vazquez, Carsten J. Schubert, Josef Zeyer, and Helmut Bürgmann.

Bacterial chitin hydrolysis in two lakes with contrasting trophic statuses.

Applied and Environmental Microbiology, 78(3):695–704, 2012.

[439] Florencia Biancalana and Anabela Anahí Berasategui.

Size-fractionated chitin contribution to seston, with linkages to the copepod *Acartia.*

Ocean and Coastal Research, 68:1–12, 2020.

[440] Zhanru Shao, Yann Thomas, Lea Hembach, Xiaohui Xing, Delin Duan, Bruno M. Moerschbacher, Vincent Bulone, Leila Tirichine, and Chris Bowler.
Comparative characterization of putative chitin deacetylases from *Phaeodactylum tricornutum* and *Thalassiosira pseudonana* highlights the potential for distinct chitin-based metabolic processes in diatoms.
New Phytologist, 221:1890–1905, 2018.

[441] H. Pizarro, M. S. Vera, A. Vinocur, G. Pérez, M. Ferraro, R. J. Menéndez Helman, and M. dos Santos Afonso.
Glyphosate input modifies microbial community structure in clear and turbid freshwater systems.
Environmental Science and Pollution Research, 23(6):5143–5153, 2016.

[442] Tao Lu, Nuohan Xu, Qi Zhang, Zhenyan Zhang, Andries Debognies, Zhigao Zhou, Liwei Sun, and Haifeng Qian.
Understanding the influence of glyphosate on the structure and function of freshwater microbial community in a microcosm.
Environmental Pollution, 260:114012, 2020.

[443] Kerri Finlay, Alain Patoine, Derek B. Donald, Matthew J. Bogard, and Peter R. Leavitt.
Experimental evidence that pollution with urea can degrade water quality in phosphorus-rich lakes of the Northern Great Plains.
Limnology and Oceanography, 55(3):1213–1230, 2010.

[444] Toshio Sakamoto, Victoria B. Delgaizo, and Donald A. Bryant.
Growth on urea can trigger death and peroxidation of the cyanobacterium *Synechococcus* sp. strain PCC 7002.
Applied and Environmental Microbiology, 64(7):2361–2366, 1998.

[445] Jonathan P. Zehr, John B. Waterbury, Patricia J. Turner, Joseph P. Montoya, Enoma Omoregie, Grieg F. Steward, Andrew Hansen, and David M. Karl.
Unicellular cyanobacteria fix N2 in the subtropical north Pacific Ocean.
Nature, 412(6847):635–638, 2001.

[446] Anne Soisig Steunou, Devaki Bhaya, Mary M. Bateson, Melanie C. Melendrez, David M. Ward, Eric Brecht, John W. Peters, Michael Kühl, and Arthur R. Grossman.

In situ analysis of nitrogen fixation and metabolic switching in unicellular thermophilic cyanobacteria inhabiting hot spring microbial mats.

Proceedings of the National Academy of Sciences of the United States of America, 103(7):2398–2403, 2006.

[447] Henrique F. Santos, Flávia L. Carmo, Gustavo Duarte, Francisco Dini-Andreote, Clovis B. Castro, Alexandre S. Rosado, Jan Dirk Van Elsas, and Raquel S. Peixoto.

Climate change affects key nitrogen-fixing bacterial populations on coral reefs.

ISME Journal, 8(11):2272–2279, 2014.

[448] Mausmi P. Mehta, David A. Butterfield, and John A. Baross.

Phylogenetic diversity of nitrogenase (*nifH*) genes in deep-sea and hydrothermal vent environments of the Juan de Fuca Ridge.

Applied and Environmental Microbiology, 69(2):960–970, 2 2003.

[449] Emanuele Caravati, Cristiana Callieri, Beatriz Modenutti, Gianluca Corno, Esteban Balseiro, Roberto Bertoni, and Luigi Michaud.

Picocyanobacterial assemblages in ultraoligotrophic Andean lakes reveal high regional microdiversity.

Journal of Plankton Research, 32(3):357–366, 3 2010.

[450] Natalia V. Ivanikova, Linda C. Popels, R. Michael L. McKay, and George S. Bullerjahn.

Lake superior supports novel clusters of cyanobacterial picoplankton.

Applied and Environmental Microbiology, 73(12):4055–4065, 6 2007.

[451] Maren Voss, Hermann W. Bange, Joachim W. Dippner, Jack J. Middelburg, Joseph P. Montoya, and Bess Ward.

The marine nitrogen cycle: Recent discoveries, uncertaintiesand the potential relevance of climate change.

Philosophical Transactions of the Royal Society B: Biological Sciences, 368(1621), 2013.

[452] Emily J. Zakem, Alia Al-Haj, Matthew J. Church, Gert L. Van Dijken, Stephanie Dutkiewicz, Sarah Q. Foster, Robinson W. Fulweiler, Matthew M. Mills, and Michael J. Follows.
Ecological control of nitrite in the upper ocean.
Nature Communications, 9(1), 2018.

[453] Karl Forchhammer and Khaled A Selim.
Carbon/Nitrogen Homeostasis Control in Cyanobacteria.
FEMS Microbiology Reviews, 44(1):1–21, 2019.

[454] Mario Giordano, Charles A Goodman, Fengying Huang, John A Raven, and Zuoxi Ruan.
A mechanistic study of the influence of nitrogen and energy availability on the NH_4^+ sensitivity of nitrogen assimilation in *Synechococcus*.
Journal of Experimental Botany, 73(16):5596–5611, 2022.

[455] Antje Baier, Wiebke Winkler, Thomas Korte, Wolfgang Lockau, and Anne Karradt.
Degradation of phycobilisomes in *Synechocystis* sp. PCC6803 evidence for essential formation of an NblA1/NblA2 heterodimer and its codegradation by a Clp protease complex.
Journal of Biological Chemistry, 289(17):11755–11766, 2014.

[456] Hans C. Bernstein.
Unlocking the Constraints of Cyanobacterial Productivity :.
mBio, 7(4):1–10, 2016.

[457] María Del Carmen Muñoz-Marín, Ignacio Luque, Mikhail V. Zubkov, Polly G. Hill, Jesús Diez, and José Manuel García-Fernández.
Prochlorococcus can use the Pro1404 transporter to take up glucose at nanomolar concentrations in the Atlantic Ocean.
Proceedings of the National Academy of Sciences of the United States of America, 110(21):8597–8602, 2013.

[458] Jai Kumar Gupta, Preeti Rai, Kavish Kumar Jain, and Shireesh Srivastava.
Overexpression of bicarbonate transporters in the marine cyanobacterium *Synechococcus* sp. PCC 7002 increases growth rate and glycogen accumulation.

Biotechnology for Biofuels, 13(1):1–12, 2020.

[459] Shin Ichi Maeda, Akio Murakami, Hisashi Ito, Ayumi Tanaka, and Tatsuo Omata.
Functional characterization of the FNT family nitrite transporter of marine picocyanobacteria.
Life, 5(1):432–446, 2015.

[460] Olga A. Kutovaya, Robert Michael L. McKay, and George S. Bullerjahn.
Detection and expression of genes for phosphorus metabolism in picocyanobacteria from the Laurentian Great Lakes.
Journal of Great Lakes Research, 39(4):612–621, 2013.

[461] Suzanne Wuijts, Dico Fraters, Sandra Boekhold, and Richard Van DUIJNEN.
Monitoring of nitrogen in water in the EU: Legal framework, effects of nitrate, design principles, effectiveness and future developments Policy Department for Citizens' Rights and Constitutional Affairs Directorate-General for Internal Policies PE.
Technical report, European Parliament, 2022.

[462] Lauren E. Krausfeldt, Abigail T. Farmer, Hector F. Castro Gonzalez, Brittany N. Zepernick, Shawn R. Campagna, and Steven W. Wilhelm.
Urea Is Both a Carbon and Nitrogen Source for *Microcystis aeruginosa*: Tracking 13C Incorporation at Bloom pH Conditions.
Frontiers in Microbiology, 10:1–12, 2019.

[463] Akira Watanabe, Kiyoshi Tsutsuki, Yudzuru Inoue, Nagamitsu Maie, Lulie Melling, and Rudolf Jaffé.
Composition of dissolved organic nitrogen in rivers associated with wetlands.
Science of the Total Environment, 493:220–228, 2014.

[464] Aporn Bualuang and Aran Incharoensakdi.
Growth enhancing effect of exogenous glycine and characterization of its uptake in halotolerant cyanobacterium *Aphanothece halophytica*.
World Journal of Microbiology and Biotechnology, 31(2):379–384, 2015.

[465] Lonnie O. Ingram and Roy A. Jensen.
Growth inhibition by L-phenylalanine in *Agmenellum quadruplicatum* - A clue to some amino acid interrelationships.
Archiv für Mikrobiologie, 91(3):221–233, 1973.

[466] Beata Żyszka-Haberecht, Emilia Niemczyk, and Jacek Lipok.
Metabolic relation of cyanobacteria to aromatic compounds.
Applied Microbiology and Biotechnology, 103(3):1167–1178, 2019.

[467] Margaret R. Mulholland, Patricia M. Glibert, Gry Miné Berg, Laurie Van Heukelem, Silvio Pantoja, and Cindy Lee.
Extracellular amino acid oxidation by microplankton: A cross-ecosystem comparison.
Aquatic Microbial Ecology, 15(2):141–152, 1998.

[468] Vanessa K. Michelou, Matthew T. Cottrell, and David L. Kirchman.
Light-stimulated bacterial production and amino acid assimilation by cyanobacteria and other microbes in the North Atlantic Ocean.
Applied and Environmental Microbiology, 73(17):5539–5546, 2007.

[469] Nishant Pappireddi, Lance Martin, and Martin Wühr.
A Review on Quantitative Multiplexed Proteomics.
ChemBioChem, 20(10):1210–1224, 2019.

[470] Henk Bolhuis, Ina Severin, Veronique Confurius-Guns, Ute I.A. Wollenzien, and Lucas J. Stal.
Horizontal transfer of the nitrogen fixation gene cluster in the cyanobacterium *Microcoleus chthonoplastes*.
ISME Journal, 4(1):121–130, 2010.

[471] Kanika Mehta, Damini Jaiswal, Monalisha Nayak, Charulata B. Prasannan, Pramod P. Wangikar, and Sanjeeva Srivastava.
Elevated carbon dioxide levels lead to proteome-wide alterations for optimal growth of a fast-growing cyanobacterium, *Synechococcus elongatus* PCC 11801.
Scientific Reports, 9(1):1–14, 2019.

[472] Yuan-Yuan Li, Xiao-Huang Chen, Cheng Xue, Hao Zhang, Geng Sun, Zhang-Xian Xie, Lin Lin, and Da-Zhi Wang.
Proteomic Response to Rising Temperature in the Marine Cyanobacterium *Synechococcus* Grown in Different Nitrogen Sources.
Frontiers in Microbiology, 10:1–13, 2019.

[473] Sijia Xu, Ying Liu, Jian Zhang, and Baoyu Gao.
Proteomic mechanisms for the combined stimulatory effects of glyphosate and antibiotic contaminants on *Microcystis aeruginosa*.
Chemosphere, 267:129244, 2021.

[474] Shuaitong Liu, Ying Liu, and Jian Zhang.
Proteomic mechanisms for the regulation of growth, photosynthetic activity and nitrogen fixation in *Nostoc* sp. PCC 7120 exposed to three antibiotic contaminants.
Ecotoxicology and Environmental Safety, 225:112753, 2021.

[475] Kimberly M. Wegener, Abhay K. Singh, Jon M. Jacobs, Thanura Elvitigala, Eric A. Welsh, Nir Keren, Marina A. Gritsenko, Bijoy K. Ghosh, David G. Camp, Richard D. Smith, and Himadri B. Pakrasi.
Global proteomics reveal an atypical strategy for carbon/nitrogen assimilation by a cyanobacterium under diverse environmental perturbations.
Molecular and Cellular Proteomics, 9(12):2678–2689, 2010.

[476] Philipp Doppler, Ricarda Kriechbaum, Benedikt Singer, and Oliver Spadiut.
Make microalgal cultures axenic again – a fast and simple workflow utilizing fluorescence-activated cell sorting.
Journal of Microbiological Methods, 186:106256, 2021.

[477] C. E.M. Lloyd, L. A. Mena-Rivera, J. A. Pemberton, P. J. Johnes, D. L. Jones, C. A. Yates, F. L. Brailsford, H. C. Glanville, C. A. McIntyre, and R. P. Evershed.
Characterisation of riverine dissolved organic matter using a complementary suite of chromatographic and mass spectrometric methods.

Biogeochemistry, pages 1–24, 2022.

[478] T. P. Burt, N. J.K. Howden, F. Worrall, and M. J. Whelan.
Long-term monitoring of river water nitrate: How much data do we need?
Journal of Environmental Monitoring, 12(1):71–79, 2010.

[479] Vanessa J. Swarbrick, Nathanael T. Bergbusch, and Peter R. Leavitt.
Spatial and temporal patterns of urea content in a eutrophic stream continuum on the Northern Great Plains.
Biogeochemistry, 157(2):171–191, 2022.

[480] Kathleen Sprouffske and Andreas Wagner.
Growthcurver: An R package for obtaining interpretable metrics from microbial growth curves.
BMC Bioinformatics, 17(1):17–20, 2016.

[481] K.J. Gold, E.L. Mozurewich, K.S. Puder, and M.C. Treadwell.
The SEQUEST Family Tree.
Journal of the American Society for Mass Spectrometry, 26(11):1814–9, 2015.

[482] Shengchaoo Yu, Mingkun Yang, Jie Xiong, Qi Zhang, Xinxin Gao, Wei Miao, and Feng Ge.
Proteogenomic Analysis Provides Novel Insight into Genome Annotation and Nitrogen Metabolism in *Nostoc* sp. PCC 7120.
Microbiology Spectrum, 9(2):e0049021, 2021.

[483] Qi Zhang, Shengchao Yu, Qiang Wang, Mingkun Yang, and Feng Ge.
Quantitative Proteomics Reveals the Protein Regulatory Network of *Anabaena* sp. PCC 7120 under Nitrogen Deficiency.
Journal of Proteome Research, 20(8):3963–3976, 2021.

[484] Fallen Teoh, Bhumika Shah, Martin Ostrowski, and Ian Paulsen.
Comparative membrane proteomics reveal contrasting adaptation strategies for coastal and oceanic marine *Synechococcus* cyanobacteria.
Environmental Microbiology, 22(5):1816–1828, 2020.

[485] Olga A. Koksharova, Ivan O. Butenko, Olga V. Pobeguts, Nina A. Safronova, and Vadim M. Govorun.
Proteomic Insights into Starvation of Nitrogen-Replete Cells of *Nostoc* sp. PCC 7120 under β-N-Methylamino-L-Alanine (BMAA) Treatment.
Toxins, 12(6), 2020.

[486] Min Chen, Miguel A. Hernandez-Prieto, Patrick C. Loughlin, Yaqiong Li, and Robert D. Willows.
Genome and proteome of the chlorophyll f-producing cyanobacterium *Halomicronema hongdechloris*: Adaptative proteomic shifts under different light conditions.
BMC Genomics, 20(1):1–15, 2019.

[487] Laura L Baers, Lisa M Breckels, Lauren A Mills, Laurent Gatto, Michael Deery, Tim J Stevens, Christopher J Howe, Kathryn S Lilley, and David J Lea-Smith.
Proteome mapping of a cyanobacterium reveals distinct compartment organisation and cell-dispersed metabolism.
Plant Physiology, 181:pp.00897.2019, 2019.

[488] Dijun Du, Hendrik W. van Veen, Satoshi Murakami, Klaas M. Pos, and Ben F. Luisi.
Structure, mechanism and cooperation of bacterial multidrug transporters.
Current Opinion in Structural Biology, 33(Figure 1):76–91, 2015.

[489] Akshai Iyengar and Stephen L. Bearne.
Aspartate-107 and leucine-109 facilitate efficient coupling of glutamine hydrolysis to CTP synthesis by *Escherichia coli* CTP synthase.
Biochemical Journal, 369(3):497–507, 2 2003.

[490] Marion Eisenhut, Hermann Bauwe, and Martin Hagemann.
Glycine accumulation is toxic for the cyanobacterium *Synechocystis* sp. strain PCC 6803, but can be compensated by supplementation with magnesium ions.
FEMS Microbiology Letters, 277(2):232–237, 2007.

[491] E. B. Newman, G. Batist, J. Fraser, S. Isenberg, P. Weyman, and V. Kapoor.
The use of glycine as nitrogen source by *Escherichia coli* K12.

BBA - General Subjects, 421(1):97–105, 1976.

[492] A. C. Chakrabarti.
Permeability of membranes to amino acids and modified amino acids: Mechanisms involved in translocation.
Amino Acids, 6(3):213–229, 1994.

[493] Larry Reitzer.
Catabolism of Amino Acids and Related Compounds.
EcoSal Plus, 1(2), 2005.

[494] Sunny Park and James A. Imlay.
High levels of intracellular cysteine promote oxidative DNA damage by driving the Fenton reaction.
Journal of Bacteriology, 185(6):1942–1950, 2003.

[495] Kazuhiro Takumi and Gen Nonaka.
Bacterial cysteine-inducible cysteine resistance systems.
Journal of Bacteriology, 198(9):1384–1392, 2016.

[496] Jesse Parry and David P. Clark.
Identification of a CysB-regulated gene involved in glutathione transport in *Escherichia coli*.
FEMS Microbiology Letters, 209(1):77–81, 2002.

[497] Yong Liang Jiang, Xue Ping Wang, Hui Sun, Shu Jing Han, Wei Fang Li, Ning Cui, Gui Ming Lin, Ju Yuan Zhang, Wang Cheng, Dong Dong Cao, Zhi Yong Zhang, Cheng Cai Zhang, Yuxing Chen, and Cong Zhao Zhou.
Coordinating carbon and nitrogen metabolic signaling through the cyanobacterial global repressor NdhR.
Proceedings of the National Academy of Sciences of the United States of America, 115(2):403–408, 2017.

[498] Robert L. Bertranda.

Lag phase is a dynamic, organized, adaptive, and evolvable period that prepares bacteria for cell division.
Journal of Bacteriology, 201(7):1–21, 2019.

[499] S C Maberly, M M De Ville, S J Thackeray, G Ainsworth, F Carse, J M Fletcher, R Groben, P Hodgson, J B James, J L Kelly, C D Vincent, and D R Wilson.
A survey of the lakes of the English Lake District: The Lakes Tour 2005.
Technical report, Centre for Ecology & Hydrology, 2005.

[500] Arun K. Yadav, Adam J. Carroll, Gonzalo M. Estavillo, Greg J. Rebetzke, and Barry J. Pogson.
Wheat drought tolerance in the field is predicted by amino acid responses to glasshouse-imposed drought.
Journal of Experimental Botany, 70(18):4931–4947, 2019.

[501] Viridiana Olin-Sandoval, Jason Shu Lim Yu, Leonor Miller-Fleming, Mohammad Tauqeer Alam, Stephan Kamrad, Clara Correia-Melo, Robert Haas, Joanna Segal, David Alejandro Peña Navarro, Lucia Herrera-Dominguez, Oscar Méndez-Lucio, Jakob Vowinckel, Michael Mülleder, and Markus Ralser.
Lysine harvesting is an antioxidant strategy and triggers underground polyamine metabolism.
Nature, 572(7768):249–253, 2019.

[502] Paulo Arruda and Pedro Barreto.
Lysine Catabolism Through the Saccharopine Pathway: Enzymes and Intermediates Involved in Plant Responses to Abiotic and Biotic Stress.
Frontiers in Plant Science, 11:1–10, 2020.

[503] Shota Isogai and Hiroshi Takagi.
Enhancement of lysine biosynthesis confers high-temperature stress tolerance to *Escherichia coli* cells.
Applied Microbiology and Biotechnology, 105(18):6899–6908, 2021.

[504] Izabella Ap Neshich, Eduardo Kiyota, and Paulo Arruda.

Genome-wide analysis of lysine catabolism in bacteria reveals new connections with osmotic stress resistance.
ISME Journal, 7(12):2400–2410, 2013.

[505] Usheer Kanjee and Walid A. Houry.
Mechanisms of acid resistance in *Escherichia coli*.
Annual Review of Microbiology, 67:65–81, 2013.

[506] Vendula Krynicka, Jens Georg, Phillip J. Jackson, Mark J. Dickman, C. Neil Hunter, Matthias E. Futschik, Wolfgang R. Hess, and Josef Komenda.
Depletion of the FtsH1/3 Proteolytic Complex Suppresses the Nutrient Stress Response in the Cyanobacterium *Synechocystis* sp strain PCC 6803.
Plant Cell, 31(12):2912–2928, 2019.

[507] Rafia Mir, Shais Jallu, and T. P. Singh.
The shikimate pathway: Review of amino acid sequence, function and three-dimensional structures of the enzymes.
Critical Reviews in Microbiology, 41(2):172–189, 2015.

[508] Piyoosh Kumar Babele, Jay Kumar, and Venkatesh Chaturvedi.
Proteomic De-Regulation in Cyanobacteria in Response to Abiotic Stresses.
Frontiers in Microbiology, 10:1–22, 2019.

[509] Xi Huang, Wenji Lao, Youci Zhou, Yunwei Sun, and Qijun Wang.
Glutamate dehydrogenase enables *Salmonella* to survive under oxidative stress and escape from clearance in macrophages.
FEBS Letters, 596(1):81–94, 2022.

[510] Amel Latifi, Marion Ruiz, and Cheng Cai Zhang.
Oxidative stress in cyanobacteria.
FEMS Microbiology Reviews, 33(2):258–278, 2009.

[511] Anthony L. Shiver, Hendrik Osadnik, Jason M. Peters, Rachel A. Mooney, Peter I. Wu, Kemardo K. Henry, Hannes Braberg, Nevan J. Krogan, James C. Hu, Robert Landick, Kerwyn Casey Huang, and Carol A. Gross.
Chemical-genetic interrogation of RNA polymerase mutants reveals structure-function relationships and physiological tradeoffs.
Molecular Cell, 81(10):2201–2215, 5 2021.

[512] Hal Alper and Gregory Stephanopoulos.
Global transcription machinery engineering: A new approach for improving cellular phenotype.
Metabolic Engineering, 9(3):258–267, 5 2007.

[513] Jacob Beal, Natalie G. Farny, Traci Haddock-Angelli, Vinoo Selvarajah, Geoff S. Baldwin, Russell Buckley-Taylor, Markus Gershater, Daisuke Kiga, John Marken, Vishal Sanchania, Abigail Sison, and Christopher T. Workman.
Robust estimation of bacterial cell count from optical density.
Communications biology, 3(1):512, 2020.

[514] Suchandan Bemal and Arga Chandrashekar Anil.
Effects of salinity on cellular growth and exopolysaccharide production of freshwater *Synechococcus* strain CCAP1405.
Journal of Plankton Research, 40(1):46–58, 2018.

[515] Fariborz Faraji, Harshit Mahandra, and Ahmad Ghahreman.
Evaluation of Different Amino Acids on Growth and Cyanide Production by *Bacillus megaterium* for Gold Recovery.
Sustainability (Switzerland), 14(15), 2022.

[516] Joseph Oddy, Sarah Raffan, Mark D. Wilkinson, J. Stephen Elmore, and Nigel G. Halford.
Stress, nutrients and genotype: understanding and managing asparagine accumulation in wheat grain.
CABI Agriculture and Bioscience, 1(1):1–14, 2020.

[517] Kondethimmanahalli Chandramouli and Pei-Yuan Qian.

Proteomics: Challenges, Techniques and Possibilities to Overcome Biological Sample Complexity.
Human Genomics and Proteomics, 1(1), 2009.

[518] Dominik A. Megger, Leona L. Pott, Kristin Rosowski, Birgit Zülch, Stephanie Tautges, Thilo Bracht, and Barbara Sitek.
NHS-based tandem mass tagging of proteins at the level of whole cells: A critical evaluation in comparison to conventional TMT-labeling approaches for quantitative proteome analysis.
Analytical Sciences, 33(12):1387–1394, 2017.

[519] Angela L. Picciano and Brian R. Crane.
A nitric oxide synthase-like protein from *Synechococcus* produces NO/NO_3^- from L-arginine and NAPDH in a tetrahydrobiopterin- and Ca^{2+}-dependent manner.
Journal of Biological Chemistry, 294(3):jbc.RA119.008399, 2019.

[520] M. C. Ungerer, L. C. Johnson, and M. A. Herman.
Ecological genomics: Understanding gene and genome function in the natural environment.
Heredity, 100(2):178–183, 2008.

[521] Mats Jansson, Jan Karlsson, and Anders Jonsson.
Carbon dioxide supersaturation promotes primary production in lakes.
Ecology Letters, 15(6):527–532, 2012.

[522] Timothy D.J. Knowles, David R. Chadwick, Roland Bol, and Richard P. Evershed.
Tracing the rate and extent of N and C flow from ^{13}C,^{15}N-glycine and glutamate into individual de novo synthesised soil amino acids.
Organic Geochemistry, 41(12):1259–1268, 2010.

[523] L. R. Moore, M. Ostrowski, D. J. Scanlan, K. Feren, and T. Sweetsir.
Ecotypic variation in phosphorus-acquisition mechanisms within marine picocyanobacteria.
Aquatic Microbial Ecology, 39(3):257–269, 2005.

[524] Bhumika S. Shah, Benjamin A. Ford, Deepa Varkey, Halina Mikolajek, Christian Orr, Vitaliy Mykhaylyk, Raymond J. Owens, and Ian T. Paulsen.

Marine picocyanobacterial PhnD1 shows specificity for various phosphorus sources but likely represents a constitutive inorganic phosphate transporter.
ISME, pages 1–12, 2023.

[525] Irina N. Ilikchyan, R. Michael L. McKay, Jonathan P. Zehr, Sonya T. Dyhrman, and George S. Bullerjahn.
Detection and expression of the phosphonate transporter gene *phnD* in marine and freshwater picocyanobacteria.
Environmental Microbiology, 11(5):1314–1324, 2009.

[526] James N. Kinney, Seth D. Axen, and Cheryl A. Kerfeld.
Comparative analysis of carboxysome shell proteins.
Photosynthesis Research, 109(1-3):21–32, 2011.

[527] Avi I. Flamholz, Noam Prywes, Uri Moran, Dan Davidi, Yinon M. Bar-On, Luke M. Oltrogge, Rui Alves, David Savage, and Ron Milo.
Revisiting Trade-offs between Rubisco Kinetic Parameters.
Biochemistry, 58(31):3365–3376, 2019.

[528] Pedro J. Cabello-Yeves, David J. Scanlan, Cristiana Callieri, Antonio Picazo, Lena Schallenberg, Paula Huber, Juan J. Roda-Garcia, Maciej Bartosiewicz, Olga I. Belykh, Irina V. Tikhonova, Alberto Torcello-Requena, Paula Martin De Prado, Andrew D. Millard, Antonio Camacho, Francisco Rodriguez-Valera, and Richard J. Puxty.
α-cyanobacteria possessing form IA RuBisCO globally dominate aquatic habitats.
ISME Journal, 16:2421–2432, 2022.

[529] Kevin J Erratt, Irena F. Creed, and Charles G. Trick.
Differential Drawdown of Ammonium, Nitrate, and Urea by Freshwater Chlorophytes and Cyanobacteria.
Journal of Phycology, 56(2):458–468, 2020.

[530] Rui Wang, Zilian Zhang, Jia Sun, and Nianzhi Jiao.
Differences in bioavailability of canonical and non-canonical D-amino acids for marine microbes.

Science of the Total Environment, 733:139216, 2020.

[531] J. D. Thomas.
The role of dissolved organic matter, particularly free amino acids and humic substances, in freshwater ecosystems.
Freshwater Biology, 38(1):1–36, 1997.

[532] Rachel L. Sleighter and Patrick G. Hatcher.
The application of electrospray ionization coupled to ultrahigh resolution mass spectronomy for the molecular chatacterization of natural organic matter.
Journal of Mass Spectrometry, 42:559–574, 2007.

[533] A. P. Hunt, J. D. Parry, and J. Hamilton-Taylor.
Further evidence of elemental composition as an indicator of the bioavailability of humic substances to bacteria.
Limnology and Oceanography, 45(1):237–241, 2000.

[534] P. Carlsson, E. Granéli, and A. Z. Segatto.
Cycling of biologically available nitrogen in riverine humic substances between marine bacteria, a heterotrophic nanoflagellate and a photosynthetic dinoflagellate.
Aquatic Microbial Ecology, 18(1):23–36, 1999.

[535] Shanika L Amarasinghe, Shian Su, Xueyi Dong, Luke Zappia, Matthew E Ritchie, and Quentin Gouil.
Opportunities and challenges in long-read sequencing data analysis.
Genome Biology, 21(30):1–16, 2020.

[536] M. A Vega-Palas, E. Flores, and A. Herrero.
NtcA, a global nitrogen regulator from the cyanobacterium *Synechococcus* that belongs to the Crp family of bacterial regulators.
Molecular Microbiology, 6(13):1853–1859, 1992.

[537] Michael A. Harrison, Jeffrey N. Keen, John B.C. Findlay, and John F. Allen.

Modification of a *glnB*-like gene product by photosynthetic electron transport in the cyanobacterium *Synechococcus* 6301.
FEBS Letters, 264(1):25–28, 1990.

[538] M. Isabel Muro-Pastor, Jose C. Reyes, and Francisco J. Florencio.
The NADP$^+$-isocitrate dehydrogenase gene (*icd*) is nitrogen regulated in cyanobacteria.
Journal of Bacteriology, 178(14):4070–4076, 1996.

[539] Cheng Cai Zhang, Jichan Jang, Samer Sakr, and Li Wang.
Protein phosphorylation on Ser, Thr and Tyr residues in cyanobacteria.
Journal of Molecular Microbiology and Biotechnology, 9(3-4):154–166, 2006.

[540] Toshio Sakamoto, Kaori Inoue-Sakamoto, and Donald A. Bryant.
A novel nitrate/nitrite permease in the marine cyanobacterium *Synechococcus* sp. strain PCC 7002.
Journal of Bacteriology, 181(23):7363–7372, 1999.

[541] Nina A. Kamennaya, Mark Chernihovsky, and Anton F. Post.
The cyanate utilization capacity of marine unicellular Cyanobacteria.
Limnology and Oceanography, 53(6):2485–2494, 2008.

[542] Yoshimi Harano, Iwane Suzuki, Shin Ichi Maeda, Takakazu Kaneko, Satoshi Tabata, and Tatsuo Omata.
Identification and nitrogen regulation of the cyanase gene from the cyanobacteria *Synechocystis* sp. strain PCC 6803 and *Synechococcus* sp. strain PCC 7942.
Journal of Bacteriology, 179(18):5744–5750, 1997.

[543] Hiroaki Okuhara, Tomohiro Matsumura, Yuichi Fujita, and Toshiharu Hase.
Cloning and inactivation of genes encoding ferredoxin- and NADH-dependent glutamate synthases in the cyanobacterium *Plectonema boryanum*. Imbalances in nitrogen and carbon assimilations caused by deficiency of the ferredoxin-dependent enzyme.
Plant Physiology, 120(1):33–41, 1999.

[544] Mario García-Domínguez, José C. Reyes, and Francisco J. Florencio.
NtcA represses transcription of *gifA* and *gifB*, genes that encode inhibitors of glutamine synthetase type I from *Synechocystis* sp. PCC 6803.
Molecular Microbiology, 35(5):1192–1201, 2000.

[545] L. M. Rubio, A. Herrero, and E. Flores.
A cyanobacterial *narB* gene encodes a ferredoxin-dependent nitrate reductase.
Plant Molecular Biology, 30(4):845–850, 1996.

[546] I. Suzuki, H. Kikuchi, S. Nakanishi, Y. Fujita, T. Sugiyama, and T. Omata.
A novel nitrite reductase gene from the cyanobacterium *Plectonema boryanum.*
Journal of Bacteriology, 177(21):6137–6143, 1995.

[547] Belén Brito, Rosa Isabel Prieto, Ezequiel Cabrera, Marie Andrée Mandrand-Berthelot, Juan Imperial, Tomás Ruiz-Argüeso, and José Manuel Palacios.
Rhizobium leguminosarum hupE encodes a nickel transporter required for hydrogenase activity.
Journal of Bacteriology, 192(4):925–935, 2010.

[548] Masahiro Kasahara and Masayuki Ohmori.
Carbamoyl Phosphate Synthetase of the Cyanobacterium *Anabaena cylindrica.*
Plant and Cell Physiology, 38(6):734–739, 1997.

[549] Ramon Terrado, Adam Monier, Robyn Edgar, and Connie Lovejoy.
Diversity of nitrogen assimilation pathways among microbial photosynthetic eukaryotes.
Journal of Phycology, 51(3):490–506, 2015.

[550] M. E. Mulligan and R. Haselkorn.
Nitrogen fixation (*nif*) genes of the cyanobacterium *Anabaena* species strain PCC 7120. The *nifB-fdxN-nifS-nifU* operon.
The Journal of biological chemistry, 264(32):19200–19207, 1989.

[551] Shin Ichi Maeda and Tatsuo Omata.
A Novel Gene (*narM*) Required for Expression of Nitrate Reductase Activity in the Cyanobacterium *Synechococcus elongatus* Strain PCC7942.

Journal of Bacteriology, 186(7):2107–2114, 2004.

[552] Silvia Picossi, Ana Valladares, Enrique Flores, and Antonia Herrero.
Nitrogen-regulated genes for the metabolism of cyanophycin, a bacterial nitrogen reserve polymer: expression and mutational analysis of two cyanophycin synthetase and cyanophycinase gene clusters in heterocyst-forming cyanobacterium *Anabaena* sp. PCC 7120.
Journal of Biological Chemistry, 279(12):11582–11592, 2004.

[553] Anne Karradt, Johanna Sobanski, Jens Mattow, Wolfgang Lockau, and Kerstin Baier.
NblA, a key protein of phycobilisome degradation, interacts with ClpC, a HSP100 chaperone partner of a cyanobacterial Clp protease.
Journal of Biological Chemistry, 283(47):32394–32403, 2008.

[554] Peter J. Lammers and Robert Haselkorn.
Sequence of the *nifD* gene coding for the α subunit of dinitrogenase from the cyanobacterium Anabaena .
Proceedings of the National Academy of Sciences, 80(15):4723–4727, 1983.

[555] Barbara J Mazur, Douglas Rice, and Robert Haselkorn.
Identification of blue-green algal nitrogen fixation genes by using heterologous DNA hybridization probes.
Proceedings of the National Academy of Sciences, 77(1):186–190, 1980.

[556] Walter Arnold, Andreas Rump, Werner Klipp, Ursula B. Priefer, and Alfred Pühler.
Nucleotide sequence of a 24,206-base-pair DNA fragment carrying the entire nitrogen fixation gene cluster of *Klebsiella pneumoniae.*
Journal of Molecular Biology, 203(3):715–738, 1988.

[557] Tan Chi Huang, Rong Fong Lin, Mu Kuei Chu, and Horng Ming Chen.
Organization and expression of nitrogen-fixation genes in the aerobic nitrogen-fixing unicellular cyanobacterium *Synechococcus* sp. Strain RF-1.
Microbiology, 145(3):743–753, 1999.

[558] Olaf Stricker, Bernd Masepohl, Werner Klipp, and Herbert Böhme.
Identification and characterization of the *nifV-nifZ-nifT* gene region from the filamentous cyanobacterium *Anabaena* sp. strain PCC 7120.
Journal of Bacteriology, 179(9):2930–2937, 1997.

[559] Ji-Won Hong, Han-Gu Choi, Sung-Ho Kang, and Ho-Sung Yoon.
Axenic purification and cultivation of an Arctic cyanobacterium, *Nodularia spumigena* KNUA005, with cold tolerance potential for sustainable production of algae-based biofuel.
Algae, 25(2):99–104, 2010.

[560] Guennadi Sezonov, Danièle Joseleau-Petit, and Richard D'Ari.
Escherichia coli physiology in Luria-Bertani broth.
Journal of Bacteriology, 189(23):8746–8749, 2007.

[561] D. J. Reasoner and E. E. Geldreich.
A new medium for the enumeration and subculture of bacteria from potable water.
Applied and Environmental Microbiology, 49(1):1–7, 1985.

[562] Suqin Gao, Yun Kong, Jing Yu, Lihong Miao, Lipeng Ji, Lirong Song, and Chi Zeng.
Isolation of axenic cyanobacterium and the promoting effect of associated bacterium on axenic cyanobacterium.
BMC Biotechnology, 20(1):1–11, 2020.