Author:
Hopton, Cassie M

Title:
Genomic analysis of astrobiology-relevant adaptations to low light in far-red light utilising cyanobacteria

Genomic analysis of astrobiology-relevant adaptations to low light in far-red light utilising cyanobacteria

Cassie Hopton

School of Geographical Sciences
University of Bristol
February 2023

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Geography (MSc) (R) in the Faculty of Science, School of Geographical Sciences

Word Count: 22,135

Abstract

Cyanobacteria have been utilised as astrobiology models and show a promise as tools that could support growth of life and production of materials on other planets. Recently, a small subset of cyanobacteria have been found to be capable of using far-red light for photosynthesis, as mediated by two known processes: far-red light photoacclimation (FARLIP) and low-light-photoacclimation (LOLIP). Due to its penetrative properties, cyanobacteria capable of harvesting far-red light can survive in low light conditions. Such extremophile cyanobacteria may be more suitable candidates within applications beyond Earth, or could be used as analogues to study the potential for life on planets receiving low luminosity. However, the genomics, evolution and application of these far-red utilising cyanobacteria are still unclear. Through an extensive genomics analysis with key marker genes, this study has identified 102 strains of cyanobacteria containing FARLIP and/or LOLIP genes. Comparative genomics of FARLIP clusters from five new genera not previously reported to be FARLIP-capable reveals conservation of a unique gene cluster containing 20 genes, as reported for other genera, with some minimal alterations. Likewise, comparative genomics of cyanobacteria containing LOLIP clusters revealed consistent appearance of *apcD4* and *apcB3* genes with variable appearance of *LHCB* and/or *isiX* genes. BLAST analysis also revealed several low light tolerant cyanobacteria exhibit genes necessary for providing bioavailable organics, hydrogen production and cold temperature survival that would be necessary in astrobiological life support systems or for use as analogue organisms. Through the expanded repertoire of low light tolerant cyanobacteria, confirmation of genetic conservations and exploration of astrobiological significance, we hope this study introduces the use of low light tolerant cyanobacteria in astrobiology.

Acknowledgments

I would like to express my great appreciation to Dr Patricia Sanchez-Baracaldo for enthusiastically helping me design an astrobiology-focussed project, as well as her guidance, mentoring and invaluable support provided this past year. I would also like to offer my special thanks to Jo Boden and Giorgio Bianchini for sharing their phylogenetic and computational knowledge, as well as Elliot Druce, Cathrine Bayliss, Thomas Turpin-Jelfs and Dr Christopher Williamson for sharing their cyanobacteria and laboratory knowledge. I would not have been able to complete my work without their help and support.

Author 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: 09th February 2023

List of Abbreviations

Chl	Chlorophyll
FARLIP	Far-red light photoacclimation
Fd	Ferredoxin
FNR	Ferredoxin—NADP(+) reductase
FRL	Far-red light
LHCB	Light harvesting chlorophyll binding protein
LLAC	Low-light adapted cyanobacteria
LOLIP	Low-light photoacclimation
MAD	Minimal ancestor deviation
OEC	Oxygen-evolving centre
PAR	Photosynthetically active radiation
Pheo-a	Pheophytin-a
Pc	Plastocyanin
PQ	Plastoquinone
PQH•	Plastosemiquinone
PQH_2	Plastoquinol
PSI	Photosystem I
PSII	Photosystem II

List of Figures

List of Tables

1. Introduction

Cyanobacteria represent one of the most influential organisms on the development of modern life on Earth. The appearance of cyanobacteria ancestors (3.6-3.8 billion years ago, Archean period) was followed by a rise in oxygen (Sánchez-Baracaldo and Cardona, 2020). This lead to the Great Oxidation Event (2.45-2.32 billion years ago, the Paleoproterozoic era), the development of multicellular eukaryotes and the formation of the ozone layer 600 million years ago (Schopf, 1993; Duarte, 2012; Schirrmeister *et al.*, 2013).

Cyanobacteria are known as ecological pioneers, and this is reflected in a ubiquitous global distribution and an occupation of various extremes. Hypersaline (Kirkwood *et al.*, 2008; Caumette *et al.*, 1994; Voß *et al.*, 2013), frigid (Nadeau and Castenholz, 2000; Mueller *et al.*, 2005; Quesada and Vincent, 2012), thermal (Pentecost, 2003; Kaštovský *et al.*, 2014; Strunecký *et al.*, 2019) and desiccated (Tiwari *et al.*, 2005; Wierzchos, Ascaso and McKay, 2006; Azua-Bustos *et al.*, 2014) environments have been colonised by cyanobacteria. Cyanobacteria have also been found to tolerate Earth and Martian ionizing and UV radiation (Cockell *et al.*, 2005; Sinha and Häder, 2007; Billi *et al.*, 2013, 2019), as well as occupy the deep sub-surface (Puente-Sánchez *et al.*, 2018).

A surprising recent discovery was cyanobacteria that could pioneer into low light conditions. These conditions are defined by a limited access to visible light in the range of 400 – 700 nm that is typically used by most photosynthetic organisms. Cyanobacterial isolates have been retrieved from such low light conditions including underneath plant canopy, within crowded microbial mats and underneath invertebrates (Kühl *et al.*, 2005; Ohkubo and Miyashita, 2017; Ho and Bryant, 2020). These low-light extremophiles survive by instead capturing far-red light (FRL) (>700 nm). FRL is scarcely utilized by other organisms and has been proven to penetrate more deeply into microbial mats and soils (Bliss and Smith, 1985; M. Kühl & Fenchel, 2000; Pierson et al., 1990). Through rearrangement of the photosystem by low-light photoacclimation (LOLIP) and/or FRL photoacclimation (FARLIP), cyanobacteria can negate competition for light access. This enables survival in darker conditions than thought possible (Gan *et al.*, 2014; Gan, Shen and Bryant,

2014; Antonaru *et al.*, 2020). Cyanobacteria capable of either LOLIP, FARLIP, or both, will be referred to as low-light adapted cyanobacteria (LLAC) from here forward.

The discovery of LLAC expands the known limits of photosynthesis on Earth, and has applications to life beyond Earth. Cyanobacteria have supported the exploration for potential life by research of the genetic, molecular and ecological factors that enable survival in extremes. For example, in extreme cold (Chrismas, Anesio and Sánchez-Baracaldo, 2015; Chrismas *et al.*, 2016a, 2018) or desiccated environments (Baqué, Viaggiu, *et al.*, 2013; Azua-Bustos *et al.*, 2014; Urrejola *et al.*, 2019). Such insights have helped us understand the potential for habitability within icy moons such as Europa, Titan or Enceladus (Vance *et al.*, 2021), and Mars (Flaim *et al.*, 2014; Baqué *et al.*, 2016; Lalić *et al.*, 2020). LLAC present a unique addition to cyanobacteria as astrobiological analogues; LLAC may be used as research models for organisms on other planets that receive low luminosity, such as the terrestrial planets proximal to near infrared-emitting dwarf star TRAPPIST-1 (Gillon *et al.*, 2016, 2017).

LLAC could also have more practical applications in astrobiology. Due to oxygen production, nitrogen- and carbon-fixation capabilities, cyanobacteria have been proposed as tools to help build a breathable atmosphere on other planets and grow heterotrophic bacteria/plants for the production of drugs, food or biomaterials (Verseux *et al.*, 2016). Cyanobacteria have also been postulated as a source of hydrogen or other materials needed to produce biofuels, such is the case for *Synechococcus* sp. PCC 7002 (Carr, 2019; Desai, 2015). These functionalities would be employed in life support systems known as bioregenerative life support systems (BLSS) (Verseux *et al.*, 2016). While FRL is not as efficient for photosynthesis (Mascoli, Bersanini and Croce, 2020), use of LLAC in colonisation would confer additional advantages over use of standard light adapted cyanobacteria. LLAC could survive and continue to perform desired functionalities despite lack of access to light; this may occur when placed under shielding from harsh and UV-intense planetary surfaces or when in crowded microbial production chambers.

However, the potential applications of LLAC cannot currently be realised as the fundamental genetic and molecular factors that influence low-light survival are not yet fully understood. The scope of FARLIP and LOLIP-capable species, their diversity and genomics still requires investigation. For this reason, this study aims to (i) understand the diversity and geographic and ecological distribution of LLAC, (ii) understand the evolution of FARLIP and LOLIP through stringent Bayesian approach, (iii) understand the conservation of FARLIP and LOLIP gene clusters among LLAC identified, and (iv), for the first time, consider the significance and utilization of LLAC in astrobiology. Through this, we hope to introduce LLAC as potential astrobiology models and tools, understand the diversity of LLAC available, and clarify the genetics of FARLIP and LOLIP processes which would be necessary for genetic engineering and optimisation in astrobiology.

2. Literature Review

2.1 The photosynthetic process in cyanobacteria

Cyanobacteria are among the oldest creatures on Earth; microfossils aged at 1.9 billion years have been attributed to cyanobacteria (Wacey *et al.*, 2013). The last common ancestor for cyanobacteria evolved around 3 billion years ago, yet the lineage of organisms capable of oxygenic photosynthesis evolved much earlier (Schirrmeister, Sanchez-Baracaldo and Wacey, 2016; Sánchez-Baracaldo and Cardona, 2020; Oliver *et al.*, 2021). Oxygenic photosynthesis describes the process of using both light-dependent and light-independent reactions to capture visible light energy in the range of 450 – 660 nm. This energy is used to reduce carbon dioxide to carbohydrates with release of oxygen, thus converting light energy to chemical energy (Scherer, Almon and Böger, 1988; Vermaas, 2001; Srivastava, Rai and Neilan, 2013; Artur and Min, 2020). This highly specific process provides cyanobacteria with vital sugars, NADPH and ATP molecules, and relies on several systems that are explained in detail below.

2.1.1 Light harvesting

The utilization of visible light begins with light harvesting antenna systems. One such system, phycobilisomes, are hemidiscoidal structures anchored to the outer cyanobacterial thylakoid membrane. These phycobilisomes contain stacks of light harvesting pigments or antenna, such as phycoerythrin, phycocyanin, phycoerythrocyanin, allophycocyanin and chlorophyll a (Chl a) (Fig. 1) (Glazer, 1985; MacColl, 1998; Govindjee and Shevela, 2011). Each of these pigments absorb and emit energy such that energy is migrated towards the reaction centre in photosystem II (PSII). Upon irradiance, photon resonance energy is absorbed and transferred unidirectionally along the light harvesting pigments as permitted by thermodynamic laws; energy is flowed from highest potential energy to lowest potential energy pigments as follows phycoerythrin → phycocyanin → allophycocyanin → Chl a. Absorbance and fluorescence emission maxima of phycobilisome pigments are shown in Table 1.

Energy therefore flows from phycoerythrin at the periphery of the phycobilisome to allophycocyanin at the core with Chl a in the thylakoid membrane adjacent to PSII. The core consists of two allophycocyanin with a 650 nm absorbance maximum and two allophycocyanin with a lower maximum, known as the L_{cm} and α^{B} polypeptide, (MacColl, 1998). Energy transcends from the higher energy to lower energy allophycocyanin to a Chl a pair known as P680 at the reaction centre of PSII (Gantt and Lipschultz, 1973; Glazer, 1985; MacColl, 1998). It should also be noted the content of phycobilisomes varies between strains (Basheva *et al.*, 2018), and can diverge. For example, it is apparent no strains synthesize both phycoerythrin and phycoerythrocyanin (Bryant, 1982).

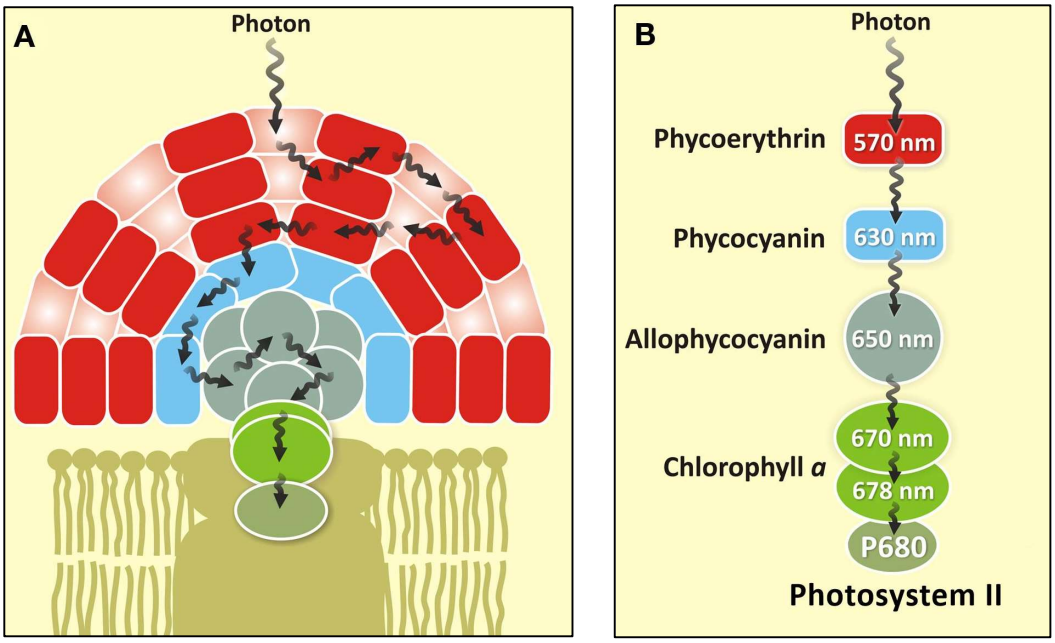


Figure 1 | Architecture of phycobilisomes. (A) Phycobilisomes consist of rod-like arrays and light harvesting phycobiliproteins, such as phycoerythrin, phycocyanin, phycoerythrocyanin, allophycocyanin, and molecules of chlorophyll. These bound at the outer thylakoid membrane but connect to the stromal side of PSII (MacColl, 1998; Hankamer *et al.*, 2001). (B) photon energy transduction occurs from phycoerythrin to Chl a in the reaction centre of PSII. Schematic modified from (Govindjee and Shevela, 2011).

Another light harvesting system employed by cyanobacteria is Chl-binding systems. As opposed to membrane-bound phycobilisomes, these light harvesting structures are comprised of transmembrane Chl-binding proteins, such as Pcb and IsiA, that

bind all available Chl – the type of Chl bound is species-dependent (Artur and Min, 2020). For example, the marine picocyanobacteria *Prochlorococcus marina* sp. binds both Chl a2 and b2, while *Acaryochloris marina* sp predominately binds Chl d (Chen, Zhang and Blankenship, 2008). Chl-binding proteins may associate extrinsically with PSI (Jordan *et al.*, 2001).

Table 1 | Absorbance and emission maxima of cyanobacteria phycobilisome proteins

Pigment	**Absorption λmax (nm)**	**Fluorescence λmax (nm)**	**Reference**
C-phycoerythrin	560>675	577	(Zickendraht-Wendelstadt, Friedrich and Rüdiger, 1980)
b-phycoerythrin	545>563	570	(Macdowall, Bednar and Rosenberg, 1968; Gantt and Lipschultz, 1973)
B-phycoerythrin	545>563>498	575	(Macdowall, Bednar and Rosenberg, 1968; Gantt and Lipschultz, 1973)
R-phycoerythrin	565>540>498	578	(Macdowall, Bednar and Rosenberg, 1968)
phycoerythrocyanin	570>595	625	(Bryant, Glazer and Eiserling,

			1976; Glazer, 1976)
C-phycocyanin	620	640	(Gantt and Lipschultz, 1973)
R-phycocyanin	617>555	636	(Glazer and Bryant, 1975; Glazer and Hixson, 1975)
Allophycocyanin	650	660	(Gantt and Lipschultz, 1973; Glazer and Bryant, 1975)
Allophycocyanin B	671>618	675	(Glazer and Bryant, 1975; Ley *et al.*, 1977)

2.1.2 Photosystem II and cytochrome b6f

Although carotenoids are thought to aid the light harvesting process (Berera *et al.*, 2010; Govindjee and Shevela, 2011), evidence only indicates a regulatory role supporting photosynthesis by providing photoprotective activities (Bailey and Grossman, 2008; Kusama *et al.*, 2014; Sedoud *et al.*, 2014) and supporting functional PSII reaction centres (Sozer *et al.*, 2010; Tóth *et al.*, 2015; Zakar *et al.*, 2016). PSII consists of twenty subunits, as well as 35 Chl a, two pheophytin-a (Pheo-a), two heme, 12 β-carotene and three plastoquinone (PQ) molecules, in addition to a number of lipids and ions (Fig. 2A) (Guskov *et al.*, 2009). The PSII is a dimeric structure formed of D1 and D2 proteins, known as PsbA and PsbD, respectively, within the thylakoid membrane. Chl a-binding antenna subunits CP43 and CP47, also known as PsbC and PsbB, surround the dimers and bind 16 and 13 molecules of Chl a, respectively (Gabdulkhakov and Dontsova, 2013). Several intrinsic proteins

subunits (PsbE, PsbF, PsbH-M, PsbN, PsbX, PsbY, PsbZ, and PsbYcf12) and extrinsic protein subunits (PsbO, PsbU, and PsbV) also contribute to the structure (Gabdulkhakov and Dontsova, 2013; Gao *et al.*, 2018).

The D1 and D2 heterodimers each consist of five α-helices and bound at the interface by four chlorophyll a (PD1, PD2, Chl_{D1} and Chl_{D2}) and two Pheo-a (pheoD1 and pheoD2) that constitute the reaction centre. Photon resonance energy is transferred from the low energy allophycocyanin to primary electron donors, the Chl a dimers PD1 and PD2, denoted P680. P680 is excited to the singlet excited state P680* that acts as an electron donor to an adjacent electron acceptor, pheoD1, the primary electron acceptor. It should be noted Chl_{D1} may also act as the primary electron donor and reduce pheoD1 (Holzwarth *et al.*, 2006; Keisuke Kawashima and Hiroshi Ishikita, 2018).

An oxygen-evolving centre (OEC) consisting of four manganese ions and a divalent calcium ion catalyse water splitting; two molecules of water, the terminal electron donor, are oxidised to molecular oxygen and four protons, which are released into the lumen. Electrons from water are transferred to a D1 tyrosine residue, Y_Z, which re-reduces P680* to the singlet ground state. Formation of molecular oxygen requires the manganese cluster to cycle through five oxidation states, S0-S4, as stimulated by four photons. S1 is generated by oxidation and proton loss from S0 manganese, induced by initial P680 photon excitation. S2 is generated from a second oxidation, S3 from a third oxidation and loss of a photon, and S4 from a fourth oxidation. In the recovery from S4 to S0, two protons and molecular oxygen are released and two water molecules bound. Electrons are successively passed from pheoD1 to a D2-bound PQ at site Q_A, followed by transfer to non-heme iron Fe^{2+} and finally a D1-bound PQ at site Q_B. From excitation of P680 by two photons, two electrons and two stroma protons are accepted at Q_B, and a second Q_B following excitation by a further two photons; as a result, in total two plastoquinol (PQH_2) are individually formed and shuttled to the cytochrome b6f complex (Artur and Min, 2020).

The cytochrome b6f complex is a dimer of eight subunits (Fig. 2B). PQH_2 binds to the positive lumen region of the complex and reduces a Rieske [2Fe-2s] centre, while

also releasing two protons into the thylakoid lumen, becoming oxidised to plastosemiquinone (PQH•). Electron transfer from PQH_2 also occurs to cytochrome f, containing a c-type cytochrome, which subsequently oxidises a membrane-bound, copper containing monomer plastocyanin (Pc). Semiquinone reduces heme b_p of cytochrome b6, becoming PQ, which in turn reduces heme b_n. Heme b_n reduces an adjacent PQ to PQH•. In a second round of PQH_2 binding, electron transfer and proton influx from stroma fully oxidises PQH• to PQ, and a second Pc . Plastocyanin is released and shuttled to photosystem I (Baniulis *et al.*, 2008).

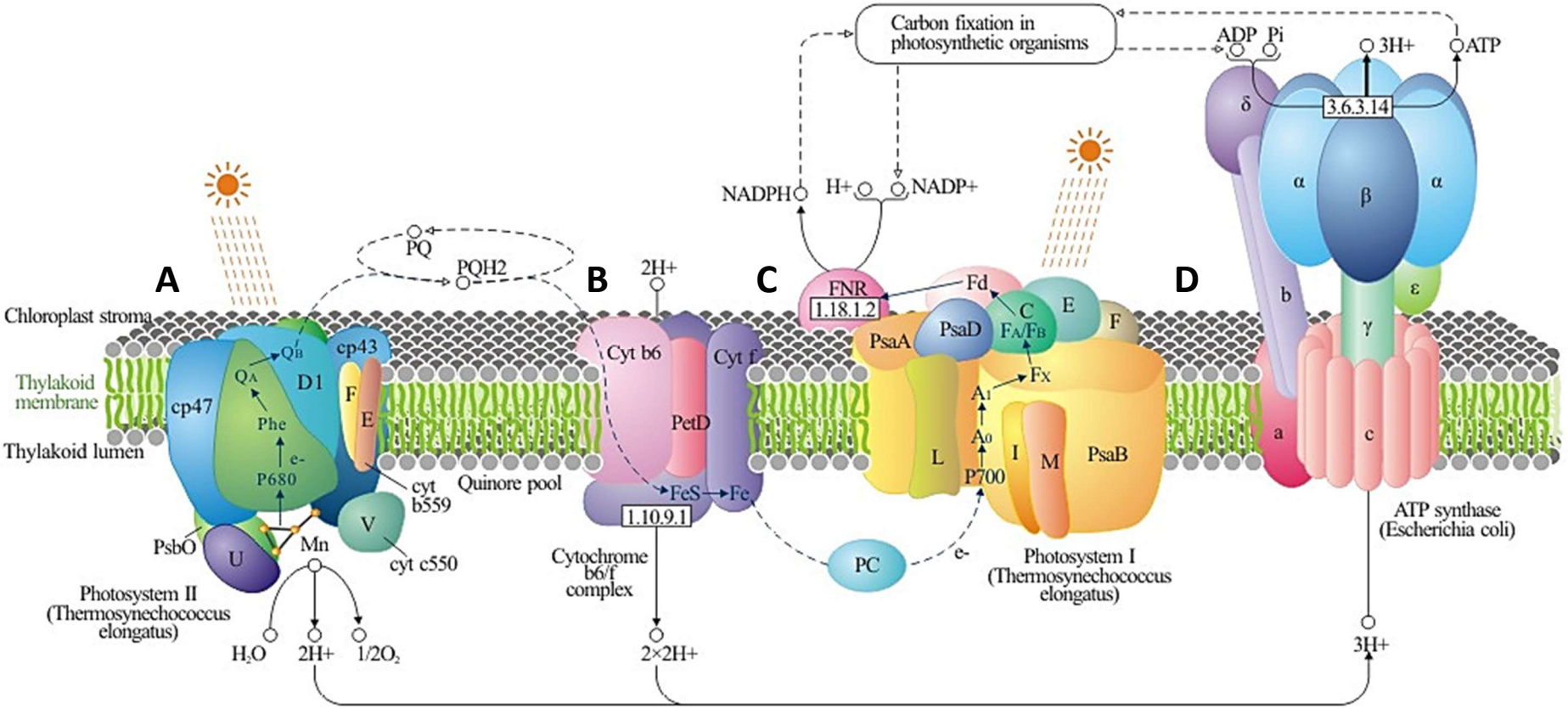


Figure 2 | Cyanobacterial photosynthetic apparatus. Schematic illustration of the (A) PSII, (B) Cytochrome b_6f (C) PSI and (D) ATP synthase. These structures are based on those derived from *Thermosynechococcus elongatus* and *Escherichia coli.* Figure is taken from PhytoAb (https://www.phytoab.com/). (A) PSII utilises photon energy to extract electrons from water molecules and translocate hydrogen ions into the thylakoid lumen. (B) Cytochrome b_6f transfers electrons from plastoquinone (PQ) to plastocyanin (PC), forming plastoquinol (PQH_2) and translocating further hydrogen ions into the thylakoid lumen. (C) PSI accepts and transfers electrons from PC to ferredoxin (Fd), which subsequently activates the formation of NADPH via a ferredoxin—NADP(+) reductase

(FNR). (D) Transfer of hydrogen ions into the thylakoid lumen allow the establishment of a proton motive force. This force is utilised by the ATP synthase to catalyse the formation of ATP.

2.1.3 Photosystem I, ATPase and the Calvin cycle

Photosystem I is a multi-subunit complex consisting of subunits PsaA-X (Fig. 2C). The reaction centre core is formed by subunits PsaA and PsaB, which have N-terminal helices that show similarity to PSII CP43 and CP47 (Jordan *et al.*, 2001). The PSI light harvesting antenna is less well defined; although, PSI has been shown to be associated with unique phycobilisomes (Liu *et al.*, 2013; Watanabe *et al.*, 2014). These phycobilisomes contain phycocyanin molecules but are absent in allophycocyanin.

Photon excitation energy is harnessed via antenna. Within the C terminal domain of PsaA and B, photon energy excites chlorophyll a and a' dimer P_A and P_B, denoted P700, to the singlet excited state, P700*. P700* reduces A, a Chl a, which initiates an electron transfer chain to Chl a A0, phylloquinone A1 and then iron sulphur cluster [4Fe-4S]-FX. [4Fe-4S]-FX reduces [4Fe-4S]-FA which in turn reduces [4Fe-4S]-FB. Pc docks at the PSI complex, and acts as an electron donor, re-reducing P700* to the singlet ground state. [4Fe-4S]-FB reduces a water soluble, iron-sulphur protein, ferredoxin, bound to the stroma side of the complex. Ferredoxin transfers the electron to ferredoxin-$NADP^+$ reductase, which in turn reduces terminal electron acceptor $NADP^+$. Two electron transfer events from ferredoxin to $NADP^+$ produces NADPH, utilising a proton as an additional substrate. In total, four photons produce the two NADPH required for the Calvin cycle (Grotjohann and Fromme, 2005; Artur and Min, 2020).

The translocation of protons into the thylakoid lumen as a result of electron transfer establishes a proton concentration gradient, which is subsequently utilised by a thylakoid embedded F_0F_1-ATPase (Fig. 2D). The F_0F_1-ATPase consists of a three subunit hydrophobic proton channel, F_0, and a hydrophilic ATPase, F_1, that protrudes from the thylakoid membrane. The ATPase of F_1 is formed by three αβ heterodimers that contain six nucleotide binding sites. Binding of ADP initiates the action of catalysis and rotation, where inorganic phosphate (Pi) is fused with ADP to form ATP. Binding of an additional ADP results in release of the previously formed ATP molecule, and as such the cycle continues. The F_1 subunit γ acts to couple ATP synthesis to proton translocation; subunit γ protrudes between the hexaheteromeric

αβ subunits, but is also linked to the rotating F_0 proton channel. In this channel, protons are taken up from the thylakoid lumen, and through a rotary mechanism transported to the thylakoid stroma. Subunit γ couples this rotation to rotation to the F_1 αβ subunits, inducing conformational changes of tightening and relaxation that allow the to formation and release of ATP into the stroma through proton motive force (Liu, 2016; Artur and Min, 2020).

The fixation of carbon dioxide into accessible organic compounds occurs during the Calvin cycle. This light-independent reaction consumes the NADPH and ATP released from the PSI and ATPase. A key enzyme in this cycle is ribulose-1,5-bisphosphate (RuBisCO); RuBisCO is activated by carbon dioxide binding near the catalytic site, and catalyses the first reaction of this cycle - addition of carbon dioxide to ribulose-1,5-bisphosphate, generating two molecules of 3-phosphoglyceric acid. The proceeding steps catalysed by phosphoglycerate kinase and glyceraldehyde-3-phosphate dehydrogenase, utilize two ATP and two NADPH, respectively (Gurrieri *et al.*, 2021). This process forms two molecules of glyceraldehyde 3-phosphate, and is followed by a ribulose-1,5-bisphosphate regeneration phase. Glyceraldehyde 3-phosphate can be used as a nutrient or to form monosaccharide sugars, such as glucose.

2.2 Divergence of key photosynthetic pigments and genes

2.2.1 Chlorophyll

In photosynthesis, chlorophylls contribute to the essential light harvesting and electron transfer events that enable efficient transferral of light energy to chemical energy. Typically, Chl a is employed for this role. Chl a consists of a nitrogen-bound Mg^{2+} ion encased in a pyrrole-derived chlorin which has a methyl group at position C7. Other forms of chlorophyll include: b, c_2, d and f. Chl b, Chl d and Chl f are similar in structure to Chl a; the molecules consist of a chlorin type ring but contain a formyl group at positions C7, C3 and C2, respectively. Chl c_2 instead consists of a porphyrin type ring structure with methyl and ethylenyl side chains. Absorption and fluorescence of each chlorophyll is given in Table 2 (Niedzwiedzki and Blankenship, 2010). It is clear Chl f is the most red-shifted chlorophyll, followed by Chl d. Chl f is

the most recent discovery within the chlorophyll family, having been detected in 2010 (Chen *et al.*, 2010).

Table 2 | Absorbance and emission maxima of cyanobacteria chlorophyll

	Absorption λmax (nm)	**Fluorescence λmax (nm)**
Chl a	443>671[1]	677[1]
Chl b	473>655[1]	662[1]
Chl c_2	456>588>634[1]	640[1]
Chl d	463>697[1]	705[1]
Chl f	706[2]	722[2]

[1] (Niedzwiedzki and Blankenship, 2010)

[2] (Chen *et al.*, 2010)

Each chlorophyll is structurally distinct (Fig. 3), and as such are synthesised by dedicated chlorophyll synthases. The most well studied chlorophyll synthase is encoded by the gene *chlg* (Shalygo *et al.*, 2009; Proctor *et al.*, 2018), which may synthesise both Chl a and Chl b. While the chlorophyll synthase for Chl f was identified in 2016 (Ho *et al.*, 2016), the chlorophyll synthase for Chl c and d is not yet known. Chl a is considered the most primordial chlorophyll and progenitor of alternative forms of chlorophyll as well as pheophytin. Indeed, these molecules have been shown to be readily produced from Chl a *in vitro* (Kobayashi *et al.*, 2013). Such a theory suggests there is a single cyanobacterial ancestor from which Chl a originated. Additionally, most cyanobacteria only employ Chl a for photosynthesis. Therefore, the use of chlorophylls alternative to Chl a must confer an advantage to the cyanobacterial photosynthetic process in order to be implemented and evolutionarily relevant.

A range of cyanobacteria have been documented to contain and utilize alternate chlorophyll. While the evolutionary relevance of Chl b is not well defined (Averina *et*

al., 2019), employment of Chl b enhances light harvesting in water environments where light penetration is deepest in the blue-green spectrum. *Prochlorococcus* can be adapted to high light or low light environments; in the case of low light, *Prochlorococcus* employ a higher ratio of Chl b to Chl to facilitate blue-green light harvesting in ocean water depths up to 200 m (Ralf and Repeta, 1992; Zinser *et al.*, 2007; Martiny *et al.*, 2009; Barrera-Rojas *et al.*, 2018). Likewise, the use of Chl b in photosystems is shared by other marine and freshwater cyanobacteria: *Synechocystis* sp. PCC 6803 (Vavilin *et al.*, 2003), *Prochlorothrix hollandica* (Van Der Staay, Yurkova and Green, 1998), *Prochloron didemni* (Hernández-Mariné, Turon and Catalan, 2019) and *Prochlorothrix scandica* (Pinevich, Velichko and Ivanikova, 2012).

Chlorophylls that fall under the c-type have been found more sparsely, and are speculated to be prominent in deep-water niches due to an enhancement for blue-light harvesting (Averina *et al.*, 2019). For example, marine-based *Prochloron didemni* has been identified as Chl c-containing (Larkum *et al.*, 1994). It should be noted these cyanobacteria do not exclusively utilize alternate forms of chlorophyll; for example, c-type chlorophyll account for 4-15% of the total chlorophyll in *Prochloron didemni,* while Chl a accounts for the remaining chlorophyll content (Larkum *et al.*, 1994). In contrast, use of Chl d and Chl f has been attributed to FRL harvesting. FRL has a wavelength of > 700 nm, that has been recorded to penetrate deep into microbial mats and soils (Bliss and Smith, 1985; M. Kühl & Fenchel, 2000; Pierson et al., 1990). Utilization of FRL therefore aids survival in low-light niches where visible light in the range of 400 – 700 nm is limited by geological or biological factors.

The diversity of cyanobacteria employing Chl d, Chl f, or both, is wide. The most extensively studied Chl d-containing organism, *Acaryochloris marina,* was discovered in 1996 within algae of *Lissoclinum patella.* Chl d in *Acaryochloris marina* accounts for more than 2% of the cell dry weight (Miyashita *et al.*, 1996). Many organisms, however, utilize both Chl d and Chl f. Examples include *Chroococcidiopsis thermalis* PCC 7203, *Leptolyngbya* sp. JSC-1, *Fischerella thermalis* PCC 7521, *Calothrix* sp. PCC 7507, *Chlorogloeopsis fritschii* PCC 6912 and *Chlorogloeopsis* sp. PCC 9212 (Averina *et al.*, 2018). The levels of Chl f comparative to other chlorophyll are light-dependent; in *Halomicronema*

hongdechloris, there are undetected levels of Chl f in visible light but the ratio of Chl f: Chl a shifts to 1:8 under red light (Chen *et al.*, 2012).

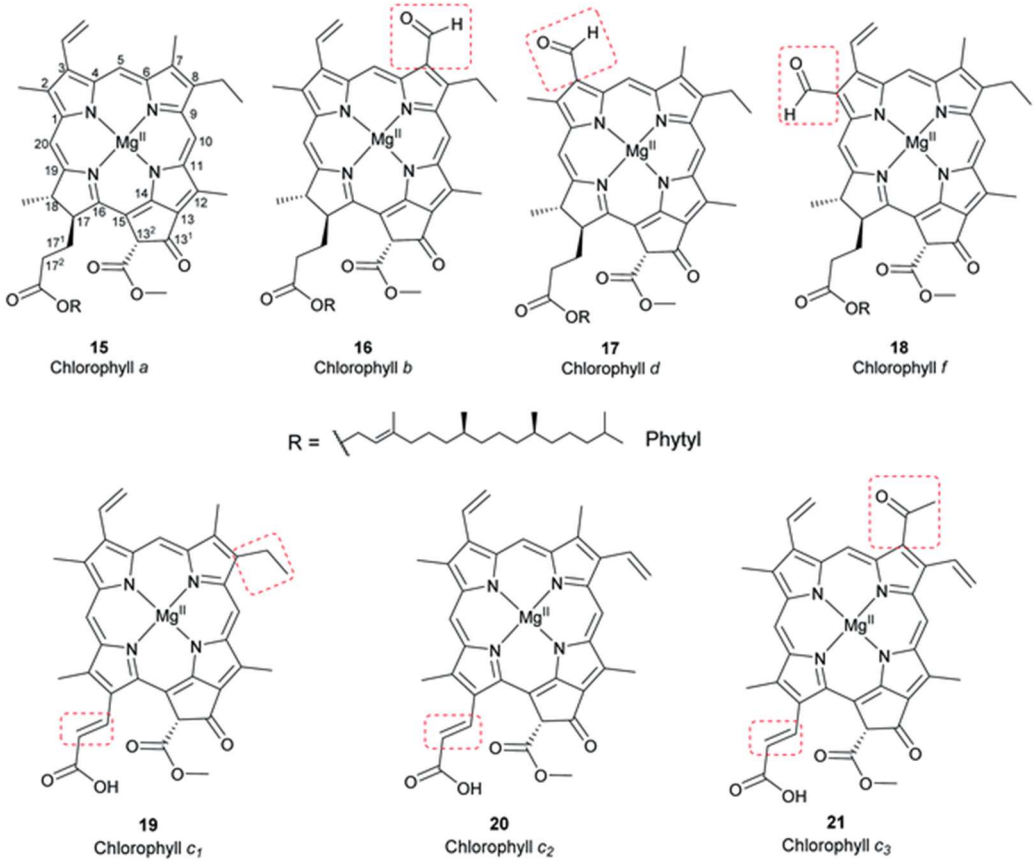


Figure 3 | Structural changes between Chl a, b, c, d and f. All chlorophyll feature a five membered cyclic ring fused to a porphyrin core structure. Chl b and f differ from Chl a by replacement of the C2 methyl with either a formyl group or the C3 vinyl group with a formyl group, respectively. The position of the formyl red-shifts absorbance. Chl c lack a phytyl tail at C17, and differ by side chains at C7 and C8 (Tahoun *et al.*, 2021). Figure taken from (Tahoun *et al.*, 2021).

2.2.2 PsbA

Photosynthesis is an intricate process that requires a multitude of factors to function and operate efficiently. The process was complicated further by the discovery of alternate forms of photosynthetic proteins encoded into cyanobacterial genomes; for example, quickly following the discovery of the D1-encoding *psbA* gene in *Synechocystis* sp. PCC 6803 (hereafter known as *psbA1*) (Jansson *et al.*, 1987), second and third *psbA* genes were found in *Synechocystis* sp. PCC 6803: *psbA2* (Ravnikar *et al.*, 1989) and *psbA3* (Metz, Nixon and Diner, 1990). As

aforementioned, PsbA is a D1 protein that forms the major structural component of PSII. Alternate forms could affect binding of reaction centres and associated components, and thus also the efficiency and mechanism underlying photosynthesis and the conversion of light energy and carbon dioxide into nutrients.

However, *psbA1* was found to be a silent gene that was not expressed for D1 formation. Although forced activation of this gene found that it encoded a functional D1 protein (Salih and Jansson, 1997), genes *psbA2 and psbA3* are dominantly expressed, with *psbA2* being attributed to more than 90% of expressed D1 in *Synechocystis* sp. PCC 6803 (Mohamed *et al.*, 1993; Tyystjärvi *et al.*, 1998). Additionally, comparison of genomic and nucleotide sequences revealed *psbA2* and *psbA3* encode almost identical proteins, with a 99.4% similarity in the coding regions (Metz, Nixon and Diner, 1990). Therefore, expression of either PsbA2 or PsbA3 would not significantly change the photosynthetic process.

While *Synechocystis* sp. PCC 6803 is a well-studied organism, these experiments only reflect expression of *psbA* variants in this strain. In fact, while common patterns of *psbA* genetics and transcription do occur, there are key differences that highlight the importance of avoiding model organisms to generalize information about the biology of organisms. For example, *Anacystis nidulans* R2 contains three *psbA* genes, with *psbA2* and *psbA3* producing identical transcripts, and *psbA1* differing by 25 residues. However, all these *psbA* genes produce transcripts and contribute to the photosynthetic D1 protein, with *psbA1* providing 94% of transcripts (Golden, Brusslan and Haselkorn, 1986). *Synechococcus* sp. PCC 7942 also displays a higher expression of *psbA1; psbA1* showed a 500- and 50-fold greater expression than *psbA2* and *psbA3,* respectively (Schaefer and Golden, 1989). The number of *psbA* genes is also not limited to three. While a fourth and fifth *psbA* gene has yet to be reported in *Synechocystis* sp. PCC 6803, such genes have been reported in *Gloeobacter violaceus* PCC 7421*, Synechococcus sp*. PCC 7942 and *Anabaena* sp. PCC 7120 (Mulo, Sicora and Aro, 2009). *Gloeobacter kilaueensis* JS-1 has also been found to encode six *psbA,* and *Leptolyngbia* sp. PCC 7375 and *Leptolyngbia* sp. Heron Island J have been found with seven and eight *psbA*, respectively (Cardona, Murray and Rutherford, 2015).

While many *psbA* genes are similar if not homologous, bioinformatical analysis has revealed highly divergent forms of *psbA* (Murray, 2012). These divergent forms of *psbA,* termed 'rogue' and 'super-rogue', contained variations in the OEC of the D1 protein required for the essential step of water-splitting. It has been established the consensus sequence at the OEC is DDEHEHDA. The 'rogue' *psbA* from *Acaryochloris marina* MBIC11017, *Crocosphaera watsonii* WH 8501, *Anabaena variabilis* ATCC 29413 and an assortment of *Synechococcus* and *Cyanothece* strains, showed the following OEC consensus: (D/E)(D/E/S)(D/A/R)H(A/S)H(T/L/V)(A/S). In comparison, the 'super-rogue' sequence from *Synechococcus* PCC 7335 showed divergence both in sequence and phylogenetically from typical *psbA* and 'rogue' *psbA* sequences. The 'super-rogue' could not be aligned with OEC sequences, lacked key residues and contained a variety of insertions. Such variation would encode a misfolded and non-functional OEC, and thus a non-functional D1 protein (Murray, 2012).

The variety of *psbA* available may have practical applications in nature. Transcript analysis in *Synechocystis* sp. PCC 6803 revealed, under high-light conditions, *psbA2* and *psbA3* produced high levels of transcripts while no transcripts for *psbA1* were detected (Mohamed and Jansson, 1989). Such light-induced effects were also observed in *Synechocystis* sp. PCC 6714 (Constant *et al.*, 1997). The increase in *psbA2* and *psbA3* transcripts within *Synechocystis* sp. PCC 6803 also appears to be UV-specific. UV-B irradiation increases the transcript level of *psbA2* and *psbA3* by 3 and 20-30 fold, respectively, while UV-A and visible light irradiation only induce a 2-3 fold increase in *psbA3* expression (Máté *et al.*, 1998). This suggests an evolutionary advantage to maintaining several variants of *psbA. Thermosynechococcus elongatus* BP-1 also contains just three forms of *psbA*, however, quantitative real time PCR analysis revealed *psbA1* expression replaces that of *psbA3* in lower-light conditions in up to 65% of PSII (Kós *et al.*, 2008; Loll *et al.*, 2008; Sander *et al.*, 2008). Expression is also UV-specific; UV-B decreases the level of *psbA1* transcripts in favour of *psbA3* (Kós *et al.*, 2008).

Induction of *psbA* may also be light-independent. Expressions assays have recently shown expression of *psbA1 in Synechocystis* sp. PCC 6803 can be induced by both microaerobic conditions and a reduced atmospheric carbon dioxide level (Chiş *et al.*,

2017). However, this is unlikely to be representative of the processes that determines *psbA1* expression in all cyanobacteria; *Synechocystis* sp. PCC 6803 contains carbon regulators in the promoter region of *psbA1* that is not apparent in other cyanobacteria, and therefore this is likely a niche approach to cyanobacterial survival. Nonetheless, this level of adaptation demonstrates each *psbA* may play a non-redundant role that allows cyanobacteria to pioneer environments that would be considered extreme for oxygenic phototrophs.

While 'super-rogue' *psbA* was thought to encode a redundant D1 protein, it is now clear the protein operates as a Chl f synthase (Ho *et al.*, 2016). 'Super-rogue' *psbA* extracted from *Chlorogloeopsis fritschii* sp. PCC 9212 and *Synechococcus* sp. PCC 7335, termed *psbA4*, appeared to synthesise Chl f upon heterologous expression. As expected, the expression of *psbA4* was also found to be FRL-dependent. It is therefore clear cyanobacteria with this divergent form of *psbA* have an evolutionary advantage of surviving in conditions where FRL is available.

2.3 FARLIP

The discovery of D1 proteins and chlorophyll highlights the need to continually explore the genomics, transcriptomics and proteomics of different organisms. Only a small amount of cyanobacteria appear to synthesise these components thus far. But, the discovery of Chl f and *psbA4* has redefined what we know of photosynthesis on Earth, and poses questions about the development of unique forms of photosynthesis beyond Earth.

The FARLIP process occurs analogous to the visible light photosynthetic process, however protein modifications in the photosynthetic apparatus red-shift energy absorbance and emission of light energy to the far-red spectrum. Protein replacement, additions and functions in FARLIP are outlined in Table 3. To summarise, the FARLIP response requires a set of approximately 20 genes which are encoded into a single cluster within the genome, known as the FARLIP cluster. The cluster of genes are as follows: *rfpA/B/C, apcA2/B2/D2/E2/D3, psbA3/D3/C2/B2/H2/A4, psaA2/B2/L2/I2/F2/J2* (Gan, Shen and Bryant, 2014).

Table 3 | FRL adapted proteins encoded in the FARLIP photosynthetic machinery [1]

	Protein Name	Gene Name	Function	Visible light photosynthesis paralog gene
1	RfpA	*rfpA*	Senses FRL, activates RfpB	-
2	RfpB	*rfpB*	Binds DNA and activates synthesis of FARLIP genes	-
3	RfpC	*rfpC*	Phosphate shuttle (predicted)	-
4	Allophycocyanin A2 / ApcA2	*apcA2*	FRL harvesting protein	*apcA1*
5	Allophycocyanin B2 / ApcB2	*apcB2*	FRL harvesting protein	*apcB1*
6	Allophycocyanin D2 / ApcD2	*apcD2*	FRL harvesting protein	*apcA1/D1*
7	Allophycocyanin D3 / ApcD3	*apcD3*	FRL harvesting protein	*apcA1/D1*
8	Allophycocyanin E2 / ApcE2	*apcE2*	FRL harvesting protein	*apcE1*
9	Photosystem II D1 subunit / PsbA3	*psbA3*	Light-activated electron extraction	*psbA1/A2*
10	Chlorophyll F synthase / PsbA4	*psbA4*	Chl f synthesis	-
11	CP47 / PsbB2	*psbB2*	Binds Chl a/d/f	*psbB1*
12	CP43 / PsbC2	*psbC2*	Binds Chl a/d/f	*psbC1*
13	Photosystem II D2 protein / PsbD3	*psbD3*	Light-activated electron extraction	*psbD1/D2*
14	Protein H / PsbH2	*psbH2*	Binds Chl a/d/f	*psbH1*
15	P700 chlorophyll a apoprotein A1 / PsbA2	*psaA2*	Binds primary electron donor of photosystem I	*psaA1*
16	P700 chlorophyll a apoprotein A2 / PsaB2	*psaB2*	Binds primary electron donor of photosystem I	*psaB1*

17	Photosystem I subunit 11 / PsaL2	*psaL2*	Photosystem I structural support	*psaL1*
18	Photosystem I subunit 8 / PsaI2	*psaI2*	Photosystem I structural support	*psaI1*
19	Photosystem I subunit 3 / PsaF2	*psaF2*	Supports electron transfer	*psaF1*
20	Photosystem I subunit 9 / PsaJ2	*psaJ2*	Photosystem I structural support	*psaJ1*

[1] Information taken from (Gan, Shen and Bryant, 2014)

Each gene in the FARLIP cluster encodes a modification to the light harvesting, PSII and PSI systems that allow the capture and transfer of red-shifted light energy, as well as some additional proteins that support activation of the FARLIP process. Notably, it appears only replacement of allophycocyanin is needed in the light harvesting system in order to allow the capture and transfer of FRL energy. Two major structural components of PSII, D1 and D2, are replaced by FARLIP counterparts. This change likely supports electron transfer using FRL photon energy.

The switch to Chl d and Chl f is also supported by production of the Chl f synthase and chlorophyll binding proteins that most likely have a structural synergy to Chl d and Chl f molecules. Likewise, the major structural components of PSI, A1 and A2, are also altered as well as proteins that provide structural and electron transfer support – these modifications will support transduction of FRL energy. Three additional proteins are produced that are unrelated to any gene found in visible light photosynthesis; *rfpa, rfpb* and *rfpc*. These genes, which encode photosensors and response regulators, appear to be unique to FARLIP-capable cyanobacteria and aid the signalling cascade that activates a switch from the visible light to FARLIP photosynthetic apparatus. The structure and mechanism of FARLIP in cyanobacteria as currently known is described in chapters below.

2.3.1 FARLIP activation

RfpA, RfpB and RfpC have no paralogs in cyanobacteria only capable of visible light photosynthesis. Deletion of these genes in FARLIP-capable cyanobacteria renders an inability to synthesise Chl f or genes in the FARLIP cluster, indicating a role in transcription activation of the FARLIP cluster (Zhao *et al.*, 2015). It has been established RfpA consists of a GAF domain, a far-red photoreceptor phytochrome domain, a molecular sensor/protein-protein interaction domain (PAS) and a histidine-kinase domain. RfpB contains a winged helix DNA-binding domain along with two CheY receiver domains, a family which mediates phosphate transfer (Stewart, 1997). RfpC is also a member of the CheY family (Gan *et al.*, 2014).

Based on these structures alone, the following model has been proposed and illustrated in Fig. 4: FRL activates the phytochrome domain of RfpA, which in turn activates or deactivates the RfpA histidine kinase activities. RfpA phosphorylates or halts phosphorylating RfpC at a key histidine residue. RfpC, either activate in a phosphorylated or dephosphorylated state, may shuttle the attached phosphate group to RfpB or remove an already placed phosphate group from RfpB – nonetheless, RfpB is activated by RfpC. Likely, conformational changes in RfpB occur upon attachment or detachment with the phosphate group that allow DNA binding to a unique region in the FARLIP cluster and activates transcription (Zhao *et al.*, 2015).

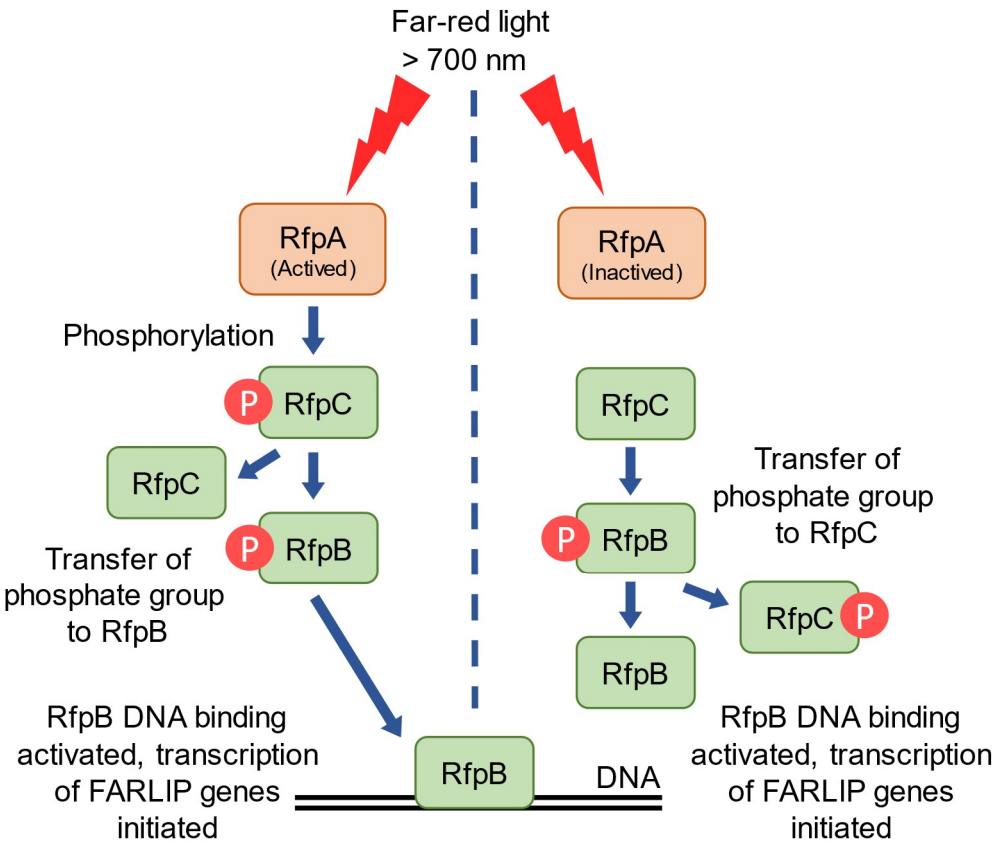


Figure 4 | Activation of FARLIP. RfpA, RfpB and RfpC undergo a cascade of reactions in response to FRL illumination. While mechanistic details are unclear, figure details two proposed models based on structural knowledge.

However, while a model has been proposed, the molecular mechanism or structural biology of this initial process is not yet known. Further studies determining the molecular role and function of these proteins are needed in order to determine the cascade of events that occur; for example, in contrary to the proposed model outlined above, RfpC may not be phosphorylated by RfpA, but instead RfpA may phosphorylate RfpB. The phosphate from RfpB may be removed by RfpC and transferred to an unknown target or removed by an unknown phosphatase. As clear above, it is also not known whether RfpB or RfpC is activated or inactivated by phosphorylation. It is also unknown how this response to FRL activates the synthesis of Chl f. Structural and mutational assays in FARLIP-capable cyanobacteria are needed to clarify these steps further.

2.3.2 FARLIP light harvesting

On the other hand, structural experiments have been completed for FRL-adapted phycobilisome. Phycobilisome components of FARLIP-capable *Leptolyngbya sp*. JSC-1, *Halomicronema hongdechloris* and *Synechococcus* sp. PCC 7335 have been determined. All these strains, once grown in FRL, replace ApcA1, ApcB1, ApcD1 and ApcE1 with ApcA2, ApcB2, ApcD2, ApcD3 and ApcE2. All phycocyanin in the visible light phycobilisome are retained (CpaA1, CpcB1, CpcA2, CpcB2, CpcH, CpcI, CpcD), but two phycoerythrin are present in the FRL phycobilisome that are not apparent in the visible light complex (CpeA, CpeB) (Herrera-Salgado *et al.*, 2018) (Gan *et al.*, 2014). However, these phycoerythrin are not detected in FRL-grown *Halomicronema hongdechloris* (Li *et al.*, 2016), indicating this is not a universal response.

The FRL-induced changes to the phycobilisomes are predominantly made to allophycocyanin. Such a change can account for the fact all FRL grown phycobilisomes studied thus far consist of a bicylindrical core as opposed to the visible light counterparts which are tri- or pentacylindrical cores (Gan *et al.*, 2014; Li *et al.*, 2016; Herrera-Salgado *et al.*, 2018). This is because while ApcE1 consists of four linker-repeat domains, ApcE2 has just two; as a result, ApcE2-containing phycobilisomes produce smaller, bi-structures instead of the larger tri- and penta-structures produced by ApcE1 (Gan *et al.*, 2014; Herrera-Salgado *et al.*, 2018). Notably, unlike *Leptolyngbya sp*. JSC-1, FRL grown phycobilisomes in *Synechococcus* PCC 7335 and *Halomicronema hongdechloris* also exhibit a lack of peripheral rods (Ho *et al.*, no date; Li *et al.*, 2016; Herrera-Salgado *et al.*, 2018) – the structural adaptation to FRL is therefore likely unique to each cyanobacteria, with exception of the presence of ApcE2, ApcA2 and ApcB2 and a bicylindrical core. These genetic and structural characteristics could therefore act as one potential biomarker for FARLIP.

The allophycocyanin changes in FRL-grown phycobilisomes also suggests the red-shifted light harvesting and energy transfer required for FARLIP is mediated predominately by allophycocyanin. Indeed, FRL-grown phycobilisome bicylindrical cores in *Synechococcus* PCC 7335 show absorption at 650 nm and 711 nm, and

emission at 730 nm, likely from ApcE2 and ApcD3. In contrast, red-light grown *Synechococcus* PCC 7335 phycobilisomes showed emission at 682 nm, likely from ApcE1 and ApcD1 (Ho *et al.*, 2017, 2020). FRL-specific allophycocyanin therefore act to emit energy that is sufficiently red-shifted to excite Chl f in PSII. However, further research is needed; the energy transfer process across multiple different FARLIP-capable strains is yet to be determined, and the structural placement of FRL-specific light harvesting molecules for energy transfer is still unclear and undefined in many FARLIP-capable strains. While great advances in knowledge have been made since the discovery of Chl f in 2010, understanding the basics of FARLIP light harvesting is required if we are to utilize FARLIP species and enhance the efficiency of this process for scientific and industry applications.

2.3.3 FARLIP PSII

Thus far, the PSII reaction centre of FARLIP-capable species has not been determined to an atomic level, and therefore the position and role of Chl d and Chl f in the reaction centre energy transfer cascade is unknown. But, photochemical and chromatography studies have provided insight into the composition and occurrence of red-shifted chlorophyll. For example, it has been deduced the PSII of FARLIP-capable *Chroococcidiopsis thermalis* contains 2 Pheo-a, 4 Chl f, 1 Chl d and 30 Chl a. Based on photochemical analysis, it appears the P_{D1} in *Chroococcidiopsis thermalis* remains a Chl a, however Chl d likely takes the position of Chl_{D1}. The following absorptions were measured for the Chl d and Chl f: 721, 727, 734, 737 and 739 nm. In *Acaryochloris marina*, Chl d is found at position Chl_{D1} with an absorption of 725 nm, and is also the primary electron donor (Shigeru Itoh *et al.*, 2007; Renger and Schlodder, 2008). The wavelength of 727 nm is similar and may indicate Chl_{D1} is a Chl d in *Chroococcidiopsis thermalis* (Nürnberg *et al.*, 2018). If it this is the case, it is likely the Chl f take on light harvesting roles, absorbing at 727 nm, 734 nm, 737 nm and 749 nm. However, this is just speculative.

Whether Chl_{D1} is also the primary electron donor in *Chroococcidiopsis thermalis* is unclear. For systems where Chl_{D1} is the primary electron donor (Holzwarth *et al.*, 2006; Keisuke Kawashima and Hiroshi Ishikita, 2018), the exchange of Chl a to Chl d or Chl f at this site is within reason in order to red-shift the electron transfer

process. Redox calculations also indicate, in typical systems, electron transfer from Chl_{D1} to $Pheo_{D1}$ is favourable while transfer from P_{D1} to chlorophyll is unfavourable (Keisuke Kawashima and Hiroshi Ishikita, 2018). Yet, circular dichroism has indicated P_{D2} as the likely primary electron donor in *Chroococcidiopsis thermalis* (Judd *et al.*, 2020). This discrepancy contrasts structural and energetic intuition, and solidifies the need for PSII reaction centre crystal structures from FARLIP-capable organisms in order to understand the FARLIP energy transfer.

Proteomic analyses have been provided for FARLIP PSII that has given insights into PSII remodelling for FARLIP. Aside from Chl f incorporation, the PSII of FRL-grown *Leptolyngbya* sp. JSC-1 also show replacement of PsbD1/2, PsbC1, PsbA2, PsbB1 and PsbH1 with FARLIP paralogs PsbD3, PsbC3, PsbA3/4, PsbB2 and PsbH2 (Ho *et al.*, 2020) (Fig. X). The PSII retains PsbX, PsbY, PsbE, PsbF, PsbJ, PsbK, PsbZ, PsbV, PsbU, PsbI, PsbO, PsbT, PsbL and PsbM (Ho *et al.*, 2020). These subunits play roles in PSII assembly, manganese and OEC stabilization, quinone binding at site Q_B, and maintaining the PSII-antenna interaction. It is therefore unfavourable to alter these proteins as they provide essential structural and function roles that do not pertain to light harvesting, red-shifted energy transfer or chlorophyll binding.

In contrast, as outlined in Table 3, the paralog proteins provide roles that predominantly focus on chlorophyll binding and electron transfer, aside from PsbA4 which functions as the chlorophyll f synthase. Although not confirmed through atomic crystal structures, presumably the paralogous PsbB2, PsbC2, PsbH2 have distinct structures that support the binding of Chl d and Chl f. PsbA3 and PsbD3, that form the D1 and D2 proteins, are also likely altered structurally to support binding of paralogous PsbB2, PsbC2, PsbH2 and a red-shifted reaction centre. The employment of these paralog proteins is therefore appropriate to form a red-shifted PSII.

However, the structural modifications that occur to PSII may represent a trade-off between survival and energy efficiency. The fluorescence lifetime of PSII is affected predominantly by two factors: (a) migration of energy from light harvesting antenna to reaction centres, and (b) trapping of the excitation energy by reaction centres. Time-resolved fluorescence measurements on FRL-grown *Chlorogloeopsis fritschii*

PCC 6912 PSII shows the quantum efficiency of FRL PSII is just 40%, while it is 70-80% for those grown in visible light. This is likely because Chl a has a 100 mV higher excited state energy than Chl f, and thus Chl f is comparatively a weaker electron donor. This may cause slower charge separation, which is involved in (a). Moreover, (b) could be slowed by Chl f that are not effectively situated to migrate energy towards the primary PSII donor (Mascoli, Bersanini and Croce, 2020). Both these possibilities could account for the reduced quantum efficiency of FARLIP PSII, although this theory would need to be confirmed by combined structural and biophysical assays.

2.3.4 FARLIP PSI

Similar structural, photochemical and efficiency assays have been conducted for FARLIP PSI. Within the FARLIP PSI of *Chroococcidiopsis thermalis*, chromatography reveals the presence of 7-8 chl f and 88-89 Chl a. No Chl d are detected. Based on structural, chemical and spectral assays, it is proposed chlorophylls occupy the following positions: P_A and P_B – Chl a; A – Chl f; and A0 – Chl a. Chl f at A is evident by a wavelength of 745 nm. The remaining Chl f show the following absorption: 745, 736, 756, 763 and 800 nm, and, if following the above model, act in FRL light harvesting. It is also proposed A-$_{1A}$ and A-$_{1B}$ act as the primary electron donors as opposed to P_A and P_B (P700) (Nürnberg *et al.*, 2018). However, this is again only speculative and would need confirmation with mechanistic assays. Proteomic assays show an equally extensive remodelling of the PSI, with FARLIP-capable *Leptolyngbya* sp. JSC-1 PSI showing replacement of PsaA1, PsaJ1, PsaF1, PsaB1, PsaL1 and PsaI1 with FARLIP counterparts PsaA2, PsaJ2, PsaF2, PsaB2, PsaL2 and PsaI2 (Fig. X).

The PSII appears to retain PsaC, PsaD, PsaE, PsaK, PsaX and PsaM (Ho *et al.*, 2020). As in PSII, the retained subunits act in structural and functional roles outside of chlorophyll binding and energy transfer; for example, PsaC acts as an apoprotein for [4Fe-4S]-FA and [4Fe-4S]-FB, PsaD forms a complex with ferredoxin and ferredoxin-oxidoreductase and PsaE stabilizes the interaction between PsaC and the PSI. Remodelling these subunits is therefore unnecessary for FARLIP. On the contrary, synthesis of paralogous PsaA2 and PsaB2, major subunits A1 and A2, is

required in order to bind the FRL adapted reaction centre and associated FRL subunits.

That being said, the roles of PsaJ1, PsaF1, PsaL1 and PsaI1 in cyanobacteria are not entirely clear. As a result, the role of the paralogous PsaJ2, PsaF2, PsaL2 and PsaI2 are also not well defined. While they are believed to have structural roles, the replacement of PsaJ1, PsaL1 and PsaI1 could indicate these subunits have functions in light harvesting and electron transport that are not yet clear. Deletion of *Synechocystis* PsaF did not affect $NADP^+$ reduction, indicating photosynthesis could complete without this subunit (Xu *et al.*, 1994). Feletion of PsaF and/or PsaE in *Synechococcus* sp. PCC 7002 lead to faster electron transfer from A_{-1} to [4Fe-4S]-FX and slower electron transfer from A_0 to A_1. It has therefore been proposed PsaF functions to stabilize an α-helical loop in the PsaA subunit that is adjacent to A_0, A_1, [4Fe-4S]-FA, [4Fe-4S]-FB and [4Fe-4S]-FX (Art van der Est *et al.*, 2004). PsaF2 could therefore be important for efficient A_0 to A_1 electron transfer in FARLIP systems. Deletion of *psaJ* in In *Synechocystis* led to a reduction in *psaF* mRNA and subsequently PsaF in membranes (Q *et al.*, 1994). PsaJ therefore acts to stabilize PsaF.

PsaL appears to enable trimerization of the PSI (Chitnis and Chitnis, 1993; Schluchter *et al.*, 1996), while PsaI stabilizes PsaL to the PSI (Xu *et al.*, 1995; Schluchter *et al.*, 1996); deletion of PsaI leads to an 80% decrease in PsaL level (Xu *et al.*, 1995). PsaI may also fixate PsaM to the PSI (Schluchter *et al.*, 1996). Many cyanobacteria exhibit trimeric PSI, although the significance is not entirely clear. Trimeric PSI have been found to offer photoprotective functions and support incorporation of longwave chlorophylls (Karapetyan, Holzwarth and Rögner, 1999; Karapetyan *et al.*, 2014). Indeed, PSI trimers in *Arthrospira platensis* exhibit chlorophylls with wavelengths of 708 and 740 – the far-red spectrum (Karapetyan *et al.*, 2014). Although this is not Chl f but a longwave Chl a (Shubin *et al.*, 1991; Shubin, Bezsmertnaya and Karapetyan, 1992), studies in both *Arthrospira platensis* and *Thermosynechococcus elongatus* show the trimeric form of PSI allows pigment-pigment interactions between chlorophyll that lead to the formation of longwave Chl a (Karapetyan *et al.*, 2014). Trimeric forms of PSI have been found in FARLIP-capable *Halomicronema hongdechloris* (Kato *et al.*, 2020)*, Fischerella thermalis*

PCC 7521 (Gisriel *et al.*, 2020) and FRL-utilizing *Acaryochloris marina* (Hamaguchi *et al.*, 2021).

Trimeric PSI may therefore offer advantages to FRL utilizing organisms, such as light harvesting or Chl d/f formation. Although FARLIP-capable *Chroococcidiopsis thermalis* PCC 7203 only exhibits monomeric and dimeric PSI (Li *et al.*, 2019), this was not studied under FRL growth. *Spirulina platensis,* for example, show higher emergences of red-shifted chlorophyll when in the trimeric state as opposed to the monomeric state (Gobets and Van Grondelle, 2001). It would be interesting to explore the relevance, if any, of trimeric PSI in FARLIP.

Despite extensive remodelling and Chl f incorporation, FARLIP PSI does not appear to suffer from degraded photosynthetic efficiency. The quantum efficiency of PSI in FRL-grown *Chlorogloeopsis fritschii* PCC 6912 is beyond 95%, despite charge separation being slowed 3.5-fold (Mascoli, Bersanini and Croce, 2020). PSI with Chl f show a less effective pigment connectivity, but can trap FRL wavelengths with high efficiency (Tros *et al.*, 2021). How the ATPase and Calvin cycle are affected by utilization of FRL is not yet known; however, given that PSI operates efficiently, it can be presumed these processes in turn are not severely effected by use of FRL. A simplistic schematic of the total FARLIP process is depicted in Fig. 5.

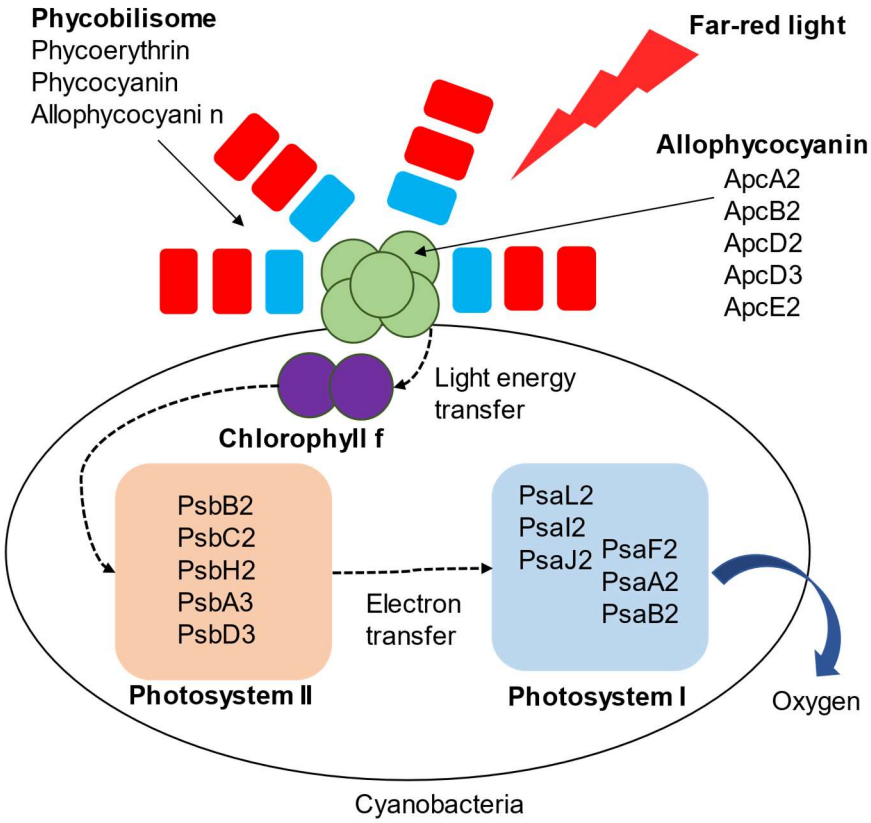


Figure 5 | The FARLIP process. FRL is absorbed by dedicated far-red absorbing pigments, such as ApcA2. This initiates an energy and electron transfer process as in the common form of photosynthesis, with an output of oxygen. However, some key proteins are replaced with FARLIP variants; these are noted within the figure. For clarity, Rfp proteins are not represented but would activate synthesis of the FARLIP proteins depicted.

2.3.5 Occurrence and distribution of FARLIP in cyanobacteria

Over 50 cyanobacteria have been identified as potentially FARLIP-capable, spanning diverse taxonomic groups and a range of ecological and geographical backgrounds (Gan, Shen and Bryant, 2014; Gan and Bryant, 2015; Antonaru *et al.*, 2020; Chen *et al.*, 2020). FARLIP appears to occur in a wide range of taxonomically diverse cyanobacterial groups. The taxonomic groups identified as having FARLIP-capable cyanobacteria are *Synechococcales, Chroococcidiopsidales, Pleurocapsales, Chroococcales* and *Nostocales*. Equally, the occurrence of FARLIP is also ecologically diverse. A large amount of FARLIP-capable cyanobacteria have been isolated from hot springs, however this is more likely a reflection of the large quantity sampling within these regions, particularly at Yellowstone National Park,

USA. Aside from this, the dispersal of FARLIP-capable cyanobacteria shows no distinguishable pattern; in a survey of 24 FARLIP cyanobacteria, the ecosystem distribution was as follows: terrestrial – 4, marine – 3, freshwater – 5, hot spring – 8, host-associated – 2, other – 1, unknown – 1 (Chen *et al.*, 2020).

Although the above study was from a small sample size, in subsequent larger surveys, the lack of ecological trends continues. Using ApcE2 as a FALRLIP biomarker, searches of large databases reveals FARLIP-capable cyanobacteria in marine, freshwater, brackish, hot spring and terrestrial ecosystems (Antonaru *et al.*, 2020). The ecosystems themselves are also varied. Marine-derived FARLIP-capable cyanobacteria have been found in marine plastic debris, coral microbial communities, algal communities and estuaries. Likewise, aquatic-derived FARLIP-capable cyanobacteria have been found in microbial lake communities, glacier meltwater and sulfidic groundwater. Of the few FARLIP-capable cyanobacteria isolated from terrestrial ecosystems, organisms have been found in soils, plant canopy and stones. There are also a few FARLIP-capable cyanobacteria isolated from more extreme environments; endolithic communities in the Atacama Desert of Chile, microbial mats in the Canadian High Arctic Lakes the hypersaline Soda Lake in Russia and hypersaline Hot Lake in Washington, USA. It should also be noted, no pattern emerges for geographic location – FARLIP-capable cyanobacteria appear in both the Arctic and the Antarctic, across the equator, and all directions of the Earth's hemisphere (Antonaru *et al.*, 2020).

It is clear ecological and geographical factors cannot determine the occurrence of FARLIP-capable cyanobacteria. It is instead likely that the ability to perform FARLIP arises from the necessity for survival in a highly specific environmental niche. For example, many of the FARLIP-capable cyanobacteria have been isolated from microbial mat communities, and some also algal and coral microbial communities. It has been established Chl d-containing *Acaryochloris marina* utilizes FRL absorption when under coral-reef sea squirts (Kühl *et al.*, 2005), and FARLIP-capable *Synechococcus* and *Leptolyngbya* strains can be found in upper and deep layers of hot spring microbial mats (Ohkubo and Miyashita, 2017). Microbial mats can range from millimetres to centimetres in thickness, and thus present a barrier to visible light if cyanobacteria are not within the outer layers (Fig. 6). Likewise, motility of

cyanobacteria is poor – if invertebrates are positioned on top of a cyanobacterial community, visible light will not be as easily accessible. This can be expanded to geological niches; low-light photic zones of caves in Carlsbad Caverns National Park have high quantities of Chl f and Chl d-utilization cyanobacteria (Behrendt *et al.*, 2020).

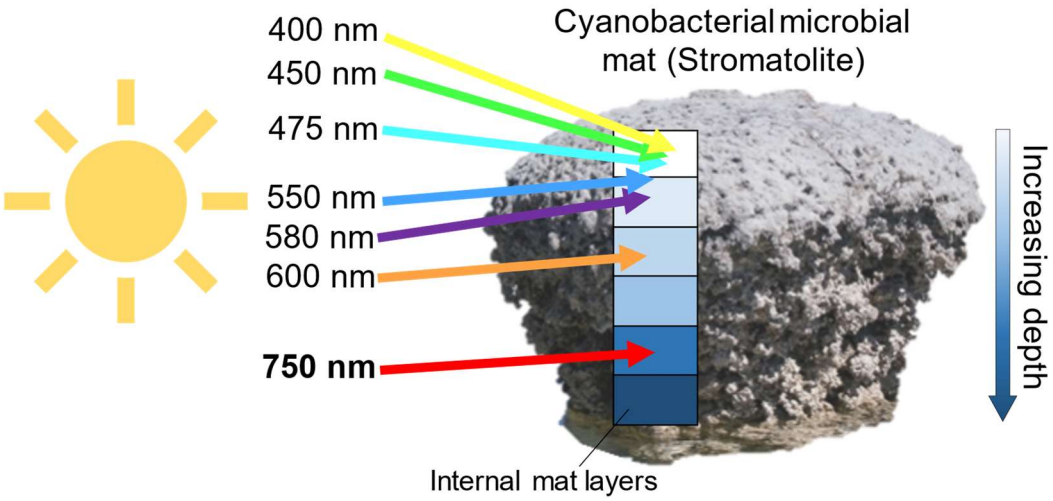


Figure 6 | FARLIP and microbial mats. FRL penetrates the most deeply into microbial mat layers, while other wavelengths of light penetrate at the surface or just below. Cyanobacteria employing FARLIP may therefore survive and grow at deeper layers. Figure is illustrative only and does not reflect actual depth penetration of the visible light spectrum.

Although the diversity of FARLIP-capable cyanobacteria has been demonstrated, the FARLIP cluster has only been characterised in 15 organisms (Ho and Bryant, 2020). Given the long evolution of cyanobacteria, further bioinformatic and genomic characterisations are required to further explore the scope of cyanobacteria capable of FARLIP, confirm conservation of the FARLIP gene cluster and determine an evolutionary origin. Phylogenetic gene trees have been produced for FARLIP-capable species and those predicted to be FARLIP-capable (Gan, Shen and Bryant, 2014; Antonaru *et al.*, 2020), both rooted and unrooted. A continued effort to explore the phylogenetic relationships between newly discovered FARLIP species, and perform comparative genomics, is needed to gain insight into the evolution of FARLIP and the FARLIP cluster.

2.4 LOLIP

The ability to survive in low-light conditions is not just limited to species containing longwave chlorophyll. *Synechococcus* found within subsurface layers of a microbial mat in Yellowstone National Park showed the ability to grow fast in low-light irradiance. These *Synechococcus* absorbed light beyond 700 nm, however, did not show synthesis of longwave chlorophylls Chl d or Chl f, only Chl a (Nowack *et al.*, 2015). These strains are therefore not FARLIP-capable or utilize Chl d for FRL utilization such as in *Acaryochloris marina*. Yet, these low-light adapted *Synechococcus* did show different pigment content compared to high-light adapted strains within the surface of the microbial mat (Nowack *et al.*, 2015), suggesting a rearrangement to light harvesting that enables the capture of FRL. Subsequent gene analysis showed these low-light adapted *Synechococcus* contained a unique set of genes: *apcD4*, *apcB3*, encoding FRL-absorbing allophycocyanin, and an IsiA-like protein, known as IsiX, encoding a Chl a-binding protein. These are summarised in Table 4. Transcript analysis showed the *apcD4, apcB3* and *IsiX* transcripts are maximal in low-light periods of the day (Olsen *et al.*, 2015). LOLIP is therefore low-light (< 50 μmol photons m^{-2} s^{-1}) as opposed to FRL-activated (Gan and Bryant, 2015).

Table 4 | Components involved in LOLIP

	Protein Name	Gene Name	Function	Visible light photosynthesis paralog gene
1	Allophycocyanin D4 / ApcD4	*apcD4*	FRL harvesting protein	*apcA1/D1*
2	Allophycocyanin B3 / ApcB3	*apcB3*	FRL harvesting protein	*apcB1*
3	IsiX	*isiX*	Chl a binding protein	*isiA*

2.4.1 The LOLIP Process

Heterologous expression of ApcB3 and ApcD4 in *E. coli* shows these pigments absorb at 615 nm and 678 nm, although purified ApcB3-ApcD4 exhibited an absorbance at 708 nm and emission at 718 nm. Replacement of ApcB3 with ApcB1 to form a ApcB1-ApcD4 complex lead to blue shifted absorbance of 615 nm and 705 nm and emission of 712 nm (Soulier, Laremore and Bryant, 2020). This confirms the FRL absorption capabilities of ApcB3-ApcD4, and suggest these allophycocyanin are responsible for the > 700 nm absorption observed in microbial mat *Synechococcus.* ApcD4 and ApcB3 contain a conserved cysteine bound to a chromophore – the binding of this chromophore is essential for FRL absorption (Gan *et al.*, 2014; Gan and Bryant, 2015; Soulier and Bryant, 2021).

Although the assembly of ApcB3-ApcD4-IsiX onto the light harvesting and photosystems is not known, it is clear ApcB3-ApcD4 takes the form of a trimer (Soulier and Bryant, 2021). This trimeric complex appears to have implications for FRL absorption. The absorbance of isolated ApcD4 is distinct from ApcD1, absorbing at 623 and 680 nm. In contrast, ApcB3 shows a similar spectral pattern to ApcB1. Yet, ApcB1-ApcD4 shows a limited absorbance at 702 nm compared to wild-type ApcB3-ApcD4 that can absorb strongly at 709 nm (Soulier and Bryant, 2021). It therefore appears, when in trimeric complex with ApcD4, ApcB3 can contribute to or increase the red-shifted absorbance of ApcD4. Understanding how and why oligomerization influences chromophore absorption would be the next step in understanding LOLIP photophysics.

Many details about this process are still unknown. It is not yet determined how LOLIP is activated, such as through response proteins as in FARLIP, nor how many subunits of ApcB3 and ApcD4 are incorporated to support light harvesting and how this incorporation remodels light harvesting. The role of IsiX, and how it works with ApcB3-ApcD4 to support FRL absorption, is also unclear. Related protein, IsiA, appears to form a light harvesting complex with PSI and provide photoprotective energy dissipation (Burnap, Troyan and Sherman, 1993; Park *et al.*, 1999; Bibby, Nield and Barber, 2001). It is possible IsiX forms a red-shifted light harvesting

complex with PSI, however this remains to be demonstrated. The possible localisation and roles of ApcB3, ApcD4 and IsiX are represented in Fig. 7.

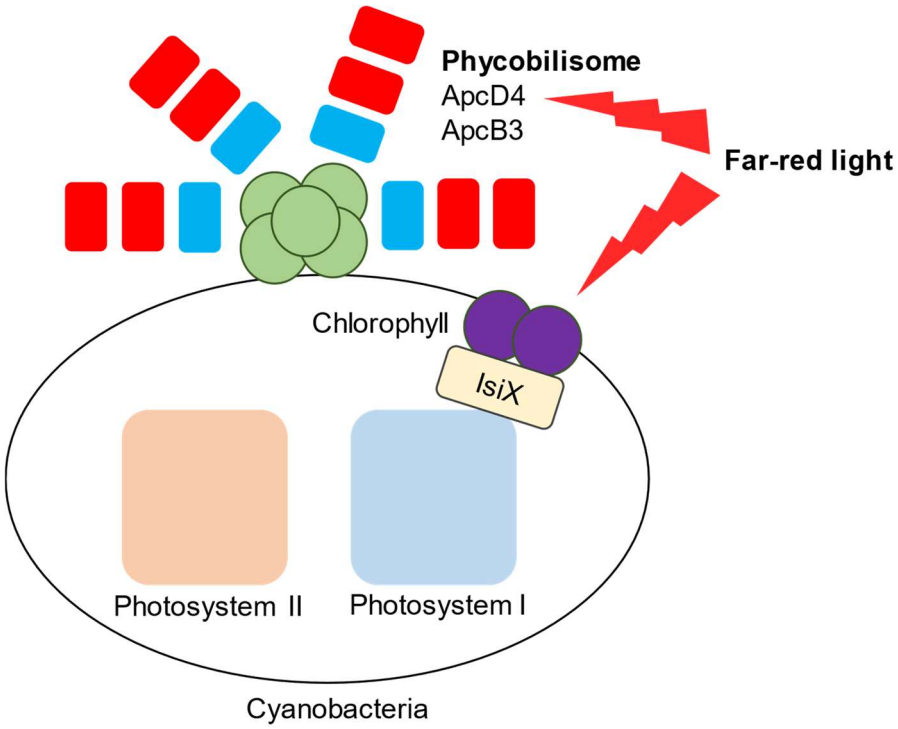


Figure 7 | Speculative LOLIP schematic. ApcD4 and ApcB3 are likely incorporated into light harvesting phycobilisomes, capturing FRL. IsiX may absorb FRL by presenting an associated far-red absorbing chlorophyll. Energy and electron transfer between photosystems and associated components likely proceeds as depicted in Fig. 3.

2.4.2 Occurrence and distribution of LOLIP

The full-scope of LOLIP-capable species, and the ecological and geographical distribution of LOLIP, has not yet been explored. LOLIP contains few components, and of these components are highly similar paralogs. This makes bioinformatic search difficult. But, it is clear the employment of FARLIP or LOLIP is not isolated among cyanobacteria; some FARLIP-capable cyanobacteria, such as *Synechococcus* sp. PCC 7335, *Chroococcidiopsis thermalis* sp. PCC 7203, *Chlorogloeopsis fritschii* PCC 9212 and *Chlorogloeopsis fritschii* sp. PCC 6912, for example, have also been found to be LOLIP-capable. Whether both these systems are employed concomitantly, or whether there is a preference for using one under certain circumstances, is not known, and needs to be explored further.

That being said, not all LOLIP-capable species are also FARLIP-capable. Species such as *Xenococcus* sp. PCC 7305, *Leptolyngbya* sp. PCC 6406 and *Gloeocapsa* sp. PCC 7428 have LOLIP genes but no FARLIP cluster (Gan and Bryant, 2015). From the species currently identified as LOLIP-capable, the ecological and geographical distribution does not appear to resemble any pattern, as in FARLIP. For example, *Xenococcus* sp. PCC 7305, *Leptolyngbya* sp. PCC 6406 and *Gloeocapsa* sp. PCC 7428 are isolated from marine, freshwater and hot spring ecosystems, respectively. The distribution of LOLIP is therefore also likely survival-dependent. Given that FARLIP appears to have no ecological and geographical pattern, it is also predicted LOLIP will show no preferences. However, this is yet to be determined.

2.5 Origin and evolution of LOLIP and FARLIP

Presumably, the need for phototrophs to survive in low light has always persisted and therefore is not confined to any cyanobacterial order. LLAC identified thus far have indeed shown a wide distribution across cyanobacterial orders (Fig. 8), such as *Pseudanabaena, Synechococcus, Pleurocapsa, Nostocales, Chroococcidiopsidales and Leptolyngbya.* But when, and how, the ability to endure low-light conditions by FRL absorption occurred is not known and has not yet been estimated.

Time-calibrated phylogenetic trees (unpublished data, Patricia Sanchez-Baracaldo research group) show that *Halomicronema hongdechloris* is among one of the oldest FARLIP- and LOLIP-capable cyanobacteria, being dated approximately 1000 million years ago. Thus, it is possible FARLIP and LOLIP could date back to 1,000 to 541 million years ago in the Neoproterozoic period. This would align with the evolution of marine picocyanobacteria that likely also emerged in the Neoproterozoic period (Sánchez-Baracaldo, 2015). Picocyanobacteria such as *Prochlorococcus* have too shown low light adaptation in marine environments. Picocyanobacteria employ Chl b in transmembrane chlorophyll as opposed to within phycobilisomes, a transition in *Prochlorococcus* that is speculated to have occurred in low oxygen oceans (Ulloa *et al.*, 2021), which also aligns with the Neoproterozoic era at the time of deep ocean oxygenation (Zhang *et al.*, 2021). As far-red light minimally penetrates water, the

need for far-red light utilisation in marine environments would be limited and thus it could be speculated marine *Prochlorococcus* separated from phycobilisome cyanobacteria in low oxygen oceans prior to the evolution of LLAC. Yet, LLAC could have evolved prior to *Prochlorococcus* if cyanobacteria were in near-surface waters. Subsequent loss of phycobilisomes from *Prochlorococcus* would result in loss of major FRL-harvesting pigments and may have allowed *Prochlorococcus* to reach deeper oceanic niches. Nonetheless, this presents a scenario where three distinct low light harvesting processes may have evolved within similar geological time.

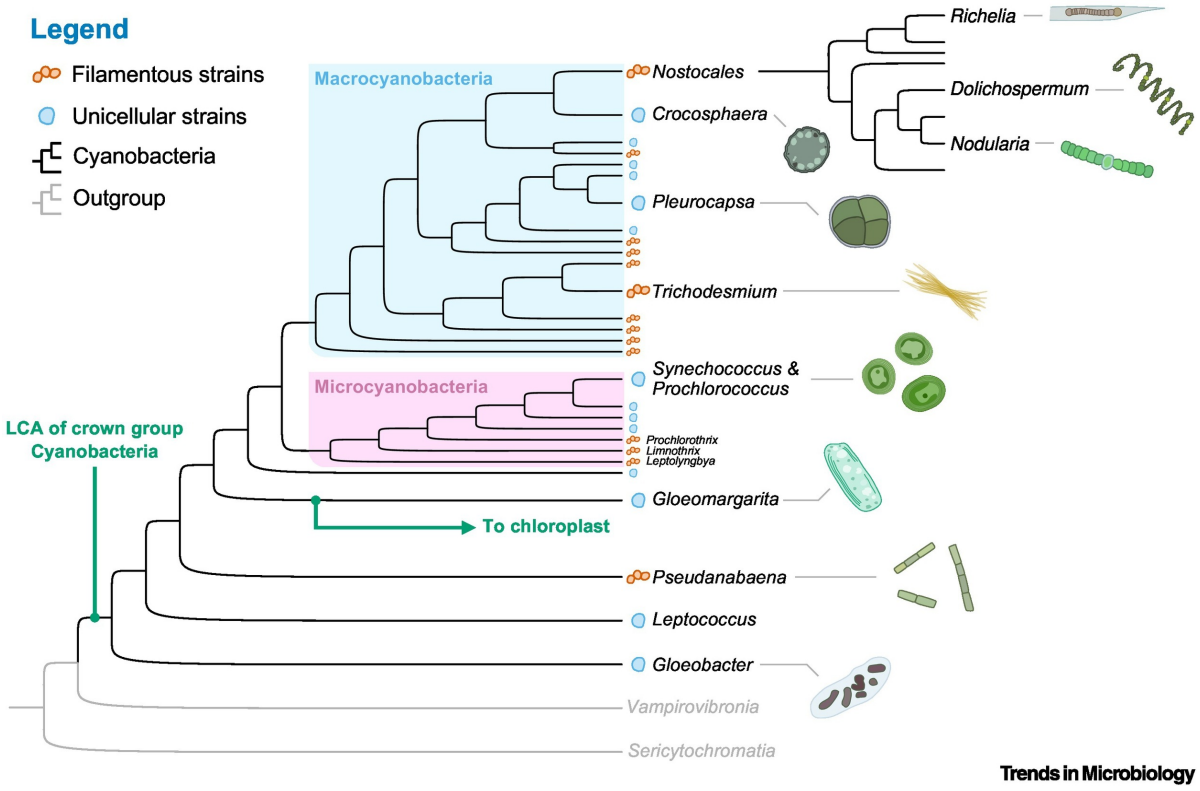


Figure 8 | Tree of life representing the main cyanobacterial groups. Closest non-photosynthetic relatives (Vampirovibrionia and Sericytochromatia) are indicated. Cladogram is taken from (Sánchez-Baracaldo *et al.*, 2022).

Whether FARLIP or LOLIP were established first, and whether they originated independent or dependent of each other, is not yet known. That being said, there is a strong argument for independent origins of FARLIP and LOLIP. There are no shared proteins among either process, except from the modification to allophycocyanin. This would be expected due to the key role of allophycocyanin in light harvesting and energy migration to the photosystems. The processes also differ greatly in activation

stimuli and the extent of photosystem remodelling. This not only suggests an independent evolution but also an evolutionary advantage to having one of these processes or both; while employment of either LOLIP or FARLIP could provide a survival advantage in low-light, employment of both could extend the time in which FRL absorption is active. Assuming LOLIP is activated by low-light and not just FRL, FRL absorption could be utilized in a variety of conditions and times of day and not limited to the presence of > 700 nm light. On the other hand, the use of one of these processes may make the other redundant, and perhaps remain inactivated. While the former may support a co-evolution theory, the latter indicates a more independent evolution. But, how LOLIP and FARLIP are employed in a single organism has yet to be demonstrated.

Regarding the origin of FARLIP, it has been speculated PsbA4, the Chl f synthase, could have arisen from a mutation in PsbA that lead to loss of OEC binding. Alternatively, PsbA may have arose from PsbA4 forming the ability to bind OEC (Ho *et al.*, 2016). However, if the latter were true, we may expect to see more species with Chl f synthase. Phylogenetic analysis has been performed for FARLIP-capable species. rRNA phylogeny of FARLIP-capable cyanobacteria show that FARLIP-capable species are diverse and have varied relatedness (Gan and Bryant, 2015; Antonaru *et al.*, 2020). Unrooted trees of FARLIP genes and proteins also match this apparent diversity (Gan, Shen and Bryant, 2014; Ho *et al.*, 2020). However, as the trees are unrooted no more information can be gained from thesis analyses.

Rooted trees have been published using maximum likelihood estimations (Antonaru *et al.*, 2020). The focus of these trees are on the characterization of metagenomes containing FARLIP gene fragments, and not to characterise the evolutionary history of FARLIP. These trees are also not rooted by stringent rooting methods, such as outgroup, mid-point or molecular clock, but by matching the tree topography with previously published unrooted trees. Further research is therefore needed to reconstruct FARLIP evolution. This could be done by characterising the full-scope of FARLIP-capable cyanobacteria, applying a Bayesian approach to determine phylogenies, and rooting with mid-point or outgroup methodologies as appropriate.

In contrast, phylogenetic analysis of LOLIP cyanobacteria is limited, and thus the origin of LOLIP is unclear. The full-scope of LOLIP species has not yet been determined, and thus rooted tree topologies of LOLIP genes or LOLIP species has not been performed. The evolution of LOLIP can not therefore be inferred and requires further insight. That being said, given that LOLIP-capable *Xenococcus* sp. PCC 7305*, Leptolyngbya sp*. PCC 6406 and *Gloeocapsa* sp. PCC 7428 are from Pleurocapsales, Synechococcales and Chroococcales orders, it is clear LOLIP may also exhibit phylogenetic diversity (Zhao *et al.*, 2015). Notably, rRNA phylogeny shows *Fischerella* sp. PCC 9605 has a common ancestor with FARLIP-capable *Fischerella* sp. NIES-4106 and *Fischerella* sp. JSC-11 (Antonaru *et al.*, 2020), yet only *Fischerella* sp. PCC 9605 exhibits both LOLIP and FARLIP genes. *Fischerella* sp. PCC 9605 therefore appears to have acquired LOLIP genes independently of lineage. Further phylogeny comparisons of LOLIP genes with cyanobacterial species trees would give insight into the acquisition of LOLIP.

2.6 Applications of cyanobacteria in astrobiology

2.6.1 As model organisms

While many extremophile bacteria are utilized as model organisms for astrobiology research, cyanobacteria capable of surviving in extremes offer unique models for understanding the limits of life on other planets. The detection of oxygen on other planets would be a strong indication for oxygen-producing life and oxygen-dependent life (Schwieterman *et al.*, 2018), and cyanobacteria represent one of the only prokaryotes on our planet to produce oxygen, survive in extreme environments and promote the development of complex life. The study of their response to non-Earth conditions could therefore give us insight into how oxygen-producing or light-utilizing life may survive, and what biomarkers remain, on other planets.

2.6.1.1 *Chroococcidiopsis*

Such astrobiology relevant research has been conducted, predominantly using cyanobacteria of the *Chroococcidiopsis* genus. Members of the *Chroococcidiopsis* genus show resistance to hot and cold temperatures (Bahl *et al.*, 2011), desiccation

(Billi, 2009; Billi *et al.*, 2011; Murik *et al.*, 2017), and ionizing radiation (Billi *et al.*, 2000; Baqué, Viaggiu, *et al.*, 2013; Verseux *et al.*, 2017). Several *Chroococcidiopsis* are also FARLIP and/or LOLIP-capable (*Chroococcidiopsis cubana* CCALA 043, *Chroococcidiopsis cubana* SAG 39.79, *Chroococcidiopsis* sp. CCALA 051, *Chroococcidiopsis* sp. FACHB-1243 and *Chroococcidiopsis thermalis* PCC 7203) (Gan, Shen and Bryant, 2014; Antonaru *et al.*, 2020). This makes them ideal for surveying the potential for life in a range of non-Earth conditions.

As a result of desiccation and temperature tolerant properties, *Chroococcidiopsis* has been used in simulated non-Earth extremes. Desiccation- and cold-tolerant *Chroococcidiopsis* CCMEE 029 and 171 were utilized to evaluate the potential for life on icy moons. Icy moon conditions of Europa were simulated by freezing solutions of either Na_2SO_4, $MgSO_4$ or NaCl to -15 ˚C, -40 ˚C or -70 ˚C. Both strains of *Chroococcidiopsis* could survive in all solutions at 15 ˚C and -40 ˚C. Neither strain could survive at -70 ˚C (Cosciotti *et al.*, 2019). Of course, while the surface temperature of Europa can go down to -227 ˚C (Ashkenazy, 2019), and salt composition more varied (Zolotov and Shock, 2001), the results indicate cyanobacterial strains can survive at the extreme boundaries of Earth life. Additionally, both strains of *Chroococcidiopsis* could survive in Na_2SO_4 and $MgSO_4$ solutions despite the absence of liquid veins between ice boundaries. This could indicate life on Europa is not limited to liquid veins, as it is thought on Earth (Mader *et al.*, 2006; Miteva, 2008; Parrilli *et al.*, 2011).

Desiccation and cold-tolerant cyanobacteria also have applications for Mars research. *Chroococcidiopsis* species not only show desiccation and cold-tolerance, but also radiation tolerance. On the surface, the average temperature is -63 ˚C (Williams, 2020), with radiation is up to 50-fold higher than Earth (Hassler *et al.*, 2014). Despite sub-surface waters being detected (Lauro *et al.*, 2020), no surface water on Mars is evident and thus only desiccation-tolerant life could survive. Studies with *Chroococcidiopsis* have revealed the limits Martian survival and indicate biomarkers such as photopigments remain preserved, indicating a potential biomarker for light-utilizing life (Baqué, Viaggiu, *et al.*, 2013; Baqué *et al.*, 2016). However, whether this applies to other cyanobacterial genera is yet to be evaluated.

On top of this, *Chroococcidiopsis* sp. have revealed potential strategies for survival by Martian life. *Chroococcidiopsis* colonised onto Antarctic rock and exposed to simulated space and Martian conditions showed resilience by colony formation and rapid DNA damage repair (Billi *et al.*, 2011). Further investigations revealed this resilience may be due to upregulation of UV-specific DNA repair genes (Mosca *et al.*, 2019), and biosynthesis of sucrose and trehalose (Fagliarone *et al.*, 2020).

The ability to survive in Martian conditions is not limited to just *Chroococcidiopsis*. The Mars Desert Research Station, located near Utah, USA, is a highly desiccated and light exposed habitat that acts as a Martian analogue site. In a recent survey, *Gloeocapsa* sp. was found as densely packed masses on sandstone and quartzite rocks within the Mars desert (Sokoloff *et al.*, 2016). The survivability of this cyanobacteria provides evidence that desiccation-tolerant and radiation-tolerant cyanobacteria can serve as models for Martian life forms.

2.6.1.2 Red dwarf star planets

The ability of cyanobacteria to perform FARLIP and LOLIP is a phenomenon only discovered within the last decade. As such, the use of LLAC as astrobiology models has not yet been fully explored. A recent study utilized LLAC as model organisms for red dwarf star planets. These red dwarf stars host super-Earths (Dressing and Charbonneau, 2015), and have a reduced luminosity, much of which falls into the infrared spectrum (Veeder, 1974; Heath *et al.*, 1999; Claudi *et al.*, 2021). An example is the TRAPPIST-1 system. The TRAPPIST-1 system features seven Earth-sized planes orbiting the ultra-cool red dwarf star, TRAPPIST-1 (Gillon *et al.*, 2017; Turbet *et al.*, 2020).

Any hypothetical photosynthetic life developing on such planets of the TRAPPIST-1 system would require an adaptation low light. LLAC could serve as model organisms in this instance; FARLIP-capable *Chlorogloeopsis fritschii* PCC 6912 and *Chroococcidiopsis thermalis* sp. PCC 7203 and FARLIP-incapable *Synechocystis* sp. PCC 6803 all displayed growth when exposed to red dwarf star light. Although only *Chlorogloeopsis fritschii* sp. PCC 6912 and *Chroococcidiopsis thermalis* sp. PCC 7203 were able to utilize the far-red end of the spectrum (Claudi *et al.*, 2021). The

study indicates low light photosynthetic life can develop under red dwarf radiation. Identifying psychrophilic LLAC may be of benefit to investigating this theory further; TRAPPIST-1e and TRAPPIST-1f are considered the most viable for habitability. Yet, both planets can reach sub-zero temperatures (Lingam sand Loeb, 2018, Wolf, 2017). A psychrophilic LLAC shown to survive under these conditions would further imply the potential for habitability of low luminosity planets by phototrophic organisms.

2.6.2 For practical application

The application of cyanobacteria for expansion into space is a developing field. Currently, there is no literature exploring the application of LLAC for practical applications in astrobiology. However, there is a wealth of speculative reviews and some explicit evidence that detail cyanobacterial applications in space, as illustrated in Fig. 9 and detailed below. Firstly, due to their production of oxygen and fixation of carbon and nitrogen, cyanobacteria are candidates as bio-fertilizer. Cyanobacteria could act to promote the stability of soils and the growth of plants and other organisms (Singh *et al.*, 2010) – such plants and organisms may be required for food and medicine beyond Earth. In a proof-of-concept experiment, cyanobacteria bio-fertilizer has been shown yield the same or improved growth of wheat (Abd-Alla, Mahmoud and Issa, 1994; Hussain and Hasnain, 2011), rice and pea plants (Osman *et al.*, 2010). Cyanobacterial bio-fertilizer has also shown to effectively support the growth of non-food organisms, such as silkworms (Chikkaswamy, 2015) and cotton (A and El-Desoukey, 2020). That being said, further studies are needed to confirm whether cyanobacteria can remain as effective bio-fertilizers under non-Earth conditions and stressors.

Cyanobacteria have also been implicated as bio-mining tools. Cyanobacteria could be utilized for the extraction of elements from rocks and planetary materials; cyanobacteria placed on volcanic and igneous rock analogous to Martian and lunar material showed an ability to biologically weather the analogue materials and gain essential nutrients for lithotrophic growth. Cyanobacteria could also survive Martian simulation and desiccation, indicating a potential application for cyanobacteria on Mars as bio-mining and bio-weathering tools (Olsson-Francis and Cockell, 2010).

However, biomining of asteroid or other non-Earth material has yet to be demonstrated, as it has for bacteria (Cockell *et al.*, 2020).

Cyanobacteria could also offer sources of fuel. Hydrogen fuel, for example, is not only a clean form of energy but also light, can be sourced abundantly and is easily stored. It is for these reasons hydrogen fuel is utilized to build space stations and power aerospace technology (Jain, 2009). Cyanobacteria can produce hydrogen from water via solar energy conversion (Hallenbeck, 2012; Thomas, 2014), and thus offer a source of hydrogen fuel production both on Earth and beyond (Bolatkhan *et al.*, 2019; Kolbe, Lechtenböhmer and Fischedick, 2019; Sadvakasova *et al.*, 2020). Indeed, *Anabaena variabilis* can continue hydrogen production even when in partial vacuum. Over five months, *Anabaena variabilis* steadily produced 0.02–0.2 mL molecular hydrogen per mg dry weight of cyanobacterial biomass per hour (Markov *et al.*, 1993). Reviews have speculated hydrogen yield could be enhanced by genetic and metabolic approaches, and the process could be made more efficient by optimising the use of a cheap nutrient substrate, such as waste water (Bolatkhan *et al.*, 2019; Sadvakasova *et al.*, 2020).

Cyanobacteria could also aid the production of bio-petrol, bio-diesel and bio-fuel cells. Such fuel can be used for energy generation, electricity, heat, and powering vehicles and machinery. It has been demonstrated *Synechocystis* sp. PCC 6803 and *Synechococcus elongatus* sp. PCC 7942 can produce and tolerate most hydrocarbon fuel and precursors, with exception of ethylene. Production of alkanes is predicted to be particularly successful (Kämäräinen *et al.*, 2012). In production of bio-diesel, *Anabaena* sp. PCC 7120 has been engineered to secrete the cyclic monoterpene limonene, which can be employed both as a bio-diesel and a jet fuel. *Anabaena* sp. PCC 7120 could produce limonene at a maximum of 3.6 ± 0.5 μg L^{-1} $O.D.^{-1}$ h^{-1} at high light intensity (Halfmann, Gu and Zhou, 2014). In a similar role, cyanobacteria have shown to be employed as fuel cells. Bio-fuel cells convert chemical energy, such as those from glucose, fructose, cellobiose, into electrical energy. Bio-fuel cells have been successfully formed with cyanobacteria (Yagishita *et al.*, 1996; Sawa *et al.*, 2017). For example, *Anabaena variabilis* M-3 derived bio-fuel cells could show a net lifetime of 80 hours (Yagishita *et al.*, 1996). However, the

application and efficiency of these cyanobacteria derived fuels within space technology has not yet been tested.

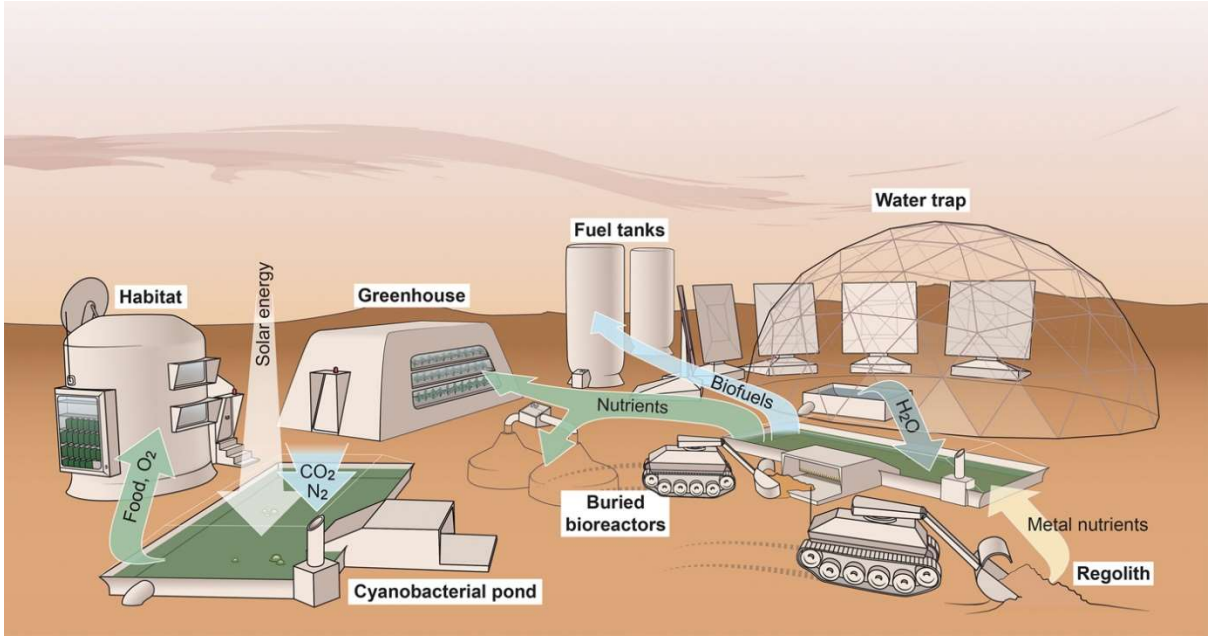


Figure 9 | Practical application of cyanobacteria on Mars. Cyanobacterial ponds could provide an output of several useful products, such as food, oxygen and nutrients, which could feed and support habitats containing heterotrophic bacteria or plants. Cyanobacteria can also output fuel and bio-fuel materials, and be used to leach crucial metals from Martian regolith. Figure taken from (Verseux *et al.*, 2016), designed by Cyprien Verseux and Sean McMahon (Yale University).

2.7 Motivation

Currently, only one study (Claudi *et al.*, 2021) has explored the use of LLAC as model organisms in astrobiology, and the application of LLAC in food or energy production are yet to be studied. This progress is hindered by several key knowledge gaps about LLAC. Firstly, the true distribution and diversity of LLAC is not yet clear; the ability to survive in low-light appears to be a survival niche and can not be predicted by geography or ecology. As LLAC contain paralogous proteins to those encoded in every cyanobacteria (Gan, Shen and Bryant, 2014), this makes it difficult, although not impossible, to identify LLAC through bioinformatics. Many LLAC that could be suitable for astrobiology applications may therefore be undiscovered. Understanding the evolutionary relationships of FARLIP and LOLIP organisms, and those that can employ both, can help us understand the acquisition of low-light

survival which may also underpin low-light survival on other planets. While some phylogenetic analyses have been performed implemeting maximum likelihood (Gan, Shen and Bryant, 2014; Gan and Bryant, 2015; Shen *et al.*, 2019; Antonaru *et al.*, 2020), more work is needed implemeting rooted methodologies to help establish directionality in evolution. Such as which genes evolved first and in which strains/lineages.

Therefore, this study aims to explore the distribution, diversity and phylogeny of both FARLIP and LOLIP cyanobacteria, with the intention of identifying new species of LLAC. Crucially, we hope to identify LLAC that may be relevant for astrobiology applications. The overreaching aim of this project is add to the growing knowledge about LLAC, and communicate the potential for LLAC in astrobiology research. This project therefore has three key aims:

i) Bioinformatically identify FARLIP and LOLIP species using unique gene identifiers
ii) Confirm the presence of FARLIP and LOLIP gene clusters in newly identified species, and explore the evolutionary relationships of key genes in these clusters with a Bayesian approach
iii) Speculate, with bioinformatical support, candidate LLAC for astrobiology research, such as those that may be useful as model organisms for astrobiology, or suitable candidates for use within life support systems in space and on other planets

3. Methods

3.1 Identification of FARLIP and LOLIP-capable cyanobacteria

To identify FARLIP-capable species, 20 cyanobacterial ApcE2 sequences (Supplementary Table S1) were submitted in a blastp and tblastn (e-value: 10^{-5}) against the NCBI non-redundant protein database (release v.203) (ncbi.nlm.nih.gov/) and JGI/IMG database (v.5.4) (img.jgi.doe.gov/) (BLOSUM62; e-value: 10^{-5}). Retrieved sequences were aligned with MAFFT v.7 (Katoh and Standley, 2013) and manually checked for the characteristic ApcE2 'VIPEDVT' motif in Jalview v.2.11.1.14 (Waterhouse *et al.*, 2009). For taxa recovered with permanent draft genomes, genome completeness was assessed by CheckM v.1.4.0 through the KBase online portal (Arkin *et al.*, 2018).

LOLIP-capable species were identified using the JGI/IMG Cassette Search online tool (v.5.4) exploring the 1,678 cyanobacteria genomes available. To fulfil the three gene minimum requirement, the following LOLIP-associated proteins from *Chlorogloeopsis fritschii* PCC 9212 were entered: ApcB3 (2550829381), ApcD4 (2550829382) and light-harvesting chlorophyll binding protein (LHCB) (2550829379). LHCB is commonly associated with the *apcD4-apcB3-isiX* cassette. Use of IsiX in place of LHCB returned no results. A maximum distance between genes (5000 nt) and a minimum distance from scaffold edge (1 nt) was applied. Retrieved genomes were identified as LOLIP by location of the *apcD4-apcB3-isiX* cassette in the JGI/IMG Gene Neighbourhood tool. Ecological and geographical information was recorded for each organism (Supplementary Table 4; Fig. 11A).

3.2 FARLIP cluster localisation and comparative genomics

The workflow pipeline is illustrated in Fig. 10. FARLIP clusters were located using the ApcE2 sequences retrieved in 3.1. To find additional FARLIP genes encoded on different contigs, sequences of genes missing from the initial search were used in a blastp. Gene cluster information around ApcE2 or other blastp query was characterised using the Gene Neighbourhood information available in the JGI/IMG online portal (v.5.4). For genomes only available on NCBI, the surrounding gene

cluster was characterised using the GenBank (release v.242) annotated record of the contig. To confirm the identity of the annotated genes, translated protein sequences were aligned against known FARLIP listed in Supplementary Table S2 using MAFFT v.7 (Katoh and Standley, 2013). Genes without protein sequences were firstly converted to amino acid using ExPASy translate tool (Gasteiger *et al.*, 2003). A maximum likelihood (ML) tree was generated using IQ-TREE v.1.6.1 (Nguyen *et al.*, 2015) with automated substitution model and ultrafast bootstrap analysis. FARLIP clusters different strains were compared using Clinker v0.0.20 (Gilchrist and Chooi, 2020).

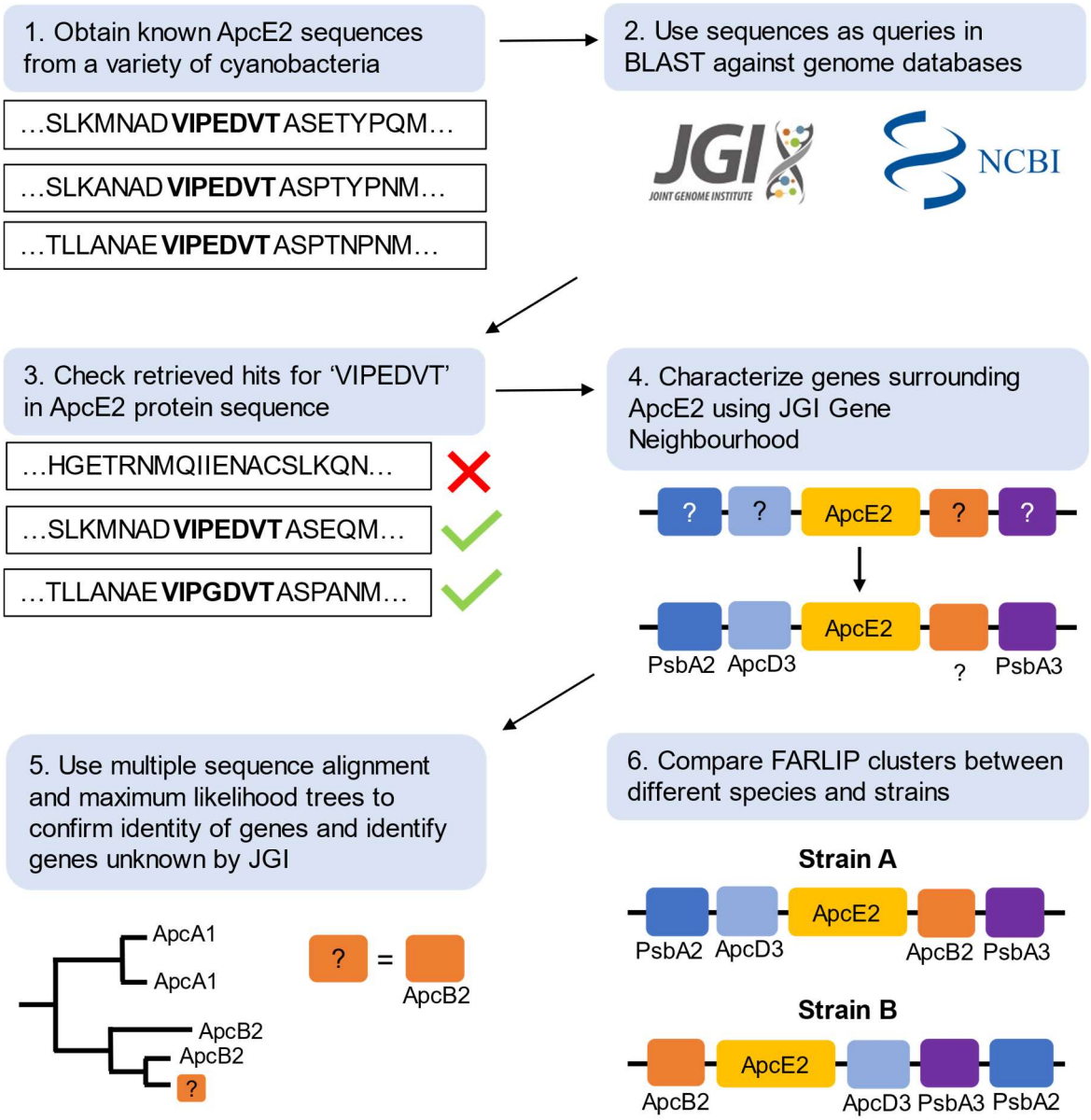


Figure 10 | Workflow pipeline for identification and comparative genomics of FARLIP-cluster containing cyanobacteria

3.3 Bayesian Inference

ApcE2 protein sequence data from 88 FARLIP-capable taxa and ApcB3 protein sequence data from 28 LOLIP-capable taxa were used to estimate phylogeny. Sequences were obtained from Genbank (release v.242) and JGI/IMG (v.5.4) using blastp (non-redundant protein sequences, BLOSUM62; e-value 10^{-5}). *Fischerella thermalis* BR2B, *Hydrococcus* sp. CSU_1_8 and *Hydrococcus* sp. RU_2_2 were

excluded due to having only partial proteins that showed a < 20% loss of sequence. Full sequences were aligned using MAFFT v.7 (BLOSUM62; Gap open penalty: 1.53; Gap extension penalty: 0.123; Maxiterate: 2) and Bayesian phylogeny estimated with MyBayes 3.1.7a (ngen=100000000; diagnfreq=10000; relburnin=yes, burninfrac=0.25; Nchains=4) (Ronquist and Huelsenbeck, 2003). JTT amino acid substitution model was estimated by MrBayes and the variation in substitution rates across sites modelled using an invariant gamma model. Trees were rooted with minimal ancestor deviation (MAD) (Tria, Landan and Dagan, 2017). MAD considers all edges of the unrooted tree as possible root positions, and derives the mean relative deviation from the molecular clock implied by each possible rooting. The edge that minimizes relative deviation is chosen as the root point.

3.4 Gene and genome size analysis

Genome size and number of genes are recorded on the JGI/IMG online portal (v.5.4). These records were collated from 810 cyanobacteria listed as high quality genomes, as well as any FARLIP and LOLIP-capable cyanobacteria. Genome and gene information of FARLIP and LOLIP-capable cyanobacteria unavailable on JGI/IMG were attained through NCBI Assembly.

3.5 BLAST analysis of astrobiology-relevant genes

Reference sequences for astrobiology-relevant genes are available in Supplementary Table 3. Reference genes were utilized in a blastp against FARLIP and FARLIP and LOLIP-capable cyanobacteria using NCBI non-redundant protein database (release v.205) and JGI/IMG database (v.5.4) (BLOSUM62; e-value 10^{-5}). LOLIP cyanobacteria were excluded due to the limited information surrounding the efficiency of this process alone. Cyanobacteria were qualified as capable of nitrogen fixation, carbon fixation, sucrose synthesis, trehalose synthesis, desaturase production or hydrogen production by having hits against all the reference genes utilized for the specified process.

4. Results

4.1 FARLIP and LOLIP-capable cyanobacteria

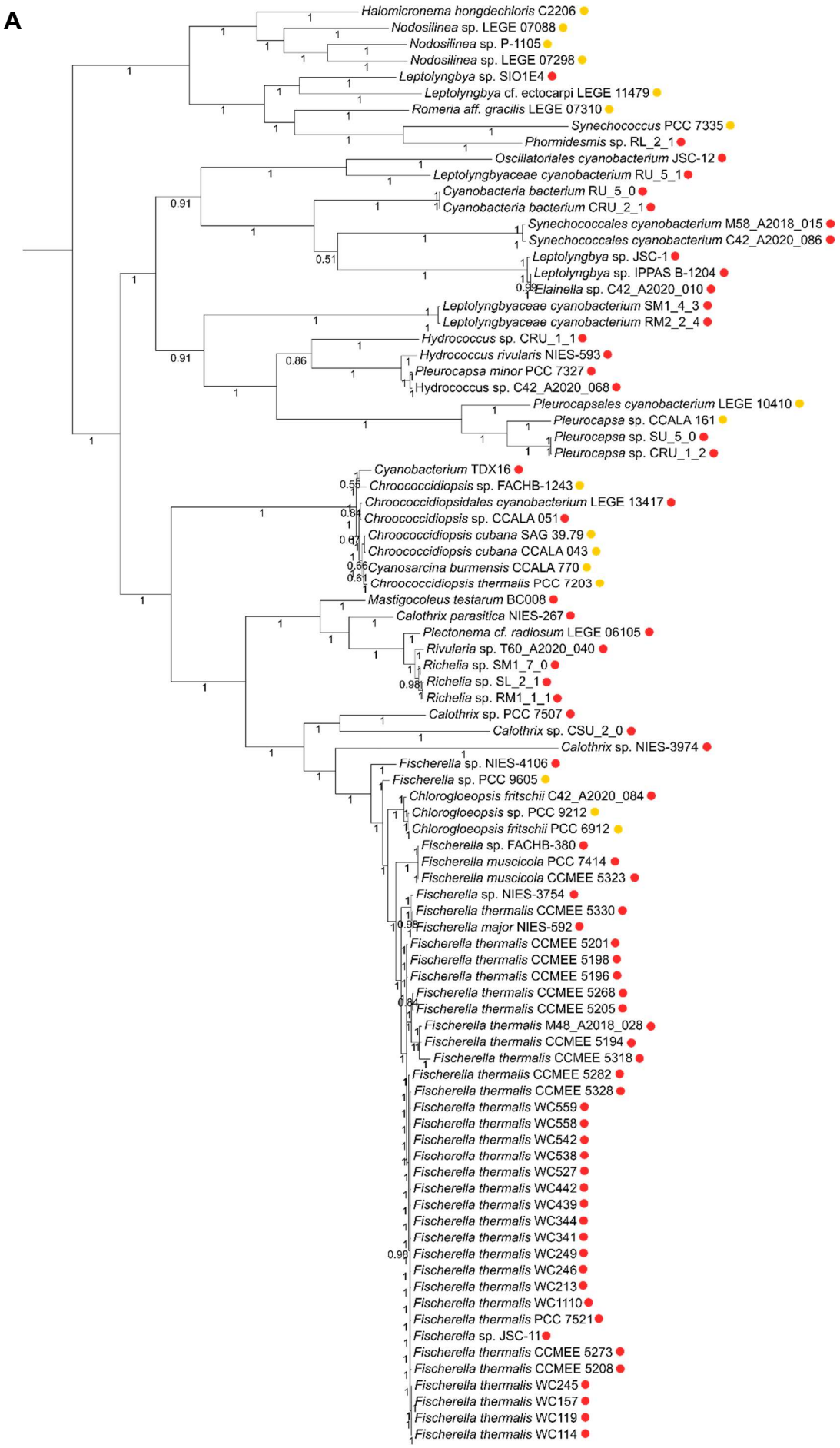
BLAST searches of JGI/IMG and NCBI databases recovered 74 strains of cyanobacteria with the FARLIP-associated protein ApcE2, 11 strains with the LOLIP-associated ApcB3, and 17 strains with both ApcE2 and ApcB3. ApcE2 was identified and distinguished from ApcE1 by a characteristic VIPEDVT motif within the phycocyanobilin binding domain, that is otherwise a IIENACS motif in ApcE1. The substitution of the cysteine in this motif constitutes an important shift in function, as it leads to non-covalent binding of the phycocyanobilin and red-shifting of light absorbance (Miao *et al.*, 2016; Ho *et al.*, 2017). The JGI/IMG Cassette Search online tool retrieved several ApcB3-containing species. Alignment analysis of the sequences revealed ApcB3 could be distinguished from ApcB1 and ApcB2 by a conserved GDITLPGGNMYP motif, where the initial glycine and final proline are otherwise serine/threonine in ApcB1 and ApcB2. The functional significance of this has not yet been characterised, but could be speculated to enhance FRL absorbance as in ApcE2. It should be noted further LOLIP strains could be retrieved by using this identifier in a BLAST search. Retrieved strains are listed in Table S4.

Cyanobacteria identified as ApcE2-containing span a wide taxonomic diversity, with ApcE2-positive strains spanning over 20 genera of cyanobacteria. The majority of species identified were *Fischerella* isolated from Yellowstone National Park. This is likely a reflection of the density of microbial research which has been conducted in this area. Of the retrieved ApcE2-positive strains, 29 strains have not yet been reported as FALRIP-capable or as containing FARLIP-associated proteins (Table S5), with five new entirely genera of cyanobacteria being identified as containing FARLIP genes: *Nodosilinea, Plectonema, Richelia, Phormidesmis* and *Romeria*. This therefore expands the taxonomic range of cyanobacteria that show an ability to perform FARLIP.

Cyanobacteria identified as ApcB3-positive also span a large diversity, featuring 14 different genera of cyanobacteria. It should be noted JGI/IMG holds over 1,600 cyanobacterial genomes. Although extensive, this search does not represent a full-

scope of LOLIP-capable cyanobacteria, but indicates LOLIP has the potential to be an equally diverse light adaption process. While there has been limited research into LOLIP, a few key cyanobacteria have been identified as LOLIP-capable, such as *Xenococcus* sp. PCC 7305, *Synechococcus* sp. PCC 7335 and *Chroococcidiopsis thermalis* PCC 7203 (Gan and Bryant, 2015; Soulier, Laremore and Bryant, 2020). On top of these, our search identified 15 cyanobacteria that have not yet been reported as LOLIP-capable or as containing LOLIP-associated proteins (Table S6). Six of these cyanobacteria are only LOLIP-capable and do not appear to contain FARLIP-associated genes (*Candidatus Gloeomargarita lithophora* D10, *Chroococcales cyanobacterium* IPPAS B-1203, *Gloeocapsopsis crepidinum* LEGE 06123, *Leptolyngbya sp*. KIOST-1, *Lusitaniella coriacea* LEGE 07157, *Scytonema millei* VB511283).

These newly identified LLAC span several taxonomic orders: *Pleurocapsales*, *Oscillatoriophycideae*, *Synechococcales, Chroococcidiopsidales*, *Nostocales* and *Gloeoemargaritales*. Phylogeny for recovered FARLIP and LOLIP strains are shown in Fig 11A and Fig 11B, respectively.


A
Halomicronema hongdechloris C2206
Nodosilinea sp. LEGE 07088
Nodosilinea sp. P-1105
Nodosilinea sp. LEGE 07298
Leptolyngbya sp. SIO1E4
Leptolyngbya cf. ectocarpi LEGE 11479
Romeria aff. gracilis LEGE 07310
Synechococcus PCC 7335
Phormidesmis sp. RL_2_1
Oscillatoriales cyanobacterium JSC-12
Leptolyngbyaceae cyanobacterium RU_5_1
Cyanobacteria bacterium RU_5_0
Cyanobacteria bacterium CRU_2_1
Synechococcales cyanobacterium M58_A2018_015
Synechococcales cyanobacterium C42_A2020_086
Leptolyngbya sp. JSC-1
Leptolyngbya sp. IPPAS B-1204
Elainella sp. C42_A2020_010
Leptolyngbyaceae cyanobacterium SM1_4_3
Leptolyngbyaceae cyanobacterium RM2_2_4
Hydrococcus sp. CRU_1_1
Hydrococcus rivularis NIES-593
Pleurocapsa minor PCC 7327
Hydrococcus sp. C42_A2020_068
Pleurocapsales cyanobacterium LEGE 10410
Pleurocapsa sp. CCALA 161
Pleurocapsa sp. SU_5_0
Pleurocapsa sp. CRU_1_2
Cyanobacterium TDX16
Chroococcidiopsis sp. FACHB-1243
Chroococcidiopsidales cyanobacterium LEGE 13417
Chroococcidiopsis sp. CCALA 051
Chroococcidiopsis cubana SAG 39.79
Chroococcidiopsis cubana CCALA 043
Cyanosarcina burmensis CCALA 770
Chroococcidiopsis thermalis PCC 7203
Mastigocoleus testarum BC008
Calothrix parasitica NIES-267
Plectonema cf. radiosum LEGE 06105
Rivularia sp. T60_A2020_040
Richelia sp. SM1_7_0
Richelia sp. SL_2_1
Richelia sp. RM1_1_1
Calothrix sp. PCC 7507
Calothrix sp. CSU_2_0
Calothrix sp. NIES-3974
Fischerella sp. NIES-4106
Fischerella sp. PCC 9605
Chlorogloeopsis fritschii C42_A2020_084
Chlorogloeopsis sp. PCC 9212
Chlorogloeopsis fritschii PCC 6912
Fischerella sp. FACHB-380
Fischerella muscicola PCC 7414
Fischerella muscicola CCMEE 5323
Fischerella sp. NIES-3754
Fischerella thermalis CCMEE 5330
Fischerella major NIES-592
Fischerella thermalis CCMEE 5201
Fischerella thermalis CCMEE 5198
Fischerella thermalis CCMEE 5196
Fischerella thermalis CCMEE 5268
Fischerella thermalis CCMEE 5205
Fischerella thermalis M48_A2018_028
Fischerella thermalis CCMEE 5194
Fischerella thermalis CCMEE 5318
Fischerella thermalis CCMEE 5282
Fischerella thermalis CCMEE 5328
Fischerella thermalis WC559
Fischerella thermalis WC558
Fischerella thermalis WC542
Fischerella thermalis WC538
Fischerella thermalis WC527
Fischerella thermalis WC442
Fischerella thermalis WC439
Fischerella thermalis WC344
Fischerella thermalis WC341
Fischerella thermalis WC249
Fischerella thermalis WC246
Fischerella thermalis WC213
Fischerella thermalis WC1110
Fischerella thermalis PCC 7521
Fischerella sp. JSC-11
Fischerella thermalis CCMEE 5273
Fischerella thermalis CCMEE 5208
Fischerella thermalis WC245
Fischerella thermalis WC157
Fischerella thermalis WC119
Fischerella thermalis WC114

B

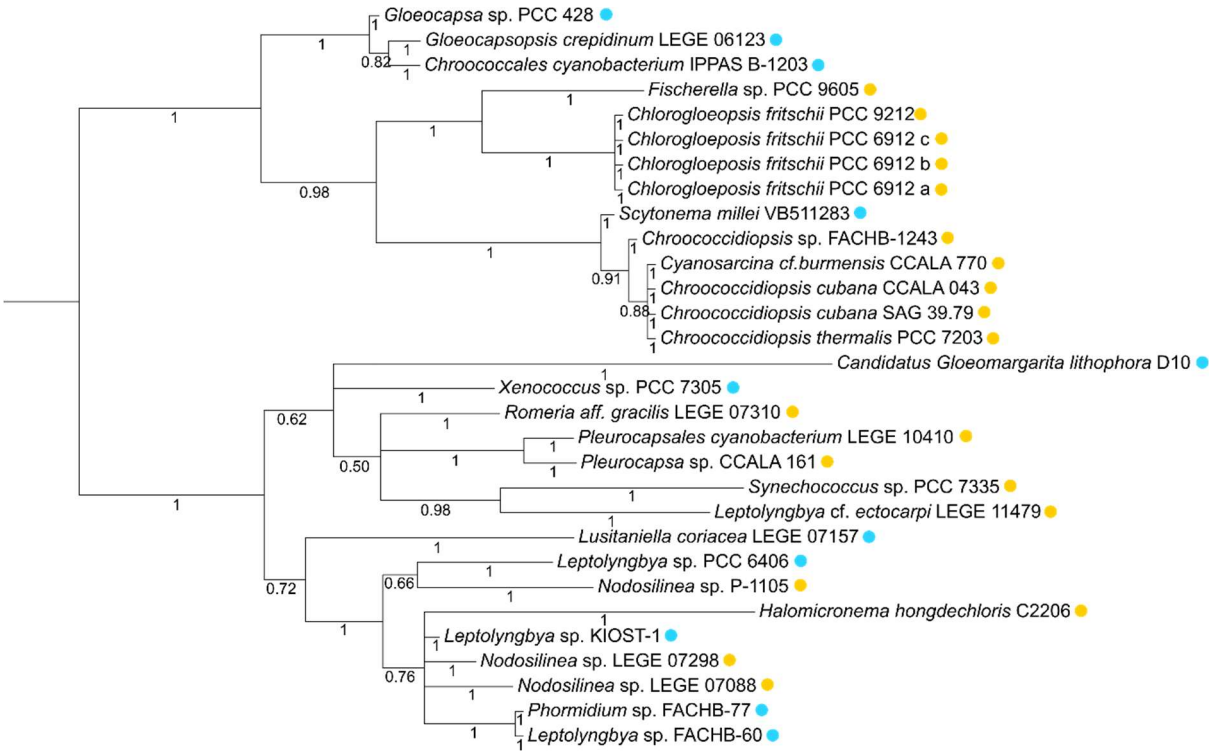


Figure 11 | Phylogenetic analysis of ApcE2 and ApcB3 sequences from identified LLAC. (A) ApcE2 phylogeny. (B) ApcB3 phylogeny. Trees were built with MrBayes and rooted with MAD. Posterior probabilities are indicated. JTT amino acid substitution model was auto-selected (DT, WR and JM, 1992). Photosynthetic capability is indicated by red (FARLIP), blue (LOLIP) and yellow (FARLIP and LOLIP).

4.2 Geographic, ecological and genetic distribution

Previous studies have shown that the occurrence of FARLIP does not appear to correlate to any geographic or ecological locations (Antonaru *et al.*, 2020). However, this has not previously been explicitly shown for LOLIP species. To investigate this further, we have completed comparative genome analysis to identify any patterns in genome size against FARLIP and/or LOLIP occurrence (Fig. 12A), and recorded the global and ecological (Fig. 12B, Fig 12C) distribution of the retrieved LLAC species including LOLIP species.

An alternative photosynthetic capability could be more likely in organisms which contain a higher gene count and/or genome size. Fig. 12A indicates the average

gene count is between 4000 to 6000, and the average genome size is between 5 and 6 Mbp. Averaging both these factors for LLAC indicate the average gene count to be 6500 genes and average genome size to be approximately 6.8 Mbp. This falls only slightly outside the average values. The likelihood of being FARLIP or LOLIP-capable, or both, can therefore not be predicted by a larger genome.

Results of the global distribution of LLAC echoes what has been found previously locations (Antonaru *et al.*, 2020). The 102 strains of LLAC identified not only span a large taxonomic range but also a diverse geological and ecological range. FARLIP- and/or LOLIP-capable strains do not appear to display disposition towards any geographic location or ecological niche (Fig. 12B). While it could be suggested LLAC tend towards areas in and around the equator, there are also a lack of cyanobacterial samples from colder regions that may account for this bias (Chrismas, Anesio and Sánchez-Baracaldo, 2015; Chrismas, Anesio and Śanchez-Baracaldo, 2018). Cyanobacterial strains isolated from the same locations all exhibit identical photosynthetic capability. For example, all cyanobacteria isolated from White Creek, Lower Geyser Basin, Yellowstone National Park, USA, exhibit FARLIP genes, likely due to intermingling of these strains. Only a single discrepancy was observed; despite both *Gloeocapsopsis crepidinum* LEGE 06123 and *Plectonema cf. radiosum* LEGE 06105 being isolated from the intertidal zone of Praia da Luz, Lagos, Portugal, *Gloeocapsopsis crepidinum* LEGE 06123 contains only FARLIP genes, while *Plectonema cf. radiosum* LEGE 06105 contains only LOLIP genes.

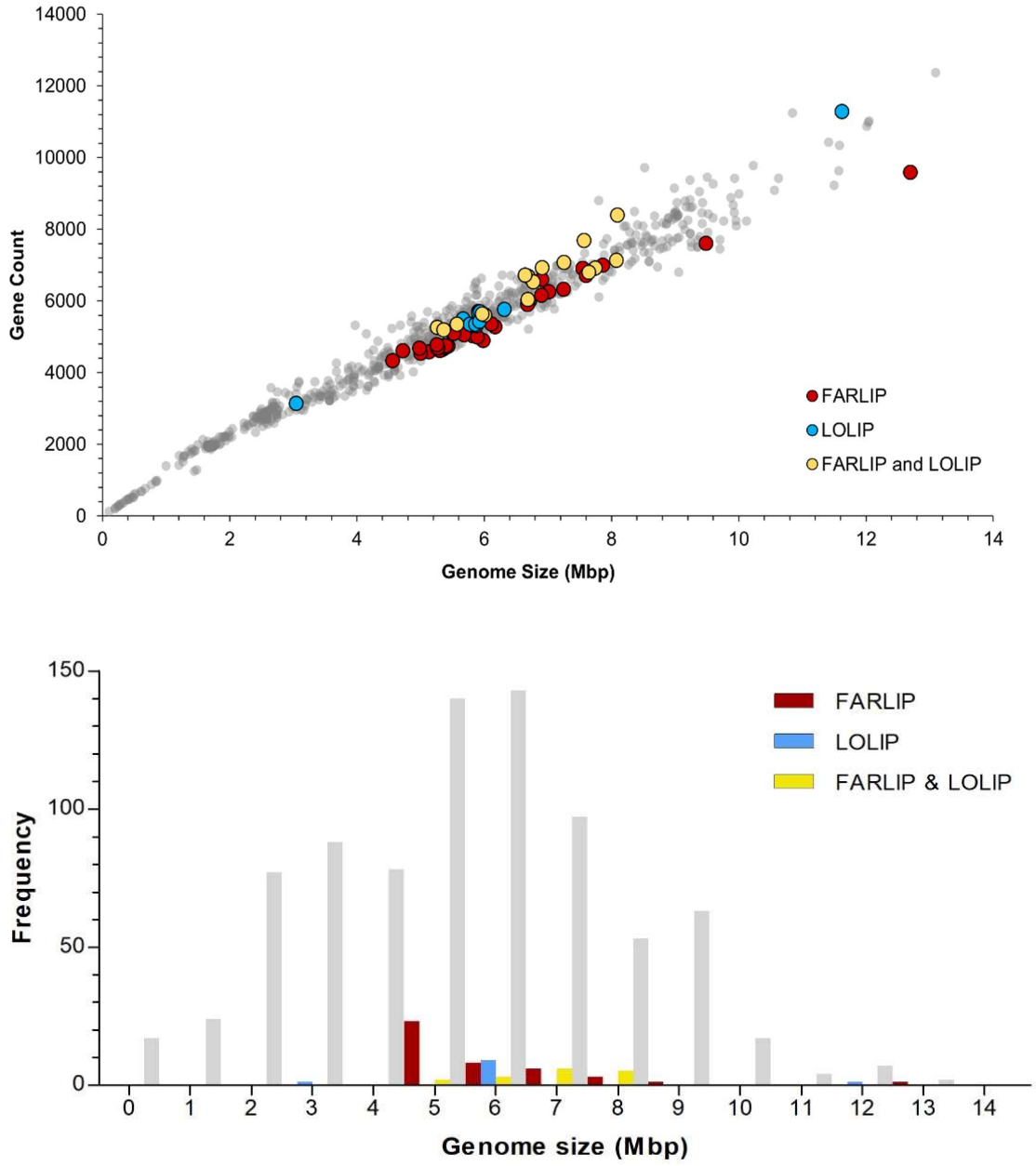
A bias towards hot springs can be observed for FARLIP strains (Fig. 12C). FARLIP-capable strains show a preference towards hot spring and marine environments, with over 26% and 55% being isolated from hot spring and marine environments, respectively. FARLIP strains were predominately isolated from Yellowstone National Park, USA; of the 74 strains identified as FARLIP-capable, 20 were derived from this region as a result a large-scale microbial surveys of the hot springs. However, this does not appear to account for the bias observed. When adjusting for duplicate locations, the ecological imbalance towards hot springs remains, whereas it balances for marine ecosystems (data not shown). This may be accounted for by the fact aquatic areas such as hot springs commonly observe the formation of microbial mats (Rozanov *et al.*, 2017; Prieto-Barajas, Valencia-Cantero and Santoyo, 2018).

Due to the deep layers that occlude light, microbial mats have been shown to inhabit Chl f-producing cyanobacteria (Ohkubo and Miyashita, 2017).

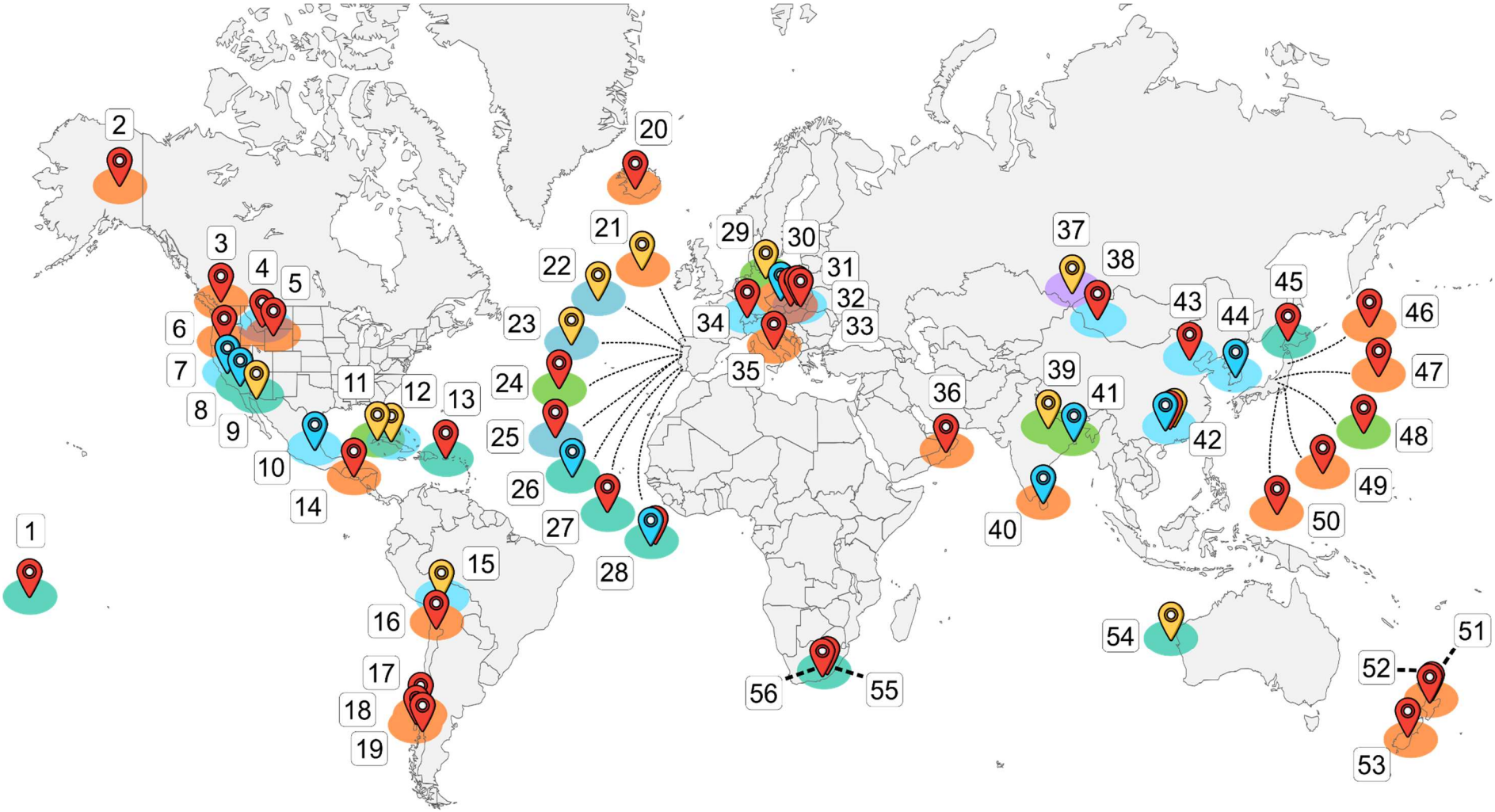
However, LOLIP-capable cyanobacteria, nor FARLIP- and LOLIP-capable cyanobacteria, show this same bias towards hot springs. In fact, these cyanobacteria are isolated least from hot spring areas. This may be in part influenced by low sample size; LOLIP-capable cyanobacteria were only located using the JGI/IMG database, which features 1,678 cyanobacteria genomes.

Ecology or geographic environment does not appear to influence the presence of an alternative photosynthetic capability, yet there appears to be a geological aspect to FARLIP and LOLIP distribution. It is apparent there are more coastal-adjacent FARLIP and LOLIP isolates as opposed to in-land isolates. While this could also be a sampling bias towards easily accessible waters, it could also suggest that FARLIP and LOLIP have an oceanic ancestral origin that has allowed widespread global distribution on modern Earth. In-land LLAC that have been isolated from lakes and soils, such as *Nodosilinea* sp. P-1105 (37) or *Chlorogloeopsis fritschii* PCC 6912 (39) may have become deposited by the water cycle. This could support a theory of evolution within oceans of the Neoproterozoic era, an era in which LLAC *Halomicronema hongdechloris* may have emerged (unpublished data, Patricia Sanchez-Baracaldo research group). If so, LLAC emergence may have coincided with the emergence of low light adapted *Prochlorococcus* which is likewise speculated to have occurred in low oxygen Neoproterozoic oceans (Ulloa *et al.*, 2021). If this were the case, it would be expected open ocean samples would contain LLAC, similar to the widespread oceanic distribution observed for *Prochlorococcus* (Flombaum *et al.*, 2013). The distribution of LLAC across open ocean samples however is yet to be explored.

A


Gene Count
Genome Size (Mbp)
FARLIP
LOLIP
FARLIP and LOLIP
Frequency
Genome size (Mbp)
FARLIP
LOLIP
FARLIP & LOLIP

B


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56

C

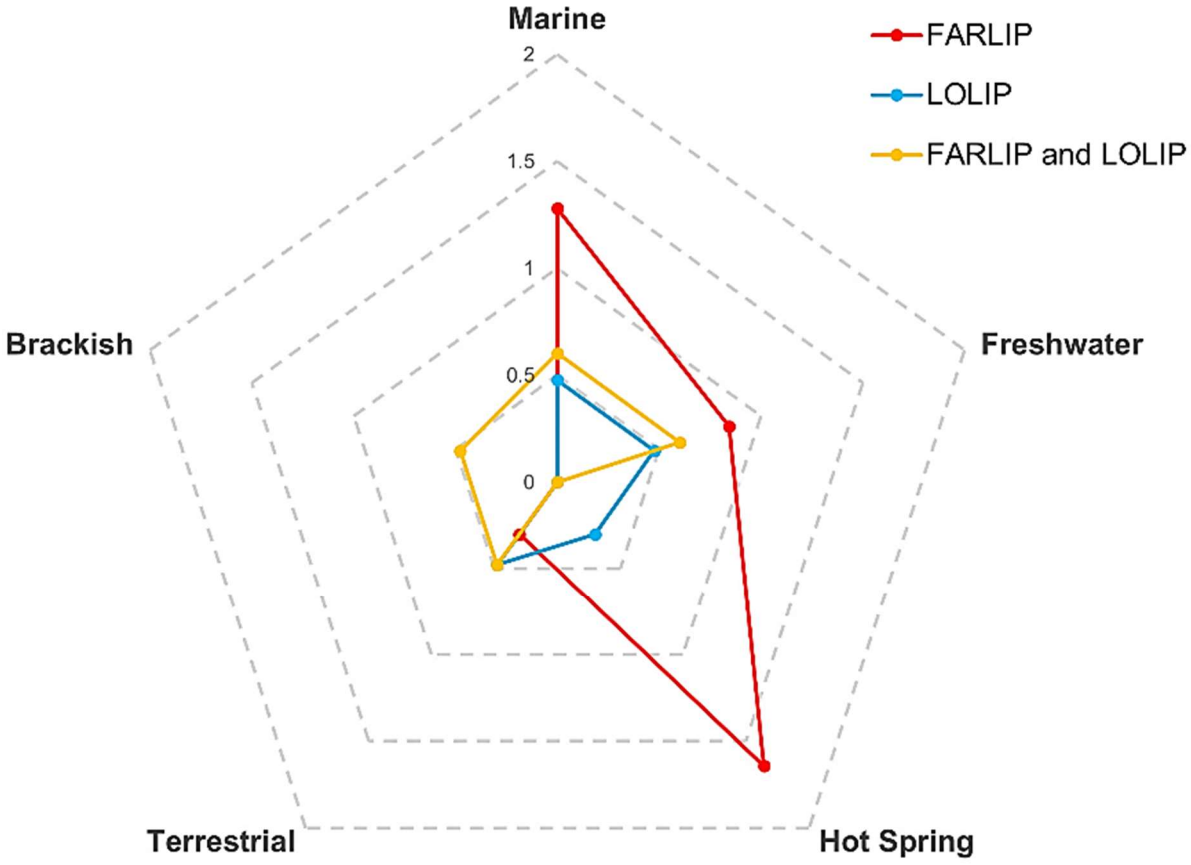


Figure 12 | Global, ecological and genomic distribution of FARLIP and LOLIP cyanobacteria. (A) Genetic comparison**.** Scatter plot (top) shows gene count versus genome size of 810 cyanobacteria (grey) against that of FARLIP, LOLIP and FARLIP and LOLIP species. Histogram (bottom) shows frequency of genome size of 810 cyanobacteria (grey) against that of FARLIP, LOLIP and FARLIP and LOLIP species. (B) Mapped locations of FARLIP and LOLIP cyanobacteria identified through ApcE2 and ApcB3 BLAST analysis, with exception of *Fischerella thermalis* JSC-11, *Fischerella* sp. PCC 9605 and *Oscillatoriales cyanobacterium* JSC-12 due to missing information regarding isolation. Geographic and ecological isolation sources were noted from NCBI Biosample. Numbers correspond to organisms listed in Supplementary Table S4. Locations are colour coded based on ecological environment (Green – terrestrial; purple – hypersaline; turquoise – marine; dark blue – brackish; cyan – freshwater; orange – hot spring), and location pins colour coded based on photosynthetic capability (red – FARLIP; blue – LOLIP; yellow – FARLIP and LOLIP). (C) Ecological comparison. The ecology of isolation sources were

recorded for each LLAC identified through BLAST. The ecology of isolation sources were tallied and quantified for each photosynthetic capability (FARLIP, LOLIP, FARLIP and LOLIP). Final total values were converted to Log10 for clarity due to the large bias towards hot spring-derived species.

4.3 Comparative genomics of FARLIP gene clusters

To date, the FARLIP clusters of 15 strains of cyanobacteria across 7 different genera have been comparatively analysed (Gan, Shen and Bryant, 2014; Ho and Bryant, 2020). Our BLAST analysis has revealed *Nodosilinea, Plectonema, Richelia, Phormidesmis* and *Romeria* as entirely new genera to be ApcE2-containing, encompassing the following strains: *Nodosilinea* sp. LEGE 07298, *Nodosilinea* sp. LEGE 07088, *Nodosilinea* sp. P-1105, *Plectonema cf. radiosum* LEGE 06105, *Romeria aff. gracilis* LEGE 07310, *Richelia* sp. SL_2_1, *Richelia* sp. SM1_7_0, *Richelia* sp. RM1_1_1 and *Phormidesmis* sp. RL_2_1. Although the FARLIP cluster shows a high level of conservation (Gan, Shen and Bryant, 2014; Ho and Bryant, 2020), it was of interest to further characterise the FARLIP cluster in these cyanobacteria to identify any genus-based additions or deletions that may be apparent (Fig. 12, Table S7). The resulting comparative genomics is illustrated in Fig. 13.

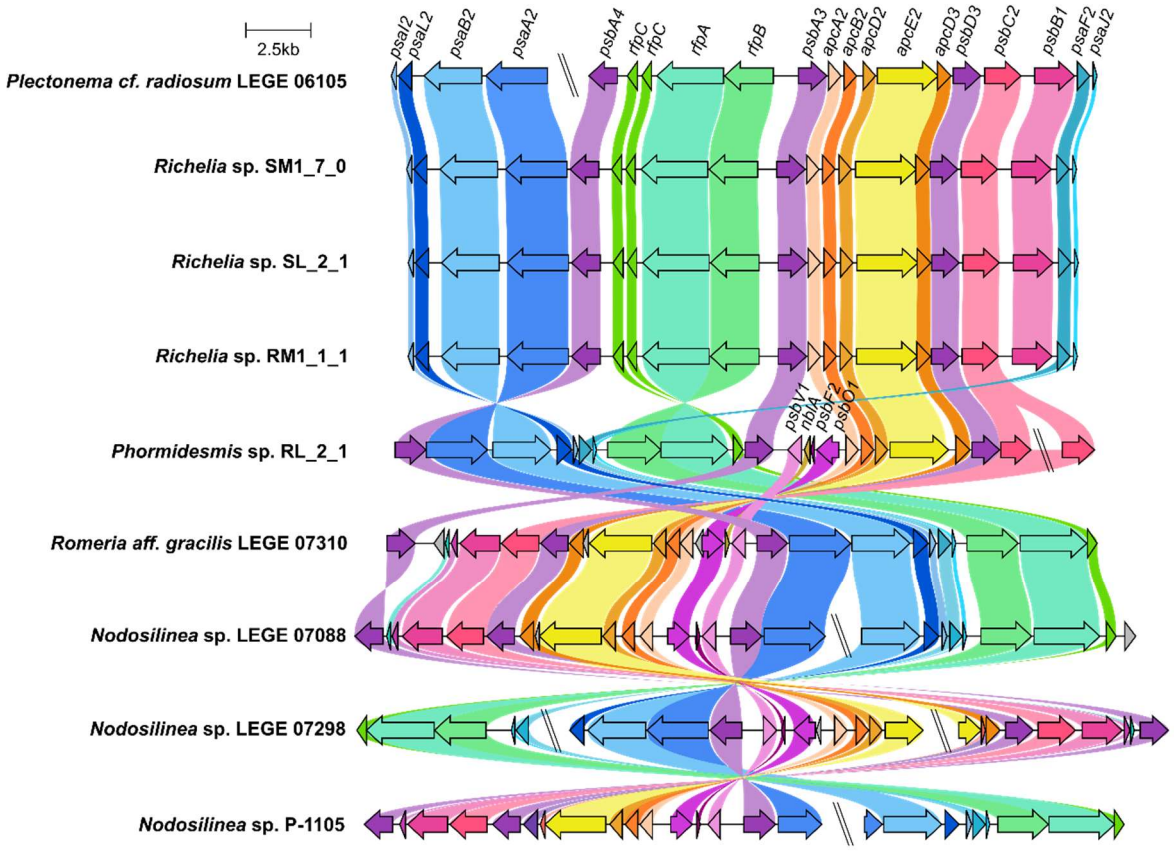


Figure 13 | Comparative genomics of FARLIP clusters across five genera of cyanobacteria. Genes are colour coded as accordingly: *psa* (blue), *psb* (purple/pink), *rfp* (green), *apc* (orange/yellow). Illustrated using Clinker (Gilchrist and Chooi, 2020).

All of the above strains had a genome completeness of over > 94%. Comparison of these clusters reveals some notable points of difference. *Romeria aff. gracilis* LEGE 07310, *Plectonema cf. radiosum* LEGE 06105, *Nodosilinea sp*. LEGE 07298, *Nodosilinea* sp. LEGE 07088 and *Nodosilinea* sp. P-1105 display all 20 FARLIP genes (*rfpA/B/C, apcA2/B2/D2/E2/D3, psbA3/D3/C2/B2/H2/A4, psaA2/B2/L2/I2/F2/J2*). *Phormidesmis* sp. RL_2_1, *Richelia* sp. SL_2_1, *Richelia* sp. SM1_7_0 and *Richelia* sp. RM1_1_1, however, appear to be missing the *psbH2* gene, and *Phormidesmis* sp. RL_2_1 is additionally missing *psbB2*. As both these genes act in chlorophyll binding, it is possible their absence genes may hinder FRL capture and FARLIP capabilities. Despite these missing genes, the conservation of the FARLIP cluster remains strong and corroborates what was found previously by Gan *et al* (Gan, Shen and Bryant, 2014; Ho and Bryant, 2020).

There are also additions into the FARLIP cluster within all strains. All *Nodosilinea* strains exhibit the addition of *psbO1-psbF2-psbV1* adjacent to the allophycocyanin gene set. This gene cluster encodes proteins with roles in the OEC, the primary site of water splitting in the PSII. The gene *psbO1* encodes the photosystem II oxygen-evolving enhancer protein - a protein that stabilizes the manganese cluster of the OEC, the primary site of water splitting (De Las Rivas and Barber, 2004). The gene *psbV1* encodes cytochrome c550, a low-potential cytochrome c that plays a role in the OEC of PSII (Katoh *et al.*, 2001). The gene *psbF2* has an unclear role; *psbF2* encodes the beta subunit of cytochrome b559 within PSII. While the function of this protein is unclear, deletion of both PsbF and PsbE, the alpha subunit, resulted in a non-functional PSII (Pakrasi, Williams and Arntzen, 1988).

Nodosilinea sp. P-1105, *Nodosilinea* sp. LEGE 07088 and *Nodosilinea* sp. LEGE 07298 all share the *psbO1-psbF2-psbV1* gene set. Yet, the isolation sources were from Russia (*Nodosilinea sp.* P-1105) and Portugal (*Nodosilinea* sp. LEGE 07088 and *Nodosilinea* sp. LEGE 07298). The shared gene set indicates there was a time whereby these strains or previous lineages of these strains co-occurred. *Halomicronema hongdechloris* C2206 also displays this cluster, but is absent of *psbF2* (Gan, Shen and Bryant, 2014; Ho and Bryant, 2020). Species tree phylogeny (unpublished data, Patricia Sanchez-Baracaldo research group) shows a common ancestor between all four strains within the same clade, with *Halomicronema hongdechloris* C2206 showing the most distant ancestry. It is therefore reasonable to assume loss of *psbF2* in *Halomicronema hongdechloris* C2206, or gain of *psbF2* in *Nodosilinea* strains, occurred in subsequent, more immediately related ancestors. How, and whether, these proteins contribute to FARLIP could be determined by gene knock-out.

A similar insertion is observed in *Phormidesmis* sp. RL_2_1 and *Romeria aff. gracilis* LEGE 07310. The insertion in *Phormidesmis* sp. RL_2_1, also adjacent to the allophycocyanin gene set, consists of *psbO1-psbF2-nblA-psbV1.* The gene *nblA* encodes a phycobilisomes degradation protein that is activated upon nitrogen deprivation (Baier *et al.*, 2004; Karradt *et al.*, 2008). In contrast, the insertion in *Romeria aff. gracilis* LEGE 07310 consists of *psbO1-nblA-psbV1.* Notably, *Romeria*

aff. gracilis LEGE 07310 lacks the *psbF2* gene commonly associated with this cluster. According to the phylogeny in Fig. 11A, *Romeria aff. gracilis* LEGE 07310 and *Halomicronema hongdechloris* C2206 ApcE2 predates that of *Phormidesmis* sp. RL_2_1. This implies relatively recent insertion of *nblA* into the FARLIP cluster in their evolutionary history. The addition of *nblA* into the FARLIP cluster may also indicate an adaptation to regulate FARLIP during nitrogen starvation.

Plectonema cf. radiosum LEGE 06105 and all *Richelia* strains exhibit two copies of *rfpC*. Species phylogeny (unpublished data, Patricia Sanchez-Baracaldo research group) indicates *Plectonema cf. radiosum* LEGE 06105 and *Richelia* share a common ancestor. Thus, the additional copy of *rfpC* likely arose from a gene duplication event in an ancestor. Notably, *Mastigocoleus testarum* BC008 and *Calothrix parasitica* NIES-267 also share a common ancestor with *Plectonema cf.radiosum* LEGE 06105 and *Richelia SL_2_1* (unpublished data, Patricia Sanchez-Baracaldo research group). Yet, the former do not have duplicate *rfpC*. This suggests any possible duplication event that occurred was specific to the ancestor from which *Plectonema cf. radiosum* LEGE 06105 and *Richelia SL_2_1* are derived. Whether this gene is active or inactive is unclear. Analysis of gene duplications within Chl d-containing *Acaryochloris* has indicated gene duplications are quickly lost from the genome unless they provide important contributions to the ecology of the organism, such as surviving in low-light conditions (Miller *et al.*, 2011). As RfpC is postulated to act as a phosphate shuffle between RfpA and RfpB (Zhao *et al.*, 2015), it is possible this gene duplication provides a more efficient FARLIP activation by increasing the amount of RfpC available to receive and release a phosphate.

One further discrepancy is observed between the strains. While *Plectonema cf. radiosum* LEGE 06105 and all *Richelia* strains contain PsbB1 in the FARLIP cluster, *Romeria aff. gracilis* LEGE 07310 and all *Nodosilinea* have PsbB2 and *Phormidesmis* sp. RL_2_1 contains no PsbB. PsbB, or CP47, is a core PSII antenna component that binds chlorophyll (Eaton-Rye and Vermaas, 1991; Hird *et al.*, 1991). It has only been documented to bind Chl a, whether it also contributes to Chl d or Chl f binding is not known. Notably, however, while deletion of partner core antenna PsbC, CP43, does not affect PSII electron transport, deletion of PsbB leads to loss of PQ and PSII activity (Vermaas *et al.*, 1987; Vermaas, Ikeuchi and Inoue, 1988).

While *Phormidesmis* sp. RL_2_1 lacks PsbB2 in the FARLIP cluster, the genome does carry a PsbB1 that thus allows retention of PQ and PSII activity. Both *Plectonema cf. radiosum* LEGE 06105 and *Richelia* strains, and *Romeria aff. gracilis* LEGE 07310 and *Nodosilinea* strains, share a common ancestor. This could indicate paralogous PsbB as a redundant protein, however the FARLIP capabilities of *Phormidesmis* sp. RL_2_1, or a related strain, have yet to be demonstrated.

The effects of utilizing either PsbB1 or PsbB2 within the FALRIP cluster is unclear. The regions between Gly-351 and Thr-365 are essential for PsbB retention in *Synechocystis* sp. PCC 6803 (Eaton-Rye and Vermaas, 1991), and PsbB contains several histidine that facilitate chlorophyll binding (Shen, Eaton-Rye and Vermaas, 1993; Barber, Morris and Büchel, 2000). Multiple sequence alignment (data not shown), indicates both PsbB1 and PsbB2 remain conserved in this region and retain all histidine required for chlorophyll binding. The use of PsbB1 or PsbB2 within the FARLIP cluster may reflect small amino acid changes that facilitate binding of Chl a, Chl d or Chl f. Further studies are needed to clarify essential components of FARLIP and quantify the impact of insertions and deletions within this well conserved cluster.

4.4 Comparative genomics of LOLIP gene clusters

Although the LOLIP cluster is small, the genomics have yet to be analysed. We have compared the LOLIP clusters of the LOLIP-capable strains listed in Table S4. All LOLIP-capable strains show conservation of the core *apcD4-apcB3* genes. However, *Chlorogloeopsis fritschii* sp. PCC 6912, *Chroococcidiopsis thermalis* PCC 7203, *Cyanosarcina cf. burmensis* CCALA 770, *Fischerella* sp. PCC 9605, *Gloeocapsa* sp. PCC 7428, *Halomicronema hongdechloris* C2206, *Leptolyngbya* sp. KIOST-1, *Pleurocapsales cyanobacterium* LEGE 10410 and *Pleurocapsa* sp. CCALA 161, *Synechococcus* sp. PCC 7335 and *Xenococcus* sp. PCC 7305 do not exhibit the presence of *isiX*. It should be noted, however, the LOLIP cluster is cut-off between the contigs of *Cyanosarcina cf. burmensis* CCALA 770 and thus this strain may exhibit *isiX* also.

The LOLIP clusters thus appear to fall into one of three genomic compositions: *apcD4-apcB3-lhcb-isiX*, *apcD4-apcB3-lhcb* and *apcD4-apcB3* (Fig. 14). The most prominent LOLIP cluster is *apcD4-apcB3-lhcb-isiX,* which was exhibited in 64% of the LOLIP-capable cyanobacteria identified. Given the prominence of LHCB over IsiX in the LOLIP-capable cyanobacteria, it appears LHCB may function as the more prominent light harvesting tool for LOLIP. The complex formed between ApcD4, ApcB3, IsiX and LHCB as yet to be elucidated, and the contribution of either IsiX or LHCB, or both together, to LOLIP is not yet known. Notably, *Pleurocapsa* sp. CCALA 161, *Pleurocapsales cyanobacterium* LEGE 10410 and *Cyanosarcina cf. burmensis* CCALA 770 show neither *isiX* nor *lhcb.* Whether FARLIP is functional in these cyanobacteria, and the impact of lacking IsiX and/or LHCB, has yet to be demonstrated.

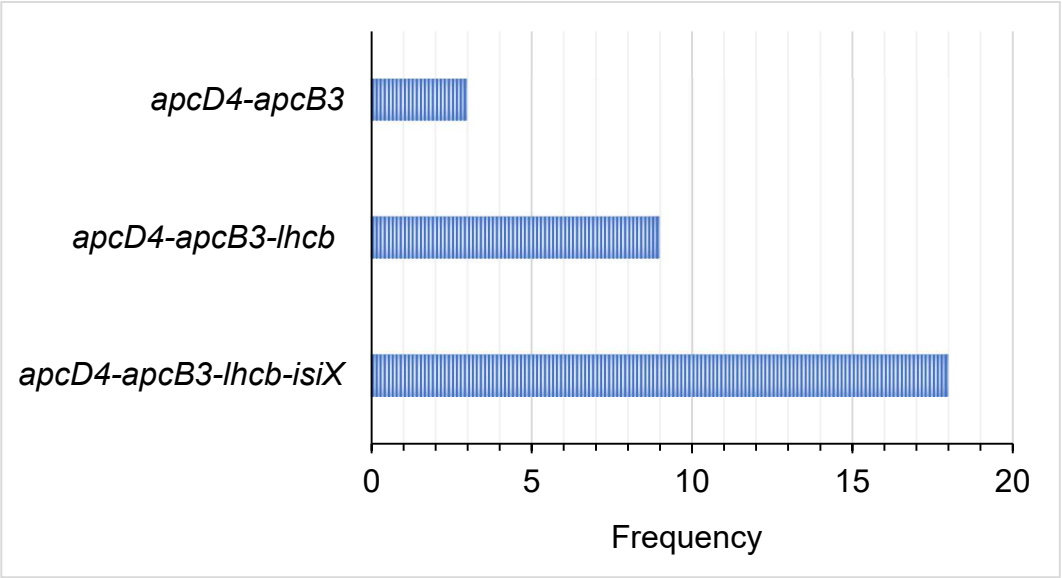


Figure 14 | Frequency of LOLIP cluster arrangements in identified LLAC. LOLIP cluster-containing cyanobacteria were identified through a BLAST analysis using ApcD4, ApcB3 and LHCB. LOLIP cluster arrangement was recorded using the JGI/IMG Gene Neighbourhood tool. Note: due to nature of the JGI/IMG Cassette search tool, identified LOLIP species are limited to those containing LOLIP clusters with adjacent LHCB.

While the LOLIP cluster is known as *apcD4-apcB3-isiX* (Gan and Bryant, 2015; Olsen *et al.*, 2015), comparative genomics of this study has revealed that LHCB is also commonly associated with the LOLIP cluster. All LOLIP-capable strains, with

exception of *Pleurocapsa* sp. CCALA 161, *Pleurocapsales cyanobacterium* LEGE 10410 and *Cyanosarcina cf. burmensis* CCALA 770, exhibit an association of the *lhcb* gene with *apcD4* and *apcB3*. LHCB encodes a chlorophyll a/b binding protein that functions in light harvesting and delivery of photon excitation energy to PSII and PSI (Cinque, Croce and Bassi, 2000; Teramoto, Ono and Minagawa, 2001). It is therefore possible LHCB functions to enhance light acquisition for LOLIP. IsiX is also predicted to function as a Chl a binding protein; unpublished results indicate ApcD4, ApcB3, ApcE and IsiX form a complex (Gan and Bryant, 2015).

The tree topology in Fig. 11B indicates that the ApcB3 gene shares a common ancestor in *Pleurocapsa* sp. CCALA 161 and *Pleurocapsales cyanobacterium* LEGE 10410, and as such both exhibit the *apcD4-apcB3* LOLIP cluster form. Common ancestors between *Phormidium* sp. FACHB-77 and *Leptolyngbya* sp. FACHB-60, *Nodosilinea* strains, *Chlorogloeopsis fritschii* PCC 6912 and *Chlorogloeopsis fritschii* PCC 9212, and *Gloeocapsopsis crepidinum* LEGE 06123 and *Chroococcales cyanobacterium* IPPAS B-1203 can also explain the presence of the extended LOLIP cluster *apcD4-apcB3-lhcb-isiX* in these strains. The occurrence of *apcD4-apcB3-lhcb* however appears sporadic; cyanobacteria that exhibit this LOLIP cluster form are not derived from a common ancestor sharing the same ApcB3. Likewise, several cyanobacteria share ApcB3 from a common ancestor and yet present different LOLIP cluster forms. For example, *Cyanosarcina cf. burmensis* CCALA 770, *Chroococcidiopsis cubana* CCALA 043 and *Chroococcidiopsis thermalis* PCC 7203 share a common ancestor but display *apcD4-apcB3, apcD4-apcB3-lhcb-isiX* and *apcD4-apcB3-lhcb*, respectively. There is not yet a species phylogeny with these strains to account for this discrepancy.

A further observation was that *Chlorogloeopsis fritschii* PCC 6912 appears to contain more than one LOLIP cluster. *Chlorogloeopsis fritschii* PCC 6912 exhibits three LOLIP clusters, two of *apcD4-apcB3-lhcb* and one *apcD4-apcB3-lhcb-isiX*. A single gene cluster of *apcD4-apcB3-lhcb-isiX* also occurs in *Chlorogloeopsis fritschii* PCC 6912. The strains are geographically separated, suggesting a lineage-based acquisition of *apcD4-apcB3-lhcb-isiX*. Multiple copies of *apcD4-apcB3-lhcb* however must have been acquired extrinsically. Notably, *Chlorogloeopsis fritschii* PCC 6912

was isolated from soil. Multiple copies of light harvesting genes may have been advantageous to acquire light when embedded under soil.

As the ApcB3 and ApcD4 sequences in all three clusters of *Chlorogloeopsis fritschii* PCC 6912 exhibit 100% amino acid homology (data not shown), evolutionary acquisition of the these LOLIP clusters can not be determined with phylogeny. Transcript analysis could indicate whether one or all of these clusters are activated in response to low-light, and the subsequent effect on *Chlorogloeopsis fritschii* PCC 6912 LOLIP compared to other species with a single LOLIP cluster.

4.5 Candidates for astrobiology research and application

FARLIP and LOLIP cyanobacteria require little light and could therefore be utilised as models to study the potential for habitability in low light receiving environments. For example, planets proximal to the near infrared-emitting dwarf star TRAPPIST-1 (Gillon *et al.*, 2016, 2017). As a result of reduced solar heat, planets such as these are likely to be cold or sub-freezing and therefore any life would be adapted to such temperatures. On Earth, cold exposed life has been shown to accumulate the sugars sucrose and trehalose (Guy, Huber and Huber, 1992; Kandror, DeLeon and Goldberg, 2002; Stitt and Hurry, 2002; Phadtare and Inouye, 2008; Grewal and Jagdale, 2010; Khani, Moharramipour and Barzegar, 2013). The cryoprotective capabilities of sucrose and trehalose has been demonstrated (Strauss and Hauser, 1986; Anchordoguy *et al.*, 1987; Rodrigues *et al.*, 2008; Zhang *et al.*, 2017; Caturla-Sánchez *et al.*, 2018), with research indicating these sugars influence the freezing point of the intracellular fluid and reduce ice crystal formation, thus preserving membrane morphology and providing whole cell stability (Phadtare and Inouye, 2008) . Fatty acid desaturases have also shown upregulation in cyanobacteria and bacteria exposed to cold (Murata and Wada, 1995; Aguilar, Cronan and De Mendoza, 1998; Los and Murata, 1999; Sakamoto and Murata, 2002). Desaturases desA, desB, desC and desD facilitate cold survival by adjusting fluidity in the cell membrane (Los and Murata, 1999), with desA being an essential desaturase for low-temperature tolerance of cyanobacteria (Wada *et al.*, 1992; Wada, Gombos and Murata, 1994). LLAC capable of sucrose, trehalose and desaturase synthesis could

therefore act as analogues to hypothetical low light harvesting organisms on planets that receive reduced solar luminosity and heat.

LLAC also show potential in BLSS. BLSS functions to reduce the requirement for supplies from Earth by *in situ* regeneration of fuel, oxygen, water and food. One key requirement for growing plants or bacteria for food and/or medicine is the availability of fixed nitrogen and carbon for growth. Fuel could be provided in the form of hydrogen (Dutta *et al.*, 2005; Jain, 2009; Bolatkhan *et al.*, 2019). Cyanobacteria have been long considered for such applications due to endogenously possessing carbon fixation enzymes RuBisCO and carbonic anhydrase, the latter of which traps inorganic carbon within cells by converting CO2 to HCO3− (Rosgaard *et al.*, 2012), as well as nitrogen fixation enzymes dinitrogenase (MoFe Protein, encoded by *nifD* and nifK) and dinitrogenase reductase (Fe Protein, encoded by *nifH*) (Dutta *et al.*, 2005). Some also show a capability to produce hydrogen, a process that is linked to nitrogen fixation. In this process, a bidirectional hydrogenase (encoded by *hoxFUYH*) can uptake molecular hydrogen produced by nitrogenase or produce molecular hydrogen from protons (Dutta *et al.*, 2005; Bolatkhan *et al.*, 2019). LLAC offer an additional advantage in that growth would not be limited in crowded microbial environments for BLSS, where other forms of light may be obstructed, and that illumination of far-red light as opposed to the spectrum of visible light requires less energy. LLAC capable of nitrogen fixation, carbon fixation and/or hydrogen production could therefore serve as valuable foundations for engineering cyanobacterial BLSS systems.

To assess the astrobiological potential of the LLAC identified in this study, an extensive BLAST analysis has been performed to identify genes of interest for BLSS (nitrogen fixation, carbon fixation, hydrogen production), as well as genes associated with cold-, saline- and desiccation-tolerance (sucrose synthesis, trehalose synthesis, desaturases). Of the LLAC analysed, 42 showed a strong genetic disposition towards the astrobiology-relevant traits analysed (Fig. 15). It is notable few cyanobacteria exhibit all common desaturase genes *desA-desD* that have been shown to be associated with cyanobacterial cold tolerance (Nishida and Murata, 1996; Los and Murata, 1999; Chrismas *et al.*, 2016a), with only *Calothrix parasitica* NIES-267, *Calothrix* sp. NIES-3974, and *Plectonema cf. radiosum* LEGE 06105

containing all four of these genes. Given that these cyanobacteria also contain sucrose and trehalose synthesis genes – the aforementioned strains therefore represent potential candidates for use as model organisms that could be present on low-irradiance, cold planets. That being said, 93.5% of the cyanobacteria genomes analysed contained *desC*, and 50% contained *desA,* which is considered the most vital desaturase in cold tolerance cyanobacteria (Wada *et al.*, 1992; Wada, Gombos and Murata, 1994)

Of note however, while *Chroococcidiopsis thermalis* PCC 7203 does not contain *desD*, the genome does contain *desA-desC*, as well as sucrose and trehalose genes. The *Chroococcidiopsis* species has been shown to not only withstand Martian conditions (Verseux *et al.*, 2017; Puente-Sánchez *et al.*, 2018; Billi *et al.*, 2019), but also extreme cold (-40 ˚C) and salinity that is akin to those on icy moons (Cosciotti *et al.*, 2019). FARLIP-capable *Chroococcidiopsis* species are therefore of interest as model organism candidates. This additionally highlights the importance of considering species-specific tolerances and performing experimental analysis to confirm tolerance against extremes.

All LLAC in Fig. 15, with exception of *Chroococcidiopsis cubana* CCALA 043*, Chroococcidiopsis sp*. FACHB-1243 and *Scytonema millei* VB511283, displayed genes necessary for nitrogen and carbon fixation. However, as only a few marker genes were selected for each trait, it is possible other nitrogen fixation genes present were missed, or genome incompleteness may have caused a lack of detection. Nonetheless, the LLAC positive for carbon and nitrogen fixation genes have been isolated from highly diverse ecological backgrounds, including hypersaline, brackish, marine, freshwater, hot spring and terrestrial environments. This could indicate an application of these LLAC in a range of ecological BLSS. 26 LLAC showed genes necessary for hydrogen production, 19 of which also show genes for nitrogen and carbon fixation, sucrose synthesis and/or trehalose synthesis. These cyanobacteria therefore show potential as multi-functional tools; utilizing fewer cyanobacteria for multiple functions would provide a space and energy saving solution, due to the reduced number of different growth chambers, and growth media, required.

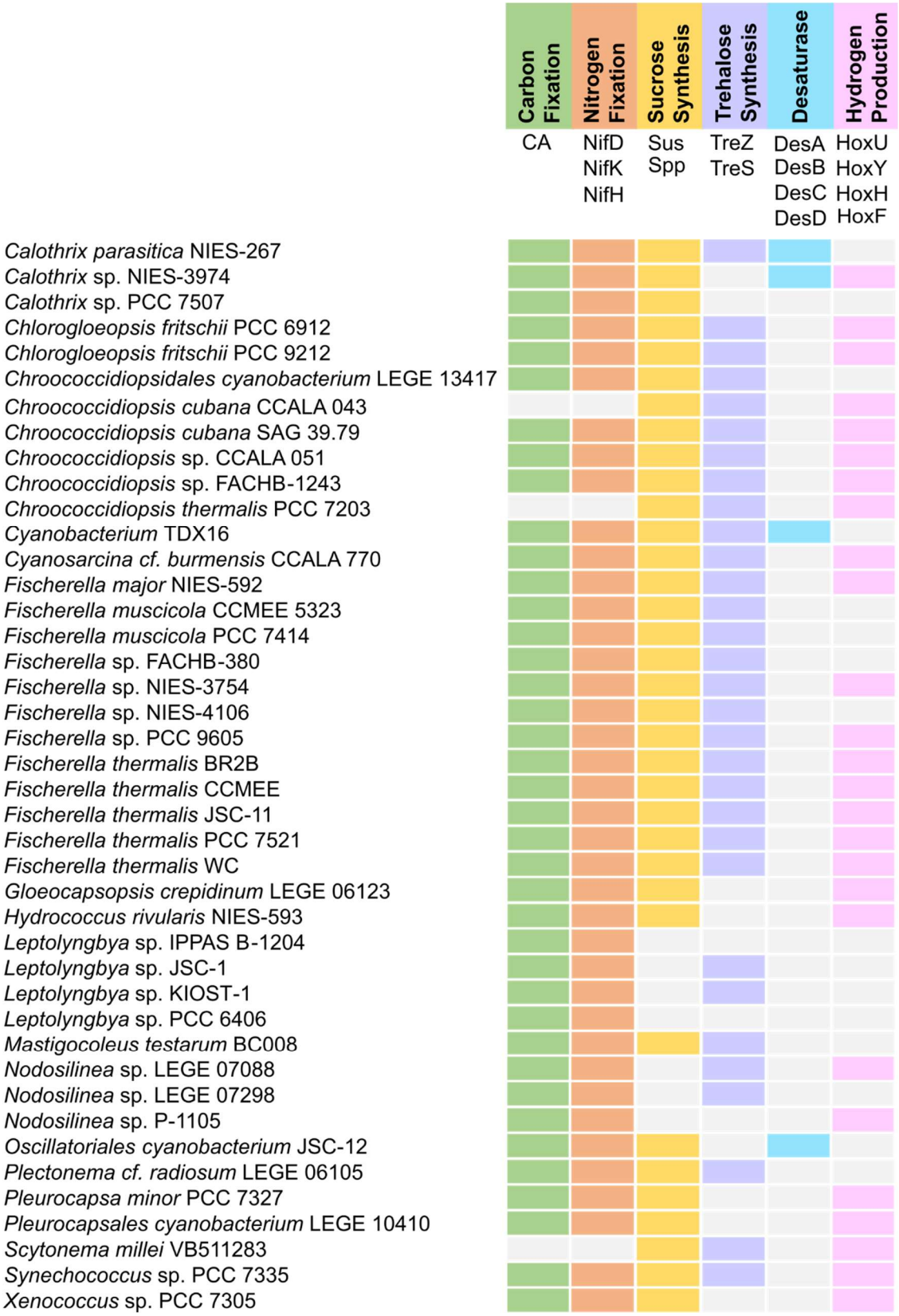


Figure 15 | BLAST analysis of astrobiology-relevant traits in LLAC. Reference proteins for each trait are listed. Proteins utilized for BLAST search and full results can be found in Supplementary Table S9. Dots indicate the LLAC is likely capable of the indicated trait. BLAST hits on all trait reference proteins was required in order to

characterise a LLAC as containing such trait. CA, carbonic anhydrase; NifH, nitrogenase; NifD, nitrogenase molybdenum-iron protein alpha chain; NifK, nitrogenase molybdenum-iron protein beta chain; Sus, sucrose synthase; Spp, sucrose-phosphate synthase; TreY, maltooligosyl trehalose synthase; TreZ, maltooligosyltrehalose trehalohydrolase; TreS, trehalose synthase; DesA, Δ12 desaturase; DesB, ω3 desaturase; DesC, Δ9 desaturase; DesD, Δ6 desaturase; HoxF, hydrogenase; HoxU, biodirectional hydrogenase; HoxY, hydrogenase small subunit; HoxH, NAD-reducing hydrogenase

It should be noted this is by no means an exhaustive list of genetic elements that contribute to each trait, particularly cold survival in cyanobacteria as this is an area not fully understood. For example, exopolysaccharide genes have not been explored here, yet provide tolerance against a range of external stress (Chrismas *et al.*, 2016a; 2016b). Furthermore, the presence of such genes also does not predict or certify a cyanobacteria as being cold-tolerant or capable of producing a certain product. Confirmation of low-light tolerance, cold tolerance assays and hydrogen production efficiency are required to further assess these LLAC as candidates for astrobiology applications.

5. Discussion

5.1 Evaluation of tools for LLAC discovery

5.1.1 FARLIP discovery tools

Previously, FARLIP-capable cyanobacteria had been identified through experimentation and comparative genomics (Gan *et al.*, 2014; Gan, Shen and Bryant, 2014; Gan and Bryant, 2015). Recently, utilisation of ApcE2 in a blastp search against the JGI/IMG database had been successful in identifying FARLIP-capable cyanobacteria from gene fragments and metagenomic data (Antonaru *et al.*, 2020). BLAST analysis has also been utilised to identify FARLIP-capable cyanobacteria among a low number of terrestrial genomes (Chen *et al.*, 2020). In the present study, expansion of this BLAST analysis to two large genome databases was effective in identifying 29 cyanobacteria not previously reported as containing the FALRIP gene cluster. The use of ApcE2 further proved to be a fast and simple tool in identifying FARLIP cluster-containing organisms, and additionally was effective for quickly locating the FARLIP cluster bioinformatically within the genome using either JGI/IMG or NCBI databases. The discovery of new FARLIP-capable cyanobacteria in this study highlights the need to revise bioinformatic results in order to meet the continually updated genome and sequence information available.

5.1.2 LOLIP discovery tools

The JGI/IMG Cassette Search tool provided a fast and efficient way to scan thousands of cyanobacterial genes for LOLIP clusters. It was not possible to perform BLAST analysis due to the limited number of confirmed ApcD4 and ApcB3 protein sequences available, and thus the resulting LOLIP-capable taxa returned were limited to those available on JGI/IMG by use of the JGI/IMG Cassette Search tool. However, a BLAST using any database could now be conducted using the sequences located by our search. Further sequence analysis of LOLIP clusters is required to identify alternative gene arrangements and commonly associated genes

so that the full scope of LOLIP-capable cyanobacteria can be identified through BLAST across multiple databases.

The JGI/IMG Cassette Search online tool was also effective in identifying LOLIP cluster-containing cyanobacteria, identifying 15 cyanobacteria not previously reported as LOLIP-capable. However, the JGI/IMG Cassette Search is limited by the minimum number of genes required to complete the search; as no sequence for IsiX is available through JGI/IMG or NCBI, LHCB was selected as the third required gene. This limits recovered taxa to those containing LOLIP clusters with adjacent LHCB. Several cyanobacteria, such as *Pleurocapsales cyanobacterium* LEGE 10410 and *Pleurocapsa* sp. CCALA 161, do not contain LHCB within the LOLIP cluster and it is likely such a variation, among other variations, in LOLIP cluster genes and gene arrangement are not uncommon.

5.2 Identification and comparative genomics of LLAC

The bioinformatic search of FARLIP and LOLIP-capable species utilised here was sufficient for both locating and performing a comparative analysis on the composition and organisation of these clusters. Formative studies previous to this study have utilised wet-lab experimental approaches to isolate, sequence and analyse the contents of FARLIP and LOLIP clusters (Gan, Shen and Bryant, 2014; Gan and Bryant, 2015; Olsen *et al.*, 2015). The experimental confidence provided by these studies was sufficient to proceed with a bioinformatic approach to finding and analysing LLAC in this study. We have deduced the organisation of the FARLIP cluster in nine cyanobacterial strains, and the LOLIP clusters of 28 cyanobacterial strains, entirely through bioinformatics.

This study provides a simplistic methodology to identify LLAC, and in turn LLAC that may be suitable for astrobiological application. For example, identification of organisms that possess tolerance to salt, drought or cold tolerance by presence of compatible solute (Klähn and Hagemann, 2011; Voß *et al.*, 2013; Urrejola *et al.*, 2019) or exopolysaccharide genes (Los and Murata, 1999; Chrismas, Anesio and Sánchez-Baracaldo, 2015; Chrismas *et al.*, 2016b), for example. Such discovered organisms may be suitable for astrobiological research – application of

bioinformatics to identify astrobiology organisms-of-interest was minimally explored in this study, but could be greatly expanded upon.

However, this is only possible provided sequence data is available and accurate and there is a reliable experimental foundation. As a result of technological advancements and accessibility, genomic sequencing has expanded. But, whether identification of extremophiles by genome alone is an effective methodology remains unclear. Analysis of the cold- and desiccation-tolerant genomes did not indicate unique genomic adaptations that were not otherwise present in temperate relatives (Chrismas *et al.*, 2016; Urrejola *et al.*, 2019). This could indicate (i) extreme survival is controlled by transcriptomics to proteomics, and/or (ii) these temperate relatives could likewise tolerate more extreme conditions. Nonetheless, the use of genomics in astrobiology screening seems a promising prospect.

5.3 Significance for future genomic engineering

Identification of relevant LLAC organisms for astrobiology application through bioinformatics is only an initial step. Any candidates determined to show potential for extreme condition tolerance at the genome level would subsequently be tested for this tolerance experimentally. After this, it is likely any organisms used for astrobiological application would be significantly engineered to meet the optimal requirements for growth, transportation and production needs (Olsson-Francis and Cockell, 2010; Verseux *et al.*, 2016; Averesch, 2021). Such synthetic engineering for resource production in space has been considered by NASA (Cumbers and Rothschild, 2010; Menezes *et al.*, 2015; Rothschild, 2019).

If utilised, it is unlikely LLAC would be employed directly; however, FARLIP and/or LOLIP clusters may be extracted from LLAC to build astrobiology-relevant organisms. For example, low-light tolerance may be implanted in biofuel producers such as *Synechococcus* sp. PCC 7002 (Carr, 2019; Desai, 2015) so that biofuel production can be retained when under energy-saving low light.

From our data, it is clear there are conservations that imply necessity, such as the presence of at least 19 of the 20 FARLIP genes, or the presence of *apcD4* and

apcB3 in the case of LOLIP. While experimental validation is required, our data in conjunction with previous experimental data indicates engineering organisms for low-light adaptation would require incorporation of these genes (Gan *et al.*, 2014; Gan, Shen and Bryant, 2014; Antonaru *et al.*, 2020).

Notably however, whether there is preferential use of LOLIP or FARLIP during low-light is yet to be experimentally established. Also, whether there is an efficiency difference between the processes is also requires experimental analysis. Presumably, utilisation of LOLIP has a lower energetic cost due to the synthesis and incorporation of just ApcD4 and ApcB3, instead of the 20 proteins produced for FARLIP. However, due to the use of just two far-red absorbing allophycocyanin, the light capture and transfer process may not be as efficient as it is in FARLIP, which incorporates at least five far-red absorbing allophycocyanin as well as Chl f. However, it was shown incorporation of Chl f causes reduced efficiency in PSII, and lowered energy output in both photosystems (Mascoli, Bersanini and Croce, 2020). As there is no known chlorophyll incorporated into the LOLIP system, it is possible this system is more energy efficient.

Further research is needed to understand the photosynthetic efficiency of LOLIP and how the two processes co-exist and interplay in organisms that have both FARLIP and LOLIP genes. For example, whether a single process is preferred or optimal and the photosynthetic efficiency when both processes are active simultaneously. Such information would provide a strong basis to engineer organisms for optimised growth under low-light.

5.4 LLAC as astrobiology models

The bioinformatic analysis provided in this study has expanded the range of LLAC that could be utilised for astrobiological application, both either as models or tools. The purpose of this analysis was to narrow the extensive, and continually increasing, list of LLAC identified in order to aid the progress of cyanobacterial use in astrobiology and provide a foundation for future research. The utilisation of LLAC as astrobiological models may require that they exhibit traits that facilitate survival in extreme conditions, such as extreme heat, cold, salinity and desiccation, as such

conditions are present on non-Earth planets. Planets that have low-light, or orbiting outside habitable zones, will be less likely to receive solar heat, and as such have a low surface temperature. Results in Fig. 15 indicate several LLAC exhibit traits that can facilitate in cold tolerance, such as sucrose synthesis, trehalose synthesis, and/or desaturase synthesis.

Both TRAPPIST-1e and TRAPPIST-1f are irradiated with low infrared light, are modelled with cold surface climates. As red dwarf star radiation additionally appears to be suitable as photosynthetically active radiation (PAR) (Claudi *et al.*, 2021), organisms such as cyanobacteria, including LLAC, may be present on planets such as TRAPPIST-1e or TRAPPIST-1f. In this study, LLAC *Calothrix parasitica* NIES-267, *Calothrix* sp. NIES-3974, *Chroococcidiopsis thermalis* PCC 7203 and *Plectonema cf. radiosum* LEGE 06105, exhibited genes that aid cold-tolerance on Earth. Organisms such as these may be suitable as basic models for studying the potential of TRAPPIST-1e or TRAPPIST-1f life. It would additionally be of interest to explore whether extremophile *Chroococcidiopsis* strains such as *Chroococcidiopsis* sp. CCMEE 029 or *Chroococcidiopsis* sp. CCMEE 171 likewise exhibit low-light tolerance; these organisms can tolerate sub-freezing temperatures (Cosciotti *et al.*, 2019), and are therefore more suitable candidates for exploring habitability of the TRAPPIST-1f dark face, if low-light tolerance is confirmed.

5.5 LLAC as astrobiology tools

5.5.1 Cyanobacteria for colonisation

LLAC could also have applications as astrobiology tools. Although variable, it is estimated the average mission launch to space costs \$300,000 kg^{-1} (Massa *et al.*, 2007). If humans are to expand onto other planetary bodies, it will be necessary to derive resources from a reliable source. Given the amount of resources required for a human colony to survive, it is not practical or economically viable to send necessities from Earth on a regular basis. As a solution, BLSS have been proposed and are currently under development. These systems utilise biological organisms and processes to mediate atmosphere revitalization, water recycling, organic water recycling and food production, thus negating the need to rely on Earth resources.

These systems consist of interconnected chambers and bioreactors that mediate production of materials and compounds, filtration, substrate supplementation, storage and output (Gitelson, 1992; Gòdia *et al.*, 2002; Tikhomirov *et al.*, 2007; Giacomelli *et al.*, 2012).

However, the bacteria for proposed use in these systems require carbon, nitrogen and oxygen, among other nutrients, in order to repopulate and produce materials. While these materials can be found on other planets, such as Mars, the compounds are not bioavailable. Cyanobacteria have therefore been proposed as a link between BLSS and local resources, allowing BLSS to work independently of materials provided by Earth (Olsson-Francis and Cockell, 2010; Verseux *et al.*, 2016). Cyanobacteria, through nitrogen and carbon fixation, can provide bioavailable nitrogen and carbon, and can produce oxygen through photosynthesis. Additionally, species of cyanobacteria, such as *Chroococcidiopsis,* have been found to survive in simulated icy moon (Cosciotti *et al.*, 2019) and Martian extreme cold (Baqué, Viaggiu, *et al.*, 2013), simulated UVC and Martian radiation (Cockell *et al.*, 2005; Billi *et al.*, 2011; Baqué, De Vera, *et al.*, 2013; Baqué, Viaggiu, *et al.*, 2013; Baqué *et al.*, 2016; Fagliarone *et al.*, 2020) and extreme desiccation (Billi, 2009; Bahl *et al.*, 2011; Billi *et al.*, 2011; Mosca *et al.*, 2019), that would suggest a suitability to survive on the surface of other planetary bodies.

5.5.2 LLAC for colonisation

Chroococcidiopsis species that are low-light tolerant; *Chroococcidiopsis cubana* CCALA 043, *Chroococcidiopsis cubana* sp. SAG 39.79, *Chroococcidiopsis* sp. CCALA 051, *Chroococcidiopsis* sp. FACHB-1243 and *Chroococcidiopsis thermalis* sp. PCC 7203 are all LLAC. The employment of LLAC for BLSS could offer several advantages. On Earth, FRL has been shown to penetrate more deeply into microbial mats and soils (Bliss and Smith, 1985; M. Kühl & Fenchel, 2000; Pierson et al., 1990), a feature that likely is also applicable to near-Earth planets such as Mars. This could suggest LLAC placed on the surface of Mars, but protected from the harsh surface conditions with a thick barrier or shield, could still effectively capture FRL, photosynthesize, and fix carbon and nitrogen while also surviving long-term on the surface. For example, barrier-protected LLAC at the surface could be used to

colonise Martian earth for the growth of crops (Verseux *et al.*, 2016; Pathak *et al.*, 2018; Chittora *et al.*, 2020), or utilised to leach essential metals (Micheletti *et al.*, 2008; Pandey, 2017; Iqbal, Javed and Baig, 2021). This study has indicated several strains of LLAC have genes necessary for nitrogen and carbon fixation (Fig. 15).

5.5.3 LLAC as production tools

Similarly, due to the penetrative properties of FRL, LLAC are more likely to grow and survive more efficiently in crowded BLSS microbial chambers. This could be useful for production of essential materials, such as hydrogen fuel. While hydrogen fuel production will likely be genetically engineered into future candidates, cyanobacteria have been proposed as sources of hydrogen fuel (Dutta *et al.*, 2005; Bolatkhan *et al.*, 2019; Sadvakasova *et al.*, 2020). It was therefore of interest to explore the scope of LLAC in this study which are readily available as candidates for hydrogen fuel production. This study has shown several LLAC are capable of hydrogen production (Fig.14), and likewise show genes for sucrose and trehalose synthesis, such as *Chlorogloeopsis fritschii* sp. PCC 6912, *Chlorogloeopsis fritschii* sp. PCC 9212, *Chroococcidiopsis cubana* SAG 39.79, *Fischerella thermalis* JSC-11, *Pleurocapsales cyanobacterium* LEGE 10410, to name a few examples.

Sucrose may also be utilized as feedstock in bioethanol production, a renewable source of fuel (Smith, 2008; Muktham *et al.*, 2016; Lin, Zhang and Pakrasi, 2020), and trehalose has been proposed in medical applications of the eye (Matsuo, Tsuchida and Morimoto, 2002; Luyckx and Baudouin, 2011). These cyanobacteria could therefore be genetically optimized to produce these compounds, while also being placed in dense microbial chambers to maximize output, without sacrificing access to PAR due to the penetrative properties of FRL.

5.5.4 LLAC as farming tools

LLAC would also be beneficial in farms, particularly those growing plant life. Such plant life may be utilized for food or medicine. The light capture process in LLAC would not be hindered by an abundance of plant canopy, allowing farms to maximize and expand crop growth and height without sacrificing light access to ground

cyanobacteria or requiring a dedicated area for cyanobacterial growth, saving space and resources. This study has additionally shown a wide range of candidate LLAC from different ecosystems, expanding the range of potential ecosystems that could receive nitrogen and carbon bioavailability as a result of LLAC action, while benefitting from the use of organisms that can tolerate crowding or shielding. For example, LLAC could be implanted in a range of different micro- or macrohabitats, such as freshwater farms or chambers growing *Nymphaea* sp., a genus of water lily that has exhibited medicinal properties for diabetes and inflammation (Chin-Lin Hsu, Song-Chwan Fang and Gow-Chin Yen, 2013), or terrestrial soil farms or chambers that support the growth of *Gossypium* plants for the production of cotton, that could be used to construct clothing and household necessities.

5.5.5 Future of LLAC in astrobiology

It should be noted any cyanobacteria implemented for astrobiological purposes will almost definitely be optimised for its intended role with genetic modification. However, if low-light tolerance is a desired trait in organisms other than cyanobacteria, FARLIP and LOLIP genes would need to be extracted from cyanobacteria initially. This study has further clarified the genetic elements that are key in low-light tolerance, however further clarification on the mechanistic details of FARLIP and LOLIP is required to ensure all factors and co-factors are incorporated into other species for efficient capture of FRL. Such details would also allow biological engineers to optimise the low-light tolerance trait, potentially expanding the limits of low-light survival and the environments where low-light tolerant species can be incorporated.

6. Conclusion

6.1 Aims of thesis and findings

Since the discovery of cyanobacterial Chl f (Chen *et al.*, 2012), the possibility of alternate forms of photosynthesis, and the limits of life, have been expanded. While not an extremophile in the traditional sense, low-light adapted cyanobacteria are a form of extremophile given their survival in the absence of light, despite the fact this very light forms the basis of their metabolism and growth. Hundreds of cyanobacteria capable of FARLIP, and less so, LOLIP, have been discovered through wet-lab and bioinformatic approaches (Gan, Shen and Bryant, 2014; Antonaru *et al.*, 2020; Chen *et al.*, 2020). The extremophile nature of LLAC, and their unique approach to photosynthesis, could indicate these organisms are potential models and tools in astrobiology. While the use of cyanobacteria in astrobiology has been suggested (Olsson-Francis and Cockell, 2010; Menezes *et al.*, 2015; Verseux *et al.*, 2016), the application of LLAC has yet to be explored, owing to their recent discovery and lack of information available. As such, the purpose of this study was to increase the knowledge base surrounding LLAC, and suggest astrobiology applications. This was achieved with three core aims:

i) Bioinformatically identify FARLIP and LOLIP species using unique gene identifiers
ii) Confirm the presence of FARLIP and LOLIP gene clusters in newly identified species, and explore the phylogenetics of key genes in these clusters with a Bayesian approach
iii) Speculate, with bioinformatical support, candidate LLAC for astrobiology research, such as those that may be useful as model organisms for astrobiology, or suitable candidates for use within life support systems in space and on other planets

The bioinformatic analysis completed in this study revealed a high diversity of cyanobacteria that appeared to be FARLIP- and/or LOLIP-capable, with ApcE2-positive strains spanning over 20 genera of cyanobacteria, and ApcB3-positive strains spanning over 14 genera. Notably, this analysis discovered five genera of

cyanobacteria not previously reported as ApcE2-containing: *Nodosilinea, Plectonema, Richelia, Phormidesmis* and *Romeria*. Mapping the discovered strains revealed a likewise highly diverse geographic and ecological distribution, with the exception of ApcE2-positive strains, which showed a slight disposition toward hot spring regions. Possible oceanic origins were speculated due to the frequency of coastal-adjacent isolates.

The gene clusters of *Nodosilinea, Plectonema, Richelia, Phormidesmis* and *Romeria* were analysed further to confirm conservation of the 20 gene FARLIP cluster as observed in other genera and species. The 20 gene cluster was present in each of the cyanobacteria analysed, with exception of *Phormidesmis* sp. RL_2_1 which lacked PsbB. There were minor additions in some genera, such as incorporation of *psbO1-psbF2-psbV1* in *Nodosilinea*, *psbO1-psbF2-nblA-psbV1* in *Phormidesmis* sp. RL_2_1, *psbO1-nblA-psbV1* in *Romeria aff. gracilis* LEGE 07310, and two copies of RfpC in *Plectonema cf. radiosum* LEGE 06105 and all *Richelia* strains. Similarly, analysis of the LOLIP cluster in ApcB3-containing species and strains revealed conservation of the core *apcD4-apcB3* genes. Not all cyanobacteria exhibited *isiX*, however this study did reveal that LHCB is also commonly associated with the *apcD4-apcB3-isiX* cluster.,It was also evident that LOLIP clusters fall into one of three genomic compositions: *apcD4-apcB3-lhcb-isiX*, *apcD4-apcB3-lhcb* and *apcD4-apcB3, apcD4-apcB3-lhcb-isiX* being the most common.

When searching for astrobiology candidates among the identified LLAC, 42 showed a strong genetic disposition toward the cold-tolerance, carbon and nitrogen fixation and fuel production traits analysed. Of note, only *Calothrix parasitica* NIES-267, *Calothrix* sp. NIES-3974, and *Plectonema cf. radiosum* LEGE 06105 showed genes associated with cold tolerance. The majority of the LLAC analysed showed genes necessary for carbon and nitrogen fixation. For hydrogen production, 26 LLAC exhibited the necessary genes, 19 of which also show genes for nitrogen and carbon fixation, sucrose synthesis and/or trehalose synthesis. LLAC that showed astrobiology traits spanned a range of ecosystems, including hypersaline, brackish, marine, freshwater, hot spring and terrestrial environments. This could allow LLAC to be employed in a range of different environments for microbe and plant growth,

production of useful materials and fuels and/or as models for a range of different extraterrestrial planets.

6.2 Limitations and next steps

This is the first study to identify and comparatively explore the genetics of LOLIP species, and likewise the first to explore the use LLAC as astrobiology candidates. However, while a wealth of previous work indicates the presence of the full FARLIP or LOLIP gene cluster does correlate with low-light tolerance. Whether the organisms studied are in fact capable of FARLIP or LOLIP still requires wet-lab validation by growing the organisms in far-red and low-light, and applying a spectroscopic analysis of chlorophyll absorbance. This would be particularly insightful for species that show divergences in key genes, such as *Phormidesmis* sp. RL_2_1 which contains no PsbB. Photosynthetic efficiency of FARLIP and LOLIP processes could also be evaluated. This could reveal whether FARLIP or LOLIP is a more efficient and therefore preferable process for incorporation into any future astrobiological tools. It could also reveal how minor changes within the FARLIP or LOLIP clusters, such as incorporation of *psbO1-psbF2-psbV1* in the FARLIP cluster of *Nodosilinea* strains, or the addition or absence of *isiX* in LOLIP clusters, has an influence on low-light photosynthesis.

Similarly, in species capable of both FARLIP and LOLIP, it would be of interest to explore how these two processes interplay – for example, where there is preferential use of one system over the other, or whether both processes are activated simultaneously and utilised in a 1:1 ratio. This is not within the scope of bioinformatic analysis and requires protein analysis. Wet-lab experiments are also required to further clarify the mechanistic details surrounding FARLIP and LOLIP processes; this knowledge is required if we are to optimise and employ these processes in future astrobiological tools. Experimental confirmation of low-light tolerance, cold tolerance assays and hydrogen production efficiency are additionally required to assess identified LLAC as candidates for astrobiology applications.

The full-scope of LOLIP species available on genomic databases has also not been realised due to limitations with the JGI/IMG Cassette Search online tool. However,

these tools have identified a number of LOLIP clusters, and the genes within these clusters identified as *apcB3* or *apcD4* using maximum likelihood trees. As such, where there was a lack of available *apcB3* or *apcD4* gene sequences, the genes identified in this study, listed in Supplementary Table 8, could be utilised in the future as BLAST queries across a range of databases to identify the full scope of LOLIP-capable species available bioinformatically.

6.3 Concluding remarks

This study has shown (a) the comparative genomic tools implemented here were successful as a bioinformatic approach, (b) a 20-gene cluster, with minor variations, is consistent among five new genera of cyanobacteria identified as carrying the FARLIP cluster, and likewise at least a 2-gene cluster is consistent in LOLIP gene clusters, and, finally, (c) LLAC may be useful as astrobiological models to assess the potential for life on planets such as those in the TRAPPIST-1 system, where planets receive low luminosity and infrared light, or be utilised as robust tools in planet colonisation, for example as biofertilizers or sources of fuel within dense plant canopy or microbial chambers. A great deal of research and technological advancements are required before human expansion into space occurs. Nonetheless, LLAC could offer many advantages over traditional cyanobacteria for in-situ resource production and utilisation on other planets, and we propose a deeper look into LLAC as astrobiological models and tools.

References

A, M. H. and El-Desoukey, H. S. (2020) ‘Effect of Nano-Fertilization and Some Bio-Fertilizer on Growth, Yield and Fiber Quality of Egyptian Cotton’, *Annals of Agric. Sci*, 58(1), pp. 35–44. Available at: https://assjm.journals.ekb.eg (Accessed: 4 August 2021).

Abd-Alla, M. H., Mahmoud, A.-L. E. and Issa, A. A. (1994) ‘Cyanobacterial Biofertilizer Improved Growth of Wheat’, *Phyton*, 34, pp. 11–18. Available at: www.biologiezentrum.at (Accessed: 4 August 2021).

Aguilar, P. S., Cronan, J. E. and De Mendoza, D. (1998) ‘A Bacillus subtilis gene induced by cold shock encodes a membrane phospholipid desaturase’, *Journal of Bacteriology*, 180(8), pp. 2194–2200. doi: 10.1128/JB.180.8.2194-2200.1998/ASSET/8F1383EA-1427-4D84-80E7-4DE6DC76BB4A/ASSETS/GRAPHIC/JB0880004006.JPEG.

Anchordoguy, T. J. *et al.* (1987) ‘Modes of interaction of cryoprotectants with membrane phospholipids during freezing’, *Cryobiology*, 24(4), pp. 324–331. doi: 10.1016/0011-2240(87)90036-8.

Antonaru, L. A. *et al.* (2020) ‘Global distribution of a chlorophyll f cyanobacterial marker’, *ISME Journal*, 14(9), pp. 2275–2287. doi: 10.1038/s41396-020-0670-y.

Arkin, A. P. *et al.* (2018) ‘KBase: The United States department of energy systems biology knowledgebase’, *Nature Biotechnology*. Nature Publishing Group, pp. 566–569. doi: 10.1038/nbt.4163.

Art van der Est, *,‡ *et al.* (2004) ‘Removal of PsaF Alters Forward Electron Transfer in Photosystem I: Evidence for Fast Reoxidation of QK-A in Subunit Deletion Mutants of Synechococcus sp. PCC 7002†’, *Biochemistry*, 43(5), pp. 1264–1275. doi: 10.1021/BI035431J.

Artur, S. and Min, C. (2020) ‘Part I: Photosynthesis and Energy Transfer’, in Wang, Q. (ed.) *MICROBIAL PHOTOSYNTHESIS*. 1st edn. Singapore: Springer Singapore, pp. 3–43. doi: 10.1007/978-981-15-3110-1.

Ashkenazy, Y. (2019) ‘The surface temperature of Europa’, *Heliyon*, 5(6), p. e01908. doi: 10.1016/J.HELIYON.2019.E01908.

Averesch, N. J. H. (2021) ‘Choice of Microbial System for In-Situ Resource Utilization on Mars’, *Frontiers in Astronomy and Space Sciences*, 0, p. 116. doi: 10.3389/FSPAS.2021.700370.

Averina, S. *et al.* (2018) ‘Far-red light photoadaptations in aquatic cyanobacteria’, *Hydrobiologia 2018 813:1*, 813(1), pp. 1–17. doi: 10.1007/S10750-018-3519-X.

Averina, S. G. *et al.* (2019) ‘Non-a chlorophylls in cyanobacteria’, *Photosynthetica*, 57(4), pp. 1109–1118. doi: 10.32615/ps.2019.130.

Azua-Bustos, A. *et al.* (2014) ‘Gloeocapsopsis AAB1, an extremely desiccation-tolerant cyanobacterium isolated from the Atacama Desert’, *Extremophiles*, 18(1), pp. 61–74. doi: 10.1007/s00792-013-0592-y.

Bahl, J. *et al.* (2011) ‘Ancient origins determine global biogeography of hot and cold

desert cyanobacteria', *Nature Communications*, 2(1), pp. 1–6. doi: 10.1038/ncomms1167.

Baier, K. *et al.* (2004) 'NblA is essential for phycobilisome degradation in Anabaena sp. strain PCC 7120 but not for development of functional heterocysts', *Microbiology*, 150(8), pp. 2739–2749. doi: 10.1099/mic.0.27153-0.

Bailey, S. and Grossman, A. (2008) 'Photoprotection in cyanobacteria: Regulation of light harvesting', in *Photochemistry and Photobiology*. John Wiley & Sons, Ltd, pp. 1410–1420. doi: 10.1111/j.1751-1097.2008.00453.x.

Baniulis, D. *et al.* (2008) 'Structure-function of the cytochrome b6f complex', in *Photochemistry and Photobiology*. John Wiley & Sons, Ltd, pp. 1349–1358. doi: 10.1111/j.1751-1097.2008.00444.x.

Baqué, M., Viaggiu, E., *et al.* (2013) 'Endurance of the endolithic desert cyanobacterium Chroococcidiopsis under UVC radiation', *Extremophiles*, 17(1), pp. 161–169. doi: 10.1007/s00792-012-0505-5.

Baqué, M., De Vera, J. P., *et al.* (2013) 'The BOSS and BIOMEX space experiments on the EXPOSE-R2 mission: Endurance of the desert cyanobacterium Chroococcidiopsis under simulated space vacuum, Martian atmosphere, UVC radiation and temperature extremes.', *Acta Astronautica*, 91(2013), pp. 180–186. doi: 10.1016/j.actaastro.2013.05.015.

Baqué, M. *et al.* (2016) 'Preservation of Biomarkers from Cyanobacteria Mixed with Mars-Like Regolith Under Simulated Martian Atmosphere and UV Flux', *Origins of Life and Evolution of Biospheres*, 46(2–3), pp. 289–310. doi: 10.1007/s11084-015-9467-9.

Barber, J., Morris, E. and Büchel, C. (2000) 'Revealing the structure of the photosystem II chlorophyll binding proteins, CP43 and CP47', *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1459(2–3), pp. 239–247. doi: 10.1016/S0005-2728(00)00158-4.

Barrera-Rojas, J. *et al.* (2018) 'The distribution of divinyl chlorophylls a and b and the presence of ferredoxin-NADP+ reductase in Prochlorococcus marinus MIT9313 thylakoid membranes', *Heliyon*, 4(12), p. e01100. doi: 10.1016/j.heliyon.2018.e01100.

Basheva, D. *et al.* (2018) 'Content of phycoerythrin, phycocyanin, alophycocyanin and phycoerythrocyanin in some cyanobacterial strains: Applications', *Engineering in Life Sciences*, 18(11), pp. 861–866. doi: 10.1002/elsc.201800035.

Behrendt, L. *et al.* (2020) 'Life in the dark: far-red absorbing cyanobacteria extend photic zones deep into terrestrial caves', *Environmental Microbiology*, 22(3), pp. 952–963. doi: 10.1111/1462-2920.14774.

Berera, R. *et al.* (2010) 'The light-harvesting function of carotenoids in the cyanobacterial stress-inducible IsiA complex', *Chemical Physics*, 373(1–2), pp. 65–70. doi: 10.1016/j.chemphys.2010.01.011.

Bibby, T. S., Nield, J. and Barber, J. (2001) 'Iron deficiency induces the formation of an antenna ring around trimeric photosystem I in cyanobacteria', *Nature 2001 412:6848*, 412(6848), pp. 743–745. doi: 10.1038/35089098.

Billi, D. *et al.* (2000) ‘Ionizing-radiation resistance in the desiccation-tolerant cyanobacterium Chroococcidiopsis’, *Applied and Environmental Microbiology*, 66(4), pp. 1489–1492. doi: 10.1128/AEM.66.4.1489-1492.2000.

Billi, D. (2009) ‘Subcellular integrities in Chroococcidiopsis sp. CCMEE 029 survivors after prolonged desiccation revealed by molecular probes and genome stability assays’, *Extremophiles*, 13(1), pp. 49–57. doi: 10.1007/s00792-008-0196-0.

Billi, D. *et al.* (2011) ‘Damage escape and repair in dried chroococcidiopsis spp. from hot and cold deserts exposed to simulated space and martian conditions’, *Astrobiology*, 11(1), pp. 65–73. doi: 10.1089/ast.2009.0430.

Billi, D. *et al.* (2013) ‘Cyanobacteria from Extreme Deserts to Space *’, *Advances in Microbiology*, 3, pp. 80–86. doi: 10.4236/aim.2013.36A010.

Billi, D. *et al.* (2019) ‘A Desert Cyanobacterium under Simulated Mars-like Conditions in Low Earth Orbit: Implications for the Habitability of Mars’, *Astrobiology*, 19(2), pp. 158–169. doi: 10.1089/ast.2017.1807.

BLISS, D. and SMITH, H. (1985) ‘Penetration of light into soil and its role in the control of seed germination’, *Plant, Cell & Environment*, 8(7), pp. 475–483. doi: 10.1111/j.1365-3040.1985.tb01683.x.

Bolatkhan, K. *et al.* (2019) ‘Hydrogen production from phototrophic microorganisms: Reality and perspectives’, *International Journal of Hydrogen Energy*, 44(12), pp. 5799–5811. doi: 10.1016/J.IJHYDENE.2019.01.092.

Bryant, D. A. (1982) *Phycoerythrocyanin and phycoerythrin: properties and occurrence in cyanobacteria*, *Journal of General Microbiology*. doi: 10.1099/00221287-128-4-835.

Bryant, D. A., Glazer, A. N. and Eiserling, F. A. (1976) ‘Characterization and structural properties of the major biliproteins of Anabaena sp.’, *Archives of Microbiology*, 110(1), pp. 61–75. doi: 10.1007/BF00416970.

Burnap, R. L., Troyan, T. and Sherman, L. A. (1993) ‘The Highly Abundant Chlorophyll-Protein Complex of Iron-Deficient Synechococcus sp. PCC7942 (CP43[prime]) Is Encoded by the isiA Gene’, *Plant Physiology*, 103(3), pp. 893–902. doi: 10.1104/PP.103.3.893.

Cardona, T., Murray, J. W. and Rutherford, A. W. (2015) ‘Origin and evolution of water oxidation before the last common ancestor of the cyanobacteria’, *Molecular Biology and Evolution*, 32(5), pp. 1310–1328. doi: 10.1093/molbev/msv024.

Caturla-Sánchez, E. *et al.* (2018) ‘Vitrification of dog spermatozoa: Effects of two cryoprotectants (sucrose or trehalose) and two warming procedures’, *Cryobiology*, 80, pp. 126–129. doi: 10.1016/J.CRYOBIOL.2017.11.001.

Caumette, P. *et al.* (1994) ‘Microbial mats in the hypersaline ponds of Mediterranean salterns (Salins-de-Giraud, France)’, *FEMS Microbiology Ecology*, 13(4), pp. 273–286. doi: 10.1111/j.1574-6941.1994.tb00074.x.

Chen, M. *et al.* (2010) ‘A red-shifted chlorophyll’, *Science*, 329(5997), pp. 1318–1319. doi: 10.1126/science.1191127.

Chen, M. *et al.* (2012) ‘A cyanobacterium that contains chlorophyll f - A red-

absorbing photopigment', *FEBS Letters*, 586(19), pp. 3249–3254. doi: 10.1016/j.febslet.2012.06.045.

Chen, M. Y. *et al.* (2020) 'Comparative genomics reveals insights into cyanobacterial evolution and habitat adaptation', *ISME Journal*, pp. 1–17. doi: 10.1038/s41396-020-00775-z.

Chen, M., Zhang, Y. and Blankenship, R. E. (2008) 'Nomenclature for membrane-bound light-harvesting complexes of cyanobacteria', in *Photosynthesis Research*. Springer, pp. 147–154. doi: 10.1007/s11120-007-9255-0.

Chikkaswamy, B. K. (2015) 'Effect of Cyanobacterial Biofertilizer on Soil Nutrients and Mulberry Leaf Quality and Its Impact on Silkworm Crops', *International Journal of Advanced Research in Engineering and Applied Sciences Impact Factor: 4*, 817(1). Available at: www.garph.co.uk (Accessed: 4 August 2021).

Chin-Lin Hsu, Song-Chwan Fang and Gow-Chin Yen (2013) 'Anti-inflammatory effects of phenolic compounds isolated from the flowers of Nymphaea mexicana Zucc.', *Food & Function*, 4(8), pp. 1216–1222. doi: 10.1039/C3FO60041F.

Chiş, C. *et al.* (2017) 'Expression of psbA1 gene in Synechocystis sp. PCC 6803 is influenced by CO2', *Open Life Sciences*, 12(1), pp. 156–161. doi: 10.1515/biol-2017-0018.

Chitnis, V. P. and Chitnis, P. R. (1993) 'PsaL subunit is required for the formation of photosystem I trimers in the cyanobacterium Synechocystis sp. PCC 6803', *FEBS Letters*, 336(2), pp. 330–334. doi: 10.1016/0014-5793(93)80831-E.

Chittora, D. *et al.* (2020) 'Cyanobacteria as a source of biofertilizers for sustainable agriculture', *Biochemistry and Biophysics Reports*, 22, p. 100737. doi: 10.1016/J.BBREP.2020.100737.

Chrismas, N. A. M. *et al.* (2016a) 'Genomic mechanisms for cold tolerance and production of exopolysaccharides in the Arctic cyanobacterium Phormidesmis priestleyi BC1401', *BMC Genomics*, 17(1), pp. 1–14. doi: 10.1186/s12864-016-2846-4.

Chrismas, N. A. M. *et al.* (2016b) 'Genomic mechanisms for cold tolerance and production of exopolysaccharides in the Arctic cyanobacterium Phormidesmis priestleyi BC1401', *BMC Genomics*, 17(1), p. 533. doi: 10.1186/s12864-016-2846-4.

Chrismas, N. A. M. *et al.* (2018) 'Photoecology of the Antarctic cyanobacterium Leptolyngbya sp. BC1307 brought to light through community analysis, comparative genomics and in vitro photophysiology', *Molecular Ecology*, 27(24), pp. 5279–5293. doi: 10.1111/mec.14953.

Chrismas, N. A. M., Anesio, A. M. and Śanchez-Baracaldo, P. (2018) 'The future of genomics in polar and alpine cyanobacteria', *FEMS Microbiology Ecology*. Oxford University Press. doi: 10.1093/femsec/fiy032.

Chrismas, N., Anesio, A. M. and Sánchez-Baracaldo, P. (2015) 'Multiple adaptations to polar and alpine environments within cyanobacteria: A phylogenomic and Bayesian approach', *Frontiers in Microbiology*, 6(OCT), p. 1070. doi: 10.3389/fmicb.2015.01070.

Cinque, G., Croce, R. and Bassi, R. (2000) 'Absorption spectra of chlorophyll a and b

in Lhcb protein environment', *Photosynthesis Research 2000 64:2*, 64(2), pp. 233–242. doi: 10.1023/A:1006467617697.

Claudi, R. *et al.* (2021) 'Super-earths, m dwarfs, and photosynthetic organisms: Habitability in the lab', *Life*, 11(1), pp. 1–18. doi: 10.3390/life11010010.

Cockell, C. S. *et al.* (2005) 'Effects of a simulated Martian UV flux on the cyanobacterium, Chroococcidiopsis sp. 029', *Astrobiology*, 5(2), pp. 127–140. doi: 10.1089/ast.2005.5.127.

Cockell, C. S. *et al.* (2020) 'Space station biomining experiment demonstrates rare earth element extraction in microgravity and Mars gravity', *Nature Communications 2020 11:1*, 11(1), pp. 1–11. doi: 10.1038/s41467-020-19276-w.

Constant, S. *et al.* (1997) 'Expression of the psbA gene during photoinhibition and recovery in Synechocystis PCC 6714: Inhibition and damage of transcriptional and translational machinery prevent the restoration of photosystem II activity', *Plant Molecular Biology*, 34(1), pp. 1–13. doi: 10.1023/A:1005754823218.

Cosciotti, B. *et al.* (2019) 'Survivability of anhydrobiotic cyanobacteria in salty ice: implications for the habitability of icy worlds', *Life*, 9(4). doi: 10.3390/life9040086.

Cumbers, J. and Rothschild, L. J. (2010) 'BISRU: Synthetic Microbes for Moon, Mars and Beyond', in *Astrobiology Science Conference 2010, LPI Contribution No. 1538*. League City, Texas, p. 5672. Available at: http://adsabs.harvard.edu/abs/2010LPICo1538.5672C (Accessed: 29 September 2021).

Dressing, C. D. and Charbonneau, D. (2015) 'The occurrence of potentially habitable planets orbiting m dwarfs estimated from the full kepler dataset and an empirical measurement of the detection sensitivity', *The Astrophysical Journal*, 807(1), p. 45. doi: 10.1088/0004-637X/807/1/45.

DT, J., WR, T. and JM, T. (1992) 'The rapid generation of mutation data matrices from protein sequences', *Computer applications in the biosciences : CABIOS*, 8(3), pp. 275–282. doi: 10.1093/BIOINFORMATICS/8.3.275.

Duarte, C. M. (2012) 'How Cyanobacteria Made Planet Earth Habitable', in Carlos M. Duarte (ed.) *The Role of Marine Biota in the Functioning of the Biosphere*. First. Spain: Fundación BBVA, pp. 55–70.

Dutta, D. *et al.* (2005) 'Hydrogen production by Cyanobacteria', *Microbial Cell Factories 2005 4:1*, 4(1), pp. 1–11. doi: 10.1186/1475-2859-4-36.

Eaton-Rye, J. J. and Vermaas, W. F. J. (1991) 'Oligonucleotide-directed mutagenesis of psbB, the gene encoding CP47, employing a deletion mutant strain of the cyanobacterium Synechocystis sp. PCC 6803', *Plant Molecular Biology*, 17, pp. 1165–1177.

Fagliarone, C. *et al.* (2020) 'Biomarker Preservation and Survivability Under Extreme Dryness and Mars-Like UV Flux of a Desert Cyanobacterium Capable of Trehalose and Sucrose Accumulation', *Frontiers in Astronomy and Space Sciences*, 7. doi: 10.3389/fspas.2020.00031.

Flaim, G. *et al.* (2014) 'Temperature-induced changes in lipid biomarkers and mycosporine-like amino acids in the psychrophilic dinoflagellate *Peridinium*

aciculiferum', *Freshwater Biology*, 59(5), pp. 985–997. doi: 10.1111/fwb.12321.

Flombaum, P. *et al.* (2013) '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), pp. 9824–9829. doi: 10.1073/PNAS.1307701110/SUPPL_FILE/PNAS.201307701SI.PDF.

Gabdulkhakov, A. G. and Dontsova, M. V. (2013) 'Structural studies on photosystem ii of cyanobacteria', *Biochemistry (Moscow)*. Maik Nauka Publishing / Springer SBM, pp. 1524–1538. doi: 10.1134/S0006297913130105.

Gan, F. *et al.* (2014) 'Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light', *Science*, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.

Gan, F. and Bryant, D. A. (2015) 'Adaptive and acclimative responses of cyanobacteria to far-red light', *Environmental Microbiology*, 17(10), pp. 3450–3465. doi: 10.1111/1462-2920.12992.

Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLiP) in Diverse Cyanobacteria', *Life*, 5(1), pp. 4–24. doi: 10.3390/life5010004.

Gantt, E. and Lipschultz, C. A. (1973) 'Energy transfer in phycobilisomes from phycoerythrin to allophycocyanin', *BBA - Bioenergetics*, 292(3), pp. 858–861. doi: 10.1016/0005-2728(73)90036-4.

Gao, J. *et al.* (2018) 'Structure and function of the photosystem supercomplexes', *Frontiers in Plant Science*. Frontiers Media S.A., p. 357. doi: 10.3389/fpls.2018.00357.

Gasteiger, E. *et al.* (2003) 'ExPASy: The proteomics server for in-depth protein knowledge and analysis', *Nucleic Acids Research*, 31(13), pp. 3784–3788. doi: 10.1093/nar/gkg563.

Giacomelli, G. A. *et al.* (2012) 'Bio-regenerative life support system development for Lunar/Mars habitats', *42nd International Conference on Environmental Systems 2012, ICES 2012*. doi: 10.2514/6.2012-3463.

Gilchrist, C. L. M. and Chooi, Y. H. (2020) 'clinker & clustermap.js: Automatic generation of gene cluster comparison figures', *bioRxiv*. bioRxiv, p. 2020.11.08.370650. doi: 10.1101/2020.11.08.370650.

Gillon, M. *et al.* (2016) 'Temperate Earth-sized planets transiting a nearby ultracool dwarf star', *Nature*, 533(7602), pp. 221–224. doi: 10.1038/nature17448.

Gillon, M. *et al.* (2017) 'Seven temperate terrestrial planets around the nearby ultracool dwarf star TRAPPIST-1', *Nature*, 542(7642), pp. 456–460. doi: 10.1038/nature21360.

Gisriel, C. *et al.* (2020) 'The structure of Photosystem I acclimated to far-red light illuminates an ecologically important acclimation process in photosynthesis', *Science Advances*, 6(6), p. eaay6415. doi: 10.1126/sciadv.aay6415.

Gitelson, J. I. (1992) 'Biological life-support systems for Mars mission', *Advances in Space Research*, 12(5), pp. 167–192. doi: 10.1016/0273-1177(92)90023-Q.

Glazer, A. N. (1976) ‘Phycocyanins: Structure and Function’, in *Photochemical and Photobiological Reviews*. Springer US, pp. 71–115. doi: 10.1007/978-1-4684-2574-1_2.

Glazer, A. N. (1985) *Light Harvesting by Phycobilisomes*, *Ann. Rev. Biophys. Biophys. Chern*.

Glazer, A. N. and Bryant, D. A. (1975) ‘Allophycocyanin B (λmax 671, 618 nm) - A new cyanobacterial phycobiliprotein’, *Archives of Microbiology*, 104(1), pp. 15–22. doi: 10.1007/BF00447294.

Glazer, A. N. and Hixson, C. S. (1975) *Characterization of R-phycocyanin. Chromophore content of R-phycocyanin and C-phycoerythrin.*, *Journal of Biological Chemistry*. doi: 10.1016/S0021-9258(19)41208-8.

Gobets, B. and Van Grondelle, R. (2001) ‘Energy transfer and trapping in photosystem I’, *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1507(1–3), pp. 80–99. doi: 10.1016/S0005-2728(01)00203-1.

Gòdia, F. *et al.* (2002) ‘MELISSA: a loop of interconnected bioreactors to develop life support in Space’, *Journal of Biotechnology*, 99(3), pp. 319–330. doi: 10.1016/S0168-1656(02)00222-5.

Golden, S. S., Brusslan, J. and Haselkorn, R. (1986) ‘Expression of a family of psbA genes encoding a photosystem II polypeptide in the cyanobacterium Anacystis nidulans R2.’, *The EMBO journal*, 5(11), pp. 2789–2798. doi: 10.1002/j.1460-2075.1986.tb04569.x.

Govindjee and Shevela, D. (2011) ‘Adventures with cyanobacteria: A personal perspective’, *Frontiers in Plant Science*, 2(JUL), p. 28. doi: 10.3389/fpls.2011.00028.

Grewal, P. S. and Jagdale, G. B. (2010) ‘Enhanced Trehalose Accumulation and Desiccation Survival of Entomopathogenic Nematodes Through Cold Preacclimation’, *http://dx.doi.org/10.1080/09583150210000162O7*, 12(5), pp. 533–545. doi: 10.1080/09583150210000162O7.

Grotjohann, I. and Fromme, P. (2005) ‘Structure of cyanobacterial Photosystem I’, *Photosynthesis Research*, 85, pp. 51–72.

Gurrieri, L. *et al.* (2021) ‘Calvin–Benson cycle regulation is getting complex’, *Trends in Plant Science*. Elsevier Ltd. doi: 10.1016/j.tplants.2021.03.008.

Guskov, A. *et al.* (2009) ‘Cyanobacterial photosystem II at 2.9-Å resolution and the role of quinones, lipids, channels and chloride’, *Nature Structural and Molecular Biology*, 16(3), pp. 334–342. doi: 10.1038/nsmb.1559.

Guy, C. L., Huber, J. L. A. and Huber, S. C. (1992) ‘Sucrose Phosphate Synthase and Sucrose Accumulation at Low Temperature’, *Plant Physiology*, 100(1), p. 502. doi: 10.1104/PP.100.1.502.

Halfmann, C., Gu, L. and Zhou, R. (2014) ‘Engineering cyanobacteria for the production of a cyclic hydrocarbon fuel from CO 2 and H 2 O’, *Green Chemistry*, 16(6), pp. 3175–3185. doi: 10.1039/C3GC42591F.

Hallenbeck, P. C. (2012) ‘Hydrogen Production by Cyanobacteria’, *Microbial Technologies in Advanced Biofuels Production*, 9781461412083, pp. 15–28. doi:

10.1007/978-1-4614-1208-3_2.

Hamaguchi, T. *et al.* (2021) ‘Structure of the far-red light utilizing photosystem I of Acaryochloris marina’, *Nature Communications*, 12(1), p. 2333. doi: 10.1038/s41467-021-22502-8.

Hankamer, B. *et al.* (2001) ‘Subunit positioning and transmembrane helix organisation in the core dimer of photosystem II’, in *FEBS Letters*. No longer published by Elsevier, pp. 142–151. doi: 10.1016/S0014-5793(01)02766-1.

Hassler, D. M. *et al.* (2014) ‘Mars’ Surface Radiation Environment Measured with the Mars Science Laboratory’s Curiosity Rover’, *Science*, 343(6169). doi: 10.1126/SCIENCE.1244797.

Heath, M. J. *et al.* (1999) ‘Habitability of Planets Around Red Dwarf Stars’, *Origins of life and evolution of the biosphere*, 29, pp. 405–424.

Hernández-Mariné, M., Turon, X. and Catalan, J. (2019) ‘A marine Synechocystis (Chroococcales, Cyanophyta) epizoic on didemnid ascidians from the Mediterranean Sea’, *Phycologia*, 29(3), pp. 275–284. doi: 10.2216/I0031-8884-29-3-275.1.

Herrera-Salgado, P. *et al.* (2018) ‘Complementary chromatic and far-red photoacclimations in Synechococcus ATCC 29403 (PCC 7335). I: The phycobilisomes, a proteomic approach’, *Photosynthesis Research 2018 138:1*, 138(1), pp. 39–56. doi: 10.1007/S11120-018-0536-6.

Hird, S. M. *et al.* (1991) ‘Differential expression of the psbB and psbH genes encoding the 47 kDa chlorophyll a-protein and the 10 kDa phosphoprotein of photosystem II during chloroplast development in wheat’, *Current Genetics 1991 19:3*, 19(3), pp. 199–206. doi: 10.1007/BF00336487.

Ho, M.-Y. *et al.* (no date) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335. II.Characterization of phycobiliproteins produced during acclimation to far-red light’. doi: 10.1007/s11120-016-0303-5.

Ho, M.-Y. and Bryant, D. A. (2020) ‘Long Wavelength Pigments in Photosynthesis’, in *Reference Module in Life Sciences*. Elsevier, pp. 1–11. doi: 10.1016/b978-0-12-819460-7.00009-8.

Ho, M. Y. *et al.* (2016) ‘Light-dependent chlorophyll f synthase is a highly divergent paralog of PsbA of photosystem II’, *Science*, 353(6302). doi: 10.1126/science.aaf9178.

Ho, M. Y. *et al.* (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, *Photosynthesis Research*, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.

Ho, M. Y. *et al.* (2020) ‘Extensive remodeling of the photosynthetic apparatus alters energy transfer among photosynthetic complexes when cyanobacteria acclimate to far-red light’, *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1861(4), p. 148064. doi: 10.1016/J.BBABIO.2019.148064.

Holzwarth, A. R. *et al.* (2006) ‘Kinetics and mechanism of electron transfer in intact photosystem II and in the isolated reaction center: Pheophytin is the primary electron acceptor’, *Proceedings of the National Academy of Sciences*, 103(18), pp. 6895–6900. doi: 10.1073/PNAS.0505371103.

Hussain, A. and Hasnain, S. (2011) ‘Phytostimulation and biofertilization in wheat by cyanobacteria’, *Journal of Industrial Microbiology and Biotechnology*, 38(1), pp. 85–92. doi: 10.1007/S10295-010-0833-3.

Iqbal, J., Javed, A. and Baig, M. A. (2021) ‘Heavy metals removal from dumpsite leachate by algae and cyanobacteria’, *Bioremediation Journal*. doi: 10.1080/10889868.2021.1884530.

Jain, I. P. (2009) ‘Hydrogen the fuel for 21st century’, *International Journal of Hydrogen Energy*, 34(17), pp. 7368–7378. doi: 10.1016/J.IJHYDENE.2009.05.093.

Jansson, C. *et al.* (1987) ‘ Construction of an Obligate Photoheterotrophic Mutant of the Cyanobacterium Synechocystis 6803 ’, *Plant Physiology*, 85(4), pp. 1021–1025. doi: 10.1104/pp.85.4.1021.

Jordan, P. *et al.* (2001) ‘Three-dimensional structure of cyanobaoterial photosystem I at 2.5 Å resolution’, *Nature*, 411(6840), pp. 909–917. doi: 10.1038/35082000.

Judd, M. *et al.* (2020) ‘The primary donor of far-red photosystem II: ChlD1 or PD2?’, *Biochimica et Biophysica Acta - Bioenergetics*, 1861(10), p. 148248. doi: 10.1016/j.bbabio.2020.148248.

Kämäräinen, J. *et al.* (2012) ‘Physiological tolerance and stoichiometric potential of cyanobacteria for hydrocarbon fuel production’, *Journal of biotechnology*, 162(1), pp. 67–74. doi: 10.1016/J.JBIOTEC.2012.07.193.

Kandror, O., DeLeon, A. and Goldberg, A. L. (2002) ‘Trehalose synthesis is induced upon exposure of Escherichia coli to cold and is essential for viability at low temperatures’, *Proceedings of the National Academy of Sciences*, 99(15), pp. 9727–9732. doi: 10.1073/PNAS.142314099.

Karapetyan, N. V. *et al.* (2014) ‘Long-wavelength chlorophylls in photosystem I of cyanobacteria: Origin, localization, and functions’, *Biochemistry (Moscow) 2014 79:3*, 79(3), pp. 213–220. doi: 10.1134/S0006297914030067.

Karapetyan, N. V., Holzwarth, A. R. and Rögner, M. (1999) ‘The photosystem I trimer of cyanobacteria: molecular organization, excitation dynamics and physiological significance’, *FEBS Letters*, 460(3), pp. 395–400. doi: 10.1016/S0014-5793(99)01352-6.

Karradt, A. *et al.* (2008) ‘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), pp. 32394–32403. doi: 10.1074/jbc.M805823200.

Kaštovský, J. *et al.* (2014) ‘Cyanocohniella calida gen. et sp. nov. (cyanobacteria: Aphanizomenonaceae) a new cyanobacterium from the thermal springs from Karlovy Vary, Czech Republic’, *Phytotaxa*, 181(5), pp. 279–292. doi: 10.11646/phytotaxa.181.5.3.

Kato, K. *et al.* (2020) ‘Structural basis for the adaptation and function of chlorophyll f in photosystem I’, *Nature Communications 2020 11:1*, 11(1), pp. 1–10. doi: 10.1038/s41467-019-13898-5.

Katoh, H. *et al.* (2001) ‘Functional Analysis of psbV and a Novel c-type Cytochrome Gene psbV2 of the Thermophilic Cyanobacterium Thermosynechococcuselongatus Strain BP-1’, *Plant and Cell Physiology*, 42(6), pp. 599–607. doi:

10.1093/PCP/PCE074.

Katoh, K. and Standley, D. M. (2013) ‘MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability’, *Molecular Biology and Evolution*, 30(4), pp. 772–780. doi: 10.1093/molbev/mst010.

Keisuke Kawashima and Hiroshi Ishikita (2018) ‘Energetic insights into two electron transfer pathways in light-driven energy-converting enzymes’, *Chemical Science*, 9(17), pp. 4083–4092. doi: 10.1039/C8SC00424B.

Khani, A., Moharramipour, S. and Barzegar, M. (2013) ‘Cold tolerance and trehalose accumulation in overwintering larvae of the codling moth, Cydia pomonella (Lepidoptera: Tortricidae)’, *http://www.eje.cz/doi/10.14411/eje.2007.057.html*, 104(3), pp. 385–392. doi: 10.14411/EJE.2007.057.

Kirkwood, A. E. *et al.* (2008) ‘Cyanobacterial Diversity and Halotolerance in a Variable Hypersaline Environment’. doi: 10.1007/s00248-007-9291-5.

Klähn, S. and Hagemann, M. (2011) ‘Compatible solute biosynthesis in cyanobacteria’, *Environmental Microbiology*, 13(3), pp. 551–562. doi: 10.1111/j.1462-2920.2010.02366.x.

Kobayashi, M. *et al.* (2013) ‘Physicochemical Properties of Chlorophylls in Oxygenic Photosynthesis — Succession of Co-Factors from Anoxygenic to Oxygenic Photosynthesis’, *Photosynthesis*. doi: 10.5772/55460.

Kolbe, K., Lechtenböhmer, S. and Fischedick, M. (2019) ‘Hydrogen derived from algae and cyanobacteria as a decentralized fueling option for hydrogen powered cars: Size, space, and cost characteristics of potential bioreactors’, *International Journal of Sustainable Transportation*, 14(5), pp. 325–334. doi: 10.1080/15568318.2018.1547935.

Kós, P. B. *et al.* (2008) ‘Differential regulation of psbA and psbD gene expression, and the role of the different D1 protein copies in the cyanobacterium Thermosynechococcus elongatus BP-1’, *Biochimica et Biophysica Acta - Bioenergetics*, 1777(1), pp. 74–83. doi: 10.1016/j.bbabio.2007.10.015.

Kühl, M. *et al.* (2005) ‘Ecology: A niche for cyanobacteria containing chlorophyll d’, *Nature*, 433(7028), p. 820. doi: 10.1038/433820a.

Kühl, M. and Fenchel, T. (2000) ‘Bio-optical characteristics and the vertical distribution of photosynthetic pigments and photosynthesis in an artificial cyanobacterial mat’, *Microbial Ecology*, 40(2), pp. 94–103. doi: 10.1007/s002480000061.

Kusama, Y. *et al.* (2014) ‘Zeaxanthin and echinenone protect the repair of Photosystem II from inhibition by singlet oxygen in synechocystis sp. PCC 6803’, *Plant and Cell Physiology*, 56(5), pp. 906–916. doi: 10.1093/pcp/pcv018.

Lalić, D. *et al.* (2020) ‘Potential of cyanobacterial secondary metabolites as biomarkers for paleoclimate reconstruction’, *Catena*, 185. doi: 10.1016/j.catena.2019.104283.

Larkum, A. W. *et al.* (1994) ‘Light-harvesting chlorophyll c-like pigment in Prochloron.’, *Proceedings of the National Academy of Sciences*, 91(2), pp. 679–683. doi: 10.1073/PNAS.91.2.679.

De Las Rivas, J. and Barber, J. (2004) ‘Analysis of the Structure of the PsbO Protein and its Implications’, *Photosynthesis Research 2004 81:3*, 81(3), pp. 329–343. doi: 10.1023/B:PRES.0000036889.44048.E4.

Lauro, S. E. *et al.* (2020) ‘Multiple subglacial water bodies below the south pole of Mars unveiled by new MARSIS data’, *Nature Astronomy*, pp. 1–8. doi: 10.1038/s41550-020-1200-6.

Ley, A. C. *et al.* (1977) ‘ Isolation and Function of Allophycocyanin B of Porphyridium cruentum ’, *Plant Physiology*, 59(5), pp. 974–980. doi: 10.1104/pp.59.5.974.

Li, M. *et al.* (2019) ‘Physiological and evolutionary implications of tetrameric photosystem I in cyanobacteria’, *Nature Plants 2019 5:12*, 5(12), pp. 1309–1319. doi: 10.1038/s41477-019-0566-x.

Li, Y. *et al.* (2016) ‘Characterization of red-shifted phycobilisomes isolated from the chlorophyll f-containing cyanobacterium Halomicronema hongdechloris’, *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1857(1), pp. 107–114. doi: 10.1016/J.BBABIO.2015.10.009.

Lin, P. C., Zhang, F. and Pakrasi, H. B. (2020) ‘Enhanced production of sucrose in the fast-growing cyanobacterium Synechococcus elongatus UTEX 2973’, *Scientific Reports*, 10(1), pp. 1–8. doi: 10.1038/s41598-019-57319-5.

Lingam, M. and Loeb, A. (2018) ‘Physical constraints on the likelihood of life on exoplanets’, *International Journal of Astrobiology*, 17(2), pp. 116–126. doi: 10.1017/S1473550417000179.

Liu, H. *et al.* (2013) ‘Phycobilisomes supply excitations to both photosystems in a megacomplex in cyanobacteria’, *Science*, 342(6162), pp. 1104–1107. doi: 10.1126/science.1242321.

Liu, L. N. (2016) ‘Distribution and dynamics of electron transport complexes in cyanobacterial thylakoid membranes’, *Biochimica et Biophysica Acta - Bioenergetics*, 1857(3), pp. 256–265. doi: 10.1016/j.bbabio.2015.11.010.

Loll, B. *et al.* (2008) ‘Modeling of variant copies of subunit D1 in the structure of photosystem II from Thermosynechococcus elongatus’, *Biological Chemistry*, 389(5), pp. 609–617. doi: 10.1515/BC.2008.058.

Los, D. A. and Murata, N. (1999) ‘Responses to cold shock in cyanobacteria’, *Journal of Molecular Microbiology and Biotechnology*, 1(2), pp. 221–230.

Luyckx, J. and Baudouin, C. (2011) ‘Trehalose: an intriguing disaccharide with potential for medical application in ophthalmology’, *Clinical Ophthalmology (Auckland, N.Z.)*, 5(1), p. 577. doi: 10.2147/OPTH.S18827.

MacColl, R. (1998) ‘Cyanobacterial phycobilisomes’, *Journal of Structural Biology*, 124(2–3), pp. 311–334. doi: 10.1006/jsbi.1998.4062.

Macdowall, F. D., Bednar, T. and Rosenberg, A. (1968) ‘Conformation dependence of intramolecular energy transfer in phycoerythrin.’, *Proceedings of the National Academy of Sciences of the United States of America*, 59(4), pp. 1356–1363. doi: 10.1073/pnas.59.4.1356.

Mader, H. M. *et al.* (2006) ‘Subsurface ice as a microbial habitat’, *Geology*, 34(3), pp.

169–172. doi: 10.1130/G22096.1.

Markov, S. A. *et al.* (1993) ‘A hollow fibre photobioreactor for continuous production of hydrogen by immobilized cyanobacteria under partial vacuum’, *International Journal of Hydrogen Energy*, 18(11), pp. 901–906. doi: 10.1016/0360-3199(93)90059-J.

Martiny, A. C. *et al.* (2009) ‘Taxonomic resolution, ecotypes and the biogeography of Prochlorococcus’, *Environmental Microbiology*, 11(4), pp. 823–832. doi: 10.1111/J.1462-2920.2008.01803.X.

Mascoli, V., Bersanini, L. and Croce, R. (2020) ‘Far-red absorption and light-use efficiency trade-offs in chlorophyll f photosynthesis’, *Nature Plants*, 6(8), pp. 1044–1053. doi: 10.1038/s41477-020-0718-z.

Massa, G. D. *et al.* (2007) ‘Plant-Growth Lighting For Space Life Support: A Review’, *Gravitational and Space Research*, 19(2), pp. 19–30.

Máté, Z. *et al.* (1998) ‘UV-B-induced differential transcription of psbA genes encoding the D1 protein of photosystem II in the cyanobacterium Synechocystis 6803’, *Journal of Biological Chemistry*, 273(28), pp. 17439–17444. doi: 10.1074/jbc.273.28.17439.

Matsuo, T., Tsuchida, Y. and Morimoto, N. (2002) ‘Trehalose eye drops in the treatment of dry eye syndrome’, *Ophthalmology*, 109(11), pp. 2024–2029. doi: 10.1016/S0161-6420(02)01219-8.

Menezes, A. A. *et al.* (2015) ‘Towards synthetic biological approaches to resource utilization on space missions’, *Journal of The Royal Society Interface*, 12(102). doi: 10.1098/RSIF.2014.0715.

Metz, J., Nixon, P. and Diner, B. (1990) ‘Nucleotide sequence of the psbA3 gene from the cyanobacterium Synechocystis PCC 6803’, *Nucleic Acids Research*. Oxford University Press, p. 6715. doi: 10.1093/nar/18.22.6715.

Miao, D. *et al.* (2016) ‘Adapting photosynthesis to the near-infrared: non-covalent binding of phycocyanobilin provides an extreme spectral red-shift to phycobilisome core-membrane linker from Synechococcus sp. PCC7335’, *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1857(6), pp. 688–694. doi: 10.1016/J.BBABIO.2016.03.033.

Micheletti, E. *et al.* (2008) ‘Selectivity in the heavy metal removal by exopolysaccharide-producing cyanobacteria’, *Journal of Applied Microbiology*, 105(1), pp. 88–94. doi: 10.1111/J.1365-2672.2008.03728.X.

Miller, S. R. *et al.* (2011) ‘Dynamics of Gene Duplication in the Genomes of Chlorophyll d-Producing Cyanobacteria: Implications for the Ecological Niche’, *Genome Biology and Evolution*, 3(1), p. 601. doi: 10.1093/GBE/EVR060.

Miteva, V. (2008) ‘Bacteria in Snow and Glacier Ice’, *Psychrophiles: From Biodiversity to Biotechnology*, pp. 31–50. doi: 10.1007/978-3-540-74335-4_3.

Miyashita, H. *et al.* (1996) ‘Chlorophyll d as a major pigment’, *Nature 1996 383:6599*, 383(6599), pp. 402–402. doi: 10.1038/383402a0.

Mohamed, A. *et al.* (1993) ‘Differential expression of the psbA genes in the

cyanobacterium Synechocystis 6803', *MGG Molecular & General Genetics*, 238(1–2), pp. 161–168. doi: 10.1007/BF00279543.

Mohamed, A. and Jansson, C. (1989) 'Influence of light on accumulation of photosynthesis-specific transcripts in the cyanobacterium Synechocystis 6803', *Plant Molecular Biology*, 13(6), pp. 693–700. doi: 10.1007/BF00016024.

Mosca, C. *et al.* (2019) 'Over-Expression of UV-Damage DNA Repair Genes and Ribonucleic Acid Persistence Contribute to the Resilience of Dried Biofilms of the Desert Cyanobacetrium Chroococcidiopsis Exposed to Mars-Like UV Flux and Long-Term Desiccation', *Frontiers in Microbiology*, 10. doi: 10.3389/fmicb.2019.02312.

Mueller, D. R. *et al.* (2005) 'Extremotrophs, extremophiles and broadband pigmentation strategies in a high arctic ice shelf ecosystem', *FEMS Microbiology Ecology*, 53(1), pp. 73–87. doi: 10.1016/j.femsec.2004.11.001.

Muktham, R. *et al.* (2016) 'A Review on 1st and 2nd Generation Bioethanol Production-Recent Progress', *Journal of Sustainable Bioenergy Systems*, 6(3), pp. 72–92. doi: 10.4236/JSBS.2016.63008.

Mulo, P., Sicora, C. and Aro, E. M. (2009) 'Cyanobacterial psbA gene family: Optimization of oxygenic photosynthesis', *Cellular and Molecular Life Sciences*. Springer, pp. 3697–3710. doi: 10.1007/s00018-009-0103-6.

Murata, N. and Wada, H. (1995) 'Acyl-lipid desaturases and their importance in the tolerance and acclimatization to cold of cyanobacteria.', *Biochemical Journal*, 308(Pt 1), p. 1. doi: 10.1042/BJ3080001.

Murik, O. *et al.* (2017) 'What distinguishes cyanobacteria able to revive after desiccation from those that cannot: the genome aspect', *Environmental Microbiology*, 19(2), pp. 535–550. doi: 10.1111/1462-2920.13486.

Murray, J. W. (2012) 'Sequence variation at the oxygen-evolving centre of photosystem II: A new class of "rogue" cyanobacterial D1 proteins', *Photosynthesis Research*, 110(3), pp. 177–184. doi: 10.1007/s11120-011-9714-5.

Nadeau, T. L. and Castenholz, R. W. (2000) 'Characterization of psychrophilic oscillatorians (cyanobacteria) from antarctic meltwater ponds', *Journal of Phycology*, 36(5), pp. 914–923. doi: 10.1046/j.1529-8817.2000.99201.x.

Nguyen, L.-T. *et al.* (2015) 'IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies', *Molecular Biology and Evolution*, 32(1), pp. 268–274. doi: 10.1093/molbev/msu300.

Niedzwiedzki, D. M. and Blankenship, R. E. (2010) 'Singlet and triplet excited state properties of natural chlorophylls and bacteriochlorophylls', *Photosynthesis Research 2010 106:3*, 106(3), pp. 227–238. doi: 10.1007/S11120-010-9598-9.

Nishida, I. and Murata, N. (1996) 'Chilling sensitivity in plants and cyanobacteria: The crucial contribution of membrane lipids', *Annual Review of Plant Physiology and Plant Molecular Biology*, 47(1), pp. 541–568. doi: 10.1146/annurev.arplant.47.1.541.

Nowack, S. *et al.* (2015) '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). doi:

10.3389/FMICB.2015.00626.

Nürnberg, D. J. *et al.* (2018) ‘Photochemistry beyond the red limit in chlorophyll f – containing photosystems’, *Science*, 1213(6394), pp. 1210–1213.

Ohkubo, S. and Miyashita, H. (2017) ‘A niche for cyanobacteria producing chlorophyll f within a microbial mat’, *ISME Journal*, 11(10), pp. 2368–2378. doi: 10.1038/ismej.2017.98.

Oliver, T. *et al.* (2021) ‘Time-resolved comparative molecular evolution of oxygenic photosynthesis’, *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1862(6), p. 148400. doi: 10.1016/J.BBABIO.2021.148400.

Olsen, M. T. *et al.* (2015) ‘The molecular dimension of microbial species: 3. Comparative genomics of Synechococcus strains with different light responses and in situ diel transcription patterns of associated putative ecotypes in the Mushroom Spring microbial mat’, *Frontiers in Microbiology*, 6(JUN). doi: 10.3389/fmicb.2015.00604.

Olsson-Francis, K. and Cockell, C. S. (2010) ‘Use of cyanobacteria for in-situ resource use in space applications’, *Planetary and Space Science*, 58(10), pp. 1279–1285. doi: 10.1016/J.PSS.2010.05.005.

Osman, M. E. H. *et al.* (2010) ‘Effect of two species of cyanobacteria as biofertilizers on some metabolic activities, growth, and yield of pea plant’, *Biology and Fertility of Soils 2010 46:8*, 46(8), pp. 861–875. doi: 10.1007/S00374-010-0491-7.

Pakrasi, H. B., Williams, J. G. and Arntzen, C. J. (1988) ‘Targeted mutagenesis of the psbE and psbF genes blocks photosynthetic electron transport: evidence for a functional role of cytochrome b559 in photosystem II.’, *The EMBO Journal*, 7(2), pp. 325–332. doi: 10.1002/J.1460-2075.1988.TB02816.X.

Pandey, V. D. (2017) ‘Cyanobacteria-mediated heavy metal remediation’, *Agro-Environmental Sustainability*, 2, pp. 105–121. doi: 10.1007/978-3-319-49727-3_6.

Park, Y.-I. *et al.* (1999) ‘Expression of the isiA gene is essential for the survival of the cyanobacterium Synechococcus sp. PCC 7942 by protecting photosystem II from excess light under iron limitation’, *Molecular Microbiology*, 32(1), pp. 123–129. doi: 10.1046/J.1365-2958.1999.01332.X.

Parrilli, E. *et al.* (2011) ‘Life in icy habitats: new insights supporting panspermia theory’, *Astrochemistry*, 22, pp. 375–383. doi: 10.1007/s12210-011-0136-2.

Pathak, J. *et al.* (2018) ‘Cyanobacterial Farming for Environment Friendly Sustainable Agriculture Practices: Innovations and Perspectives’, *Frontiers in Environmental Science*, 0(FEB), p. 7. doi: 10.3389/FENVS.2018.00007.

Pentecost, A. (2003) ‘Cyanobacteria associated with hot spring travertines’, *Canadian Journal of Earth Sciences*, 40(11), pp. 1447–1457. doi: 10.1139/e03-075.

Phadtare, S. and Inouye, M. (2008) ‘Cold-shock proteins’, *Psychrophiles: From Biodiversity to Biotechnology*, pp. 191–209. doi: 10.1007/978-3-540-74335-4_12/COVER.

Pierson, B. K., Sands, V. M. and Frederick, J. L. (1990) ‘Spectral Irradiance and Distribution of Pigments in a Highly Layered Marine Microbial Mat’, *Applied and*

Environmental Microbiology, 56(8).

Pinevich, A., Velichko, N. and Ivanikova, N. (2012) 'Cyanobacteria of the Genus Prochlorothrix', *Frontiers in Microbiology*, 3(MAY). doi: 10.3389/FMICB.2012.00173.

Prieto-Barajas, C. M., Valencia-Cantero, E. and Santoyo, G. (2018) 'Microbial mat ecosystems: Structure types, functional diversity, and biotechnological application', *Electronic Journal of Biotechnology*, 31, pp. 48–56. doi: 10.1016/J.EJBT.2017.11.001.

Proctor, M. S. *et al.* (2018) 'Plant and algal chlorophyll synthases function in Synechocystis and interact with the YidC/Alb3 membrane insertase', *FEBS Letters*, 592(18), pp. 3062–3073. doi: 10.1002/1873-3468.13222.

Puente-Sánchez, F. *et al.* (2018) 'Viable cyanobacteria in the deep continental subsurface', *PNAS*, 115(42), pp. 10702–10707. doi: 10.1073/pnas.1808176115.

Q, X. *et al.* (1994) 'Targeted deletion of psaJ from the cyanobacterium Synechocystis sp. PCC 6803 indicates structural interactions between the PsaJ and PsaF subunits of photosystem I', *Plant molecular biology*, 26(1), pp. 291–302. doi: 10.1007/BF00039540.

Quesada, A. and Vincent, W. F. (2012) 'Cyanobacteria in the Cryosphere: Snow, Ice and Extreme Cold', in *Ecology of Cyanobacteria II*, pp. 387–399. doi: 10.1007/978-94-007-3855-3.

Ralf, G. and Repeta, D. J. (1992) 'The pigments of Prochlorococcus marinus: The presence of divinylchlorophyll a and b in a marine procaryote', *Limnology and Oceanography*. John Wiley & Sons, Ltd, pp. 425–433. doi: 10.4319/lo.1992.37.2.0425.

Ravnikar, P. D. *et al.* (1989) 'Nucleotide sequence of a second psbA gene from the unicellular cyanobacterium Synechocystis 6803', *Nucleic Acids Research*. Oxford University Press, p. 3991. doi: 10.1093/nar/17.10.3991.

Renger, T. and Schlodder, E. (2008) 'The primary electron donor of photosystem II of the cyanobacterium acaryochloris marina is a chlorophyll d and the water oxidation is driven by a chlorophyll a/chlorophyll d heterodimer', *Journal of Physical Chemistry B*, 112(25), pp. 7351–7354. doi: 10.1021/JP801900E.

Rodrigues, J. P. *et al.* (2008) 'Evaluation of trehalose and sucrose as cryoprotectants for hematopoietic stem cells of umbilical cord blood', *Cryobiology*, 56(2), pp. 144–151. doi: 10.1016/J.CRYOBIOL.2008.01.003.

Ronquist, F. and Huelsenbeck, J. P. (2003) 'MrBayes 3: Bayesian phylogenetic inference under mixed models', *Bioinformatics*, 19(12), pp. 1572–1574. doi: 10.1093/BIOINFORMATICS/BTG180.

Rosgaard, L. *et al.* (2012) 'Bioengineering of carbon fixation, biofuels, and biochemicals in cyanobacteria and plants', *Journal of Biotechnology*, 162(1), pp. 134–147. doi: 10.1016/J.JBIOTEC.2012.05.006.

Rothschild, L. J. (2019) 'The Role of Synthetic Biology in NASA's Missions', in *The Team's Third Synthetic Biology for Defense Workshop*. Arlington, Virginia. Available at: http://2011.igem.org/Team:Brown-Stanford (Accessed: 29 September 2021).

Rozanov, A. S. *et al.* (2017) 'Biodiversity of the microbial mat of the Garga hot spring', *BMC Evolutionary Biology 2017 17:2*, 17(2), pp. 37–49. doi: 10.1186/S12862-017-1106-9.

Sadvakasova, A. K. *et al.* (2020) 'Bioprocesses of hydrogen production by cyanobacteria cells and possible ways to increase their productivity', *Renewable and Sustainable Energy Reviews*, 133, p. 110054. doi: 10.1016/J.RSER.2020.110054.

Sakamoto, T. and Murata, N. (2002) 'Regulation of the desaturation of fatty acids and its role in tolerance to cold and salt stress', *Current Opinion in Microbiology*, 5(2), pp. 206–210. doi: 10.1016/S1369-5274(02)00306-5.

Salih, G. F. and Jansson, C. (1997) 'Activation of the silent psbA1 gene in the cyanobacterium Synechocystis sp strain 6803 produces a novel and functional D1 protein', *Plant Cell*, 9(6), pp. 869–878. doi: 10.1105/tpc.9.6.869.

Sánchez-Baracaldo, P. (2015) 'Origin of marine planktonic cyanobacteria', *Nature Publishing Group*, 5, p. 17418. doi: 10.1038/srep17418.

Sánchez-Baracaldo, P. *et al.* (2022) 'Cyanobacteria and biogeochemical cycles through Earth history', *Trends in Microbiology*, 30(2), pp. 143–157. doi: 10.1016/J.TIM.2021.05.008.

Sánchez-Baracaldo, P. and Cardona, T. (2020) 'On the origin of oxygenic photosynthesis and Cyanobacteria', *New Phytologist*, 225(4), pp. 1440–1446. doi: 10.1111/NPH.16249.

Sander, J. *et al.* (2008) 'Role of the psbA Gene Family of PSII from the Thermophilic Cyanobacterium Thermosynechococcus elongatus', in *Photosynthesis. Energy from the Sun*. Springer Netherlands, pp. 745–748. doi: 10.1007/978-1-4020-6709-9_166.

Sawa, M. *et al.* (2017) 'Electricity generation from digitally printed cyanobacteria', *Nature Communications 2017 8:1*, 8(1), pp. 1–10. doi: 10.1038/s41467-017-01084-4.

Schaefer, M. R. and Golden, S. S. (1989) 'Differential expression of members of a cyanobacterial psbA gene family in response to light.', *Journal of bacteriology*, 171(7), pp. 3973–3981. doi: 10.1128/jb.171.7.3973-3981.1989.

Scherer, S., Almon, H. and Böger, P. (1988) 'Interaction of photosynthesis, respiration and nitrogen fixation in cyanobacteria', *Photosynthesis Research*. Martinus Nijhoff, The Hague/Kluwer Academic Publishers, pp. 95–114. doi: 10.1007/BF00035255.

Schirrmeister, B. E. *et al.* (2013) 'Evolution of multicellularity coincided with increased diversification of cyanobacteria and the Great Oxidation Event', *Proceedings of the National Academy of Sciences*, 110(5), pp. 1791–1796. doi: 10.1073/PNAS.1209927110.

Schirrmeister, B. E., Sanchez-Baracaldo, P. and Wacey, D. (2016) 'Cyanobacterial evolution during the Precambrian', *International Journal of Astrobiology*, 15(3), pp. 187–204. doi: 10.1017/S1473550415000579.

Schluchter, W. M. *et al.* (1996) 'Characterization of psal and psaL Mutants of Synechococcus sp. Strain PCC 7002: A New Model for State Transitions in Cyanobacteria', *Photochemistry and Photobiology*, 64(1), pp. 53–66. doi: 10.1111/J.1751-1097.1996.TB02421.X.

Schopf, J. W. (1993) ‘Microfossils of the early Archean apex chert: New evidence of the antiquity of life’, *Science*, 260(5108), pp. 640–646. doi: 10.1126/science.260.5108.640.

Schwieterman, E. W. *et al.* (2018) ‘Exoplanet Biosignatures: A Review of Remotely Detectable Signs of Life’, *Astrobiology*. Mary Ann Liebert Inc., pp. 663–708. doi: 10.1089/ast.2017.1729.

Sedoud, A. *et al.* (2014) ‘The cyanobacterial photoactive orange carotenoid protein is an excellent singlet oxygen quencher’, *Plant Cell*, 26(4), pp. 1781–1791. doi: 10.1105/tpc.114.123802.

Shalygo, N. *et al.* (2009) ‘Expression of chlorophyll synthase is also involved in feedback-control of chlorophyll biosynthesis’, *Plant Molecular Biology 2009 71:4*, 71(4), pp. 425–436. doi: 10.1007/S11103-009-9532-8.

Shen, G. *et al.* (2019) ‘Characterization of chlorophyll f synthase heterologously produced in Synechococcus sp. PCC 7002’, *Photosynthesis Research*, 140(1), pp. 77–92. doi: 10.1007/s11120-018-00610-9.

Shen, G., Eaton-Rye, J. J. and Vermaas, W. F. J. (1993) ‘Mutation of Histidine Residues in CP47 Leads to Destabilization of the Photosystem II Complex and to Impairment of Light Energy Transfer’, *Biochemistry*, 32, pp. 5109–5115.

Shigeru Itoh, *,‡ *et al.* (2007) ‘Function of Chlorophyll d in Reaction Centers of Photosystems I and II of the Oxygenic Photosynthesis of Acaryochloris marina†’, *Biochemistry*, 46(43), pp. 12473–12481. doi: 10.1021/BI7008085.

Shubin, V. V. *et al.* (1991) ‘Origin of the 77 K variable fluorescence at 758 nm in the cyanobacterium Spirulina platensis’, *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1060(1), pp. 28–36. doi: 10.1016/S0005-2728(05)80115-X.

Shubin, V. V., Bezsmertnaya, I. N. and Karapetyan, N. V. (1992) ‘Isolation from Spirulina membranes of two photosystem I-type complexes, one of which contains chlorophyll responsible for the 77 K fluorescence band at 760 nm’, *FEBS Letters*, 309(3), pp. 340–342. doi: 10.1016/0014-5793(92)80803-O.

Singh, S. P. *et al.* (2010) ‘Photoprotective and biotechnological potentials of cyanobacterial sheath pigment, scytonemin’, *African Journal of Biotechnology*, 9(5), pp. 580–588.

Sinha, R. P. and Häder, D.-P. (2007) ‘UV-protectants in Cyanobacteria’, *Plant Science*, 174(3), pp. 278–289. doi: 10.1016/j.plantsci.2007.12.004.

Smith, A. M. (2008) ‘Prospects for increasing starch and sucrose yields for bioethanol production’, *The Plant Journal*, 54(4), pp. 546–558. doi: 10.1111/J.1365-313X.2008.03468.X.

Sokoloff, P. C. *et al.* (2016) ‘The “Martian” flora: New collections of vascular plants, lichens, fungi, algae, and cyanobacteria from the Mars Desert Research Station, Utah’, *Biodiversity Data Journal*, 4(1). doi: 10.3897/BDJ.4.e8176.

Soulier, N. and Bryant, D. A. (2021) ‘The structural basis of far-red light absorbance by allophycocyanins’, *Photosynthesis Research*, 147(1), pp. 11–26. doi: 10.1007/s11120-020-00787-y.

Soulier, N., Laremore, T. N. and Bryant, D. A. (2020) 'Characterization of cyanobacterial allophycocyanins absorbing far-red light', *Photosynthesis Research*, 145(3), pp. 189–207. doi: 10.1007/s11120-020-00775-2.

Sozer, O. *et al.* (2010) 'Involvement of carotenoids in the synthesis and assembly of protein subunits of photosynthetic reaction centers of synechocystis sp. PCC 6803', *Plant and Cell Physiology*, 51(5), pp. 823–835. doi: 10.1093/pcp/pcq031.

Srivastava, A. K., Rai, A. N. and Neilan, B. A. (2013) *Stress Biology of Cyanobacteria: Molecular Mechanisms to Cellular Responses - Google Books*. 1st edn. Edited by A. K. Srivastava, A. N. Rai, and B. A. Neilan. Boca Raton: CRC Press.

Van Der Staay, G. W. M., Yurkova, N. and Green, B. R. (1998) 'The 38 kDa chlorophyll a/b protein of the prokaryote Prochlorothrix hollandica is encoded by a divergent pcb gene', *Plant Molecular Biology*, 36(5), pp. 709–716. doi: 10.1023/A:1005930210515.

Stewart, R. C. (1997) 'Kinetic characterization of phosphotransfer between CheA and CheY in the bacterial chemotaxis signal transduction pathway', *Biochemistry*, 36(8), pp. 2030–2040. doi: 10.1021/BI962261K.

Stitt, M. and Hurry, V. (2002) 'A plant for all seasons: alterations in photosynthetic carbon metabolism during cold acclimation in Arabidopsis', *Current Opinion in Plant Biology*, 5(3), pp. 199–206. doi: 10.1016/S1369-5266(02)00258-3.

Strauss, G. and Hauser, H. (1986) 'Stabilization of lipid bilayer vesicles by sucrose during freezing', *Proceedings of the National Academy of Sciences of the United States of America*, 83(8), pp. 2422–2426. doi: 10.1073/pnas.83.8.2422.

Strunecký, O. *et al.* (2019) 'High diversity of thermophilic cyanobacteria in Rupite hot spring identified by microscopy, cultivation, single-cell PCR and amplicon sequencing', *Extremophiles*, 23, pp. 35–48. doi: 10.1007/s00792-018-1058-z.

Tahoun, M. *et al.* (2021) 'Chemistry of porphyrins in fossil plants and animals', *RSC Advances*, 11(13), pp. 7552–7563. doi: 10.1039/D0RA10688G.

Teramoto, H., Ono, T. and Minagawa, J. (2001) 'Identification of Lhcb Gene Family Encoding the Light-harvesting Chlorophyll-a/b Proteins of Photosystem II in Chlamydomonas reinhardtii', *Plant and Cell Physiology*, 42(8), pp. 849–856. doi: 10.1093/PCP/PCE115.

Thomas, K. M. (2014) 'Hydrogen adsorption and storage on porous materials', *Handbook of Hydrogen Energy*, 88(1), pp. 707–762. doi: 10.1201/b17226.

Tikhomirov, A. A. *et al.* (2007) 'Biological life support systems for a Mars mission planetary base: Problems and prospects', *Advances in Space Research*, 40(11), pp. 1741–1745. doi: 10.1016/J.ASR.2006.11.009.

Tiwari, O. N. *et al.* (2005) 'Distribution and physiological characterization of cyanobacteria isolated from arid zones of Rajasthan ARID ZONE CYANOBACTERIA 166', *Tropical Ecology*, 46(2), pp. 165–171.

Tóth, T. N. *et al.* (2015) 'Carotenoids are essential for the assembly of cyanobacterial photosynthetic complexes', *Biochimica et Biophysica Acta - Bioenergetics*, 1847(10), pp. 1153–1165. doi: 10.1016/j.bbabio.2015.05.020.

Tria, F. D. K., Landan, G. and Dagan, T. (2017) ‘Phylogenetic rooting using minimal ancestor deviation’, *Nature Ecology & Evolution 2017 1:1*, 1(1), pp. 1–7. doi: 10.1038/s41559-017-0193.

Tros, M. *et al.* (2021) ‘Breaking the Red Limit: Efficient Trapping of Long-Wavelength Excitations in Chlorophyll-f-Containing Photosystem I’, *Chem*, 7(1), pp. 155–173. doi: 10.1016/j.chempr.2020.10.024.

Turbet, M. *et al.* (2020) ‘A Review of Possible Planetary Atmospheres in the TRAPPIST-1 System’, *Space Science Reviews 2020 216:5*, 216(5), pp. 1–48. doi: 10.1007/S11214-020-00719-1.

Tyystjärvi, T. *et al.* (1998) ‘Regulation of psbA Gene Expression in Synechocystis 6803’, in *Photosynthesis: Mechanisms and Effects*. Springer Netherlands, pp. 2909–2912. doi: 10.1007/978-94-011-3953-3_682.

Ulloa, O. *et al.* (2021) ‘The cyanobacterium Prochlorococcus has divergent light-harvesting antennae and may have evolved in a low-oxygen ocean’, *Proceedings of the National Academy of Sciences of the United States of America*, 118(11), p. e2025638118. doi: 10.1073/PNAS.2025638118/SUPPL_FILE/PNAS.2025638118.SAPP.PDF.

Urrejola, C. *et al.* (2019) ‘Genomic Features for Desiccation Tolerance and Sugar Biosynthesis in the Extremophile Gloeocapsopsis sp. UTEX B3054’, *Frontiers in Microbiology*, 10(MAY), p. 950. doi: 10.3389/fmicb.2019.00950.

Vance, S. D. *et al.* (2021) ‘The Salty Secrets of Icy Ocean Worlds’, *Journal of Geophysical Research: Planets*, 126(1), p. e2020JE006736. doi: 10.1029/2020JE006736.

Vavilin, D. *et al.* (2003) ‘Energy and electron transfer in photosystem II of a chlorophyll b-containing Synechocystis sp. PCC 6803 mutant’, *Biochemistry*, 42(6), pp. 1731–1746. doi: 10.1021/bi026853g.

Veeder, G. J. (1974) ‘Luminosities and temperatures of M dwarf stars from infrared photometry’, *The Astronomical Journal*, 79(10), pp. 1056–1072.

Vermaas, J. *et al.* (1987) ‘Sequencing and modification of psbB, the gene encoding the CP-47 protein of Photosystem II, in the cyanobacterium Synechocystis 6803’, *Plant Molecular Biology*, 8, pp. 317–326.

Vermaas, W. F. (2001) ‘Photosynthesis and Respiration in Cyanobacteria’, in *Encyclopedia of Life Sciences*. John Wiley & Sons, Ltd. doi: 10.1038/npg.els.0001670.

Vermaas, W., Ikeuchi, M. and Inoue, Y. (1988) ‘Protein composition of the photosystem II core complex in genetically engineered mutants of the cyanobacterium Synechocystis sp. PCC 6803’, *Photosynthesis research*, 17(1–2), pp. 97–113. doi: 10.1007/BF00047683.

Verseux, C. *et al.* (2016) ‘Sustainable life support on Mars - The potential roles of cyanobacteria’, in *International Journal of Astrobiology*. Cambridge University Press, pp. 65–92. doi: 10.1017/S147355041500021X.

Verseux, C. *et al.* (2017) ‘Evaluation of the resistance of chroococcidiopsis spp. To sparsely and densely ionizing irradiation’, *Astrobiology*, 17(2), pp. 118–125. doi:

10.1089/ast.2015.1450.

Voß, B. *et al.* (2013) 'Insights into the Physiology and Ecology of the Brackish-Water-Adapted Cyanobacterium Nodularia spumigena CCY9414 Based on a Genome-Transcriptome Analysis', *PLoS ONE*. Edited by P. J. Janssen, 8(3), p. e60224. doi: 10.1371/journal.pone.0060224.

Wacey, D. *et al.* (2013) 'Nanoscale analysis of pyritized microfossils reveals differential heterotrophic consumption in the ~1.9-Ga Gunflint chert', *Proceedings of the National Academy of Sciences of the United States of America*, 110(20), pp. 8020–8024. doi: 10.1073/PNAS.1221965110/SUPPL_FILE/PNAS.201221965SI.PDF.

Wada, H. *et al.* (1992) 'Genetic Manipulation of the Extent of Desaturation of Fatty Acids in Membrane Lipids in the Cyanobacterium Synechocystis PCC6803', *Plant and Cell Physiology*, 33(5), pp. 535–540. doi: 10.1093/OXFORDJOURNALS.PCP.A078287.

Wada, H., Gombos, Z. and Murata, N. (1994) 'Contribution of membrane lipids to the ability of the photosynthetic machinery to tolerate temperature stress.', *Proceedings of the National Academy of Sciences*, 91(10), pp. 4273–4277. doi: 10.1073/PNAS.91.10.4273.

Watanabe, M. *et al.* (2014) 'Attachment of phycobilisomes in an antenna-photosystem I supercomplex of cyanobacteria', *Proceedings of the National Academy of Sciences of the United States of America*, 111(7), pp. 2512–2517. doi: 10.1073/pnas.1320599111.

Waterhouse, A. M. *et al.* (2009) 'Jalview Version 2--a multiple sequence alignment editor and analysis workbench', *Bioinformatics*, 25(9), pp. 1189–1191. doi: 10.1093/bioinformatics/btp033.

Wierzchos, J., Ascaso, C. and McKay, C. P. (2006) 'Endolithic cyanobacteria in halite rocks from the hyperarid core of the Atacama Desert', *Astrobiology*, 6(3), pp. 415–422. doi: 10.1089/ast.2006.6.415.

Williams, D. R. (2020) *Mars Fact Sheet*, *NASA*. Available at: https://nssdc.gsfc.nasa.gov/planetary/factsheet/marsfact.html (Accessed: 26 July 2021).

Wolf, E. T. (2017) 'Assessing the Habitability of the TRAPPIST-1 System Using a 3D Climate Model', *The Astrophysical Journal Letters*, 839(1), p. L1. doi: 10.3847/2041-8213/AA693A.

Xu, Q. *et al.* (1994) 'Mutational analysis of photosystem I polypeptides in Synechocystis sp. PCC 6803. Subunit requirements for reduction of NADP+ mediated by ferredoxin and flavodoxin.', *Journal of Biological Chemistry*, 269(34), pp. 21512–21518. doi: 10.1016/S0021-9258(17)31834-3.

Xu, Q. *et al.* (1995) 'Mutational Analysis of Photosystem I Polypeptides in the Cyanobacterium Synechocystis sp. PCC 6803', *Journal of Biological Chemistry*, 270(27), pp. 16243–16250. doi: 10.1074/JBC.270.27.16243.

Yagishita, T. *et al.* (1996) 'Photosynthetic bio-fuel cells using cyanobacteria', *Renewable Energy*, 9(1–4), pp. 958–961. doi: 10.1016/0960-1481(96)88439-4.

Zakar, T. *et al.* (2016) ‘Carotenoids assist in cyanobacterial photosystem II assembly and function’, *Frontiers in Plant Science*. Frontiers Media S.A., p. 295. doi: 10.3389/fpls.2016.00295.

Zhang, H. *et al.* (2021) ‘Snowball Earth, population bottleneck and Prochlorococcus evolution’, *Proceedings of the Royal Society B*, 288(1963). doi: 10.1098/RSPB.2021.1956.

Zhang, M. *et al.* (2017) ‘Freeze-drying of mammalian cells using trehalose: preservation of DNA integrity’, *Scientific Reports 2017 7:1*, 7(1), pp. 1–10. doi: 10.1038/s41598-017-06542-z.

Zhao, C. *et al.* (2015) ‘RfpA, RfpB, and RfpC are the Master Control Elements of Far-Red Light Photoacclimation (FaRLiP)’, *Frontiers in Microbiology*, 6(NOV), p. 1303. doi: 10.3389/fmicb.2015.01303.

Zickendraht-Wendelstadt, B., Friedrich, J. and Rüdiger, W. (1980) ‘Spectral Characterization of Monomeric C-Phycoerythrin From Pseudanabaena W 1173 and Its A and B Subunits: Energy Transfer in Isolated Subunits and C-Phycoerythrin’, *Photochemistry and Photobiology*, 31(4), pp. 367–376. doi: 10.1111/j.1751-1097.1980.tb02555.x.

Zinser, E. R. *et al.* (2007) ‘Influence of light and temperature on Prochlorococcus ecotype distributions in the Atlantic Ocean’, *Limnology and Oceanography*, 52(5), pp. 2205–2220. doi: 10.4319/LO.2007.52.5.2205.

Zolotov, M. Y. and Shock, E. L. (2001) ‘Composition and stability of salts on the surface of Europa and their oceanic origin’, *Journal of Geophysical Research: Planets*, 106(E12), pp. 32815–32827. doi: 10.1029/2000JE001413.

Supplementary Information

Refer to associated excel document attached as PDF below ‘[Supplementary Information] Genomic analysis of astrobiology-relevant adaptations to low light in far-red light utilising cyanobacteria’.

Supplementary Table S1: Reference ApcE2 protein sequences

Table shows reference ApcE2 protein sequences utilised for initial BLAST analysis to identify FARLIP-capable cyanobacteria

Strain	ApcE2 IMG Gene ID	ApcE2 GenBank ID
Calothrix parasitica NIES-267	2776204670	
Calothrix sp. NIES-3974	2776195382	
Calothrix sp. PCC 7507	2505798509	
Chlorogloeopsis fritschii PCC 6912	2512510463	
Chlorogloeopsis fritschii PCC 9212	2550828642	
Chroococcidiopsis sp. CCALA 051		WP_106544528.1
Chroococcidiopsis thermalis PCC 7203	2503611289	
Cyanobacterium TDX16		OWY64183.1
Cyanosarcina cf. burmensis CCALA 770	2789936001	
Fischerella major NIES-592		WP_073556984.1
Fischerella sp. NIES-4106	2776279192	
Fischerella sp. PCC 9605	2516147122	
*Fischerella thermalis BR2B**	2805783684 / 2805785123	
Fischerella thermalis JSC-11	2505770972	
Halomicronema hongdechloris C2206	2758436805	
Hydrococcus rivularis NIES-593		WP_073600233.1
Leptolyngbya sp. JSC-1	2022833636	
Pleurocapsa minor PCC 7327	2509573778	
Pleurocapsa sp. CCALA 161	2790194197	
Synechococcus sp. PCC 7335	647578828	

Supplementary Table S2: Reference FARIP protein sequences
Table shows reference FARLIP protein sequences used to confirm identity of newly discovered FARLIP cluster genes through maximum liklihood phylogeny

Protein	Strain	Source
Apc		
ApcA1	*Synechococcus PCC 7335*	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Chroococcidiopsis thermalis PCC 7203	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Leptolyngbya sp. JSC-1	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Calothrix sp. PCC 7507	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Chlorogloeopsis sp. PCC 9212	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Fischerella thermalis PCC 7521	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
ApcA2	*Synechococcus PCC 7335*	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Chroococcidiopsis thermalis PCC 7203	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Leptolyngbya sp. JSC-1	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Calothrix sp. PCC 7507	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Chlorogloeopsis sp. PCC 9212	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Fischerella thermalis PCC 7521	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
ApcA3	*Synechococcus sp. PCC 7002*	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Chroococcidiopsis thermalis PCC 7203	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Calothrix sp. PCC 7507	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Fischerella thermalis PCC 7521	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
ApcB1	*Synechococcus PCC 7335*	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Chroococcidiopsis thermalis PCC 7203	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Leptolyngbya sp. JSC-1	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Calothrix sp. PCC 7507	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Chlorogloeopsis sp. PCC 9212	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Fischerella thermalis PCC 7521	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Synechococcus sp. PCC 7002	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
ApcB2	*Synechococcus PCC 7335*	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Chroococcidiopsis thermalis PCC 7203	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Leptolyngbya sp. JSC-1	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Calothrix sp. PCC 7507	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Chlorogloeopsis sp. PCC 9212	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Fischerella thermalis PCC 7521	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
ApcB3	*Synechococcus PCC 7335*	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Chroococcidiopsis thermalis PCC 7203	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Chlorogloeopsis sp. PCC 9212	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Xenococcus sp. PCC 7305	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Leptolyngbya sp. PCC 6406	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Gloeocapsa sp. PCC 7428	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
ApcD1	*Synechococcus PCC 7335*	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Chroococcidiopsis thermalis PCC 7203	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Leptolyngbya sp. JSC-1	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Calothrix sp. PCC 7507	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Chlorogloeopsis sp. PCC 9212	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Fischerella thermalis PCC 7521	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
ApcD2	*Synechococcus PCC 7335*	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Chroococcidiopsis thermalis PCC 7203	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Leptolyngbya sp. JSC-1	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Calothrix sp. PCC 7507	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Chlorogloeopsis sp. PCC 9212	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Fischerella thermalis PCC 7521	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
ApcD3	*Synechococcus PCC 7335*	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Chroococcidiopsis thermalis PCC 7203	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Leptolyngbya sp. JSC-1	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Chlorogloeopsis sp. PCC 9212	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Fischerella thermalis PCC 7521	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
ApcD4	*Synechococcus PCC 7335*	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Chroococcidiopsis thermalis PCC 7203	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Chlorogloeopsis sp. PCC 9212	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Xenococcus sp. PCC 7305	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Leptolyngbya sp. PCC 6406	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Gloeocapsa sp. PCC 7428	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
ApcF1	*Synechococcus PCC 7335*	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Chroococcidiopsis thermalis PCC 7203	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Leptolyngbya sp. JSC-1	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Calothrix sp. PCC 7507	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Chlorogloeopsis sp. PCC 9212	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Fischerella thermalis PCC 7521	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Synechococcus sp. PCC 7002	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
ApcF2	*Oscillatoriales sp. JSC-12*	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Chroococcidiopsis thermalis PCC 7203	Gan, F., Shen, G. and Bryant, D. (2014) 'Occurrence of Far-Red Light Photoacclimation (FaRLIP) in Diverse Cyanobacteria', Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
Psa		
PsaA1	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) 'Far-red light photoacclimation (FaRLIP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression', Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Fischerella thermalis PCC 7521	Gisriel, C. et al. (2020) 'The structure of Photosystem I acclimated to far-red light illuminates an ecologically important acclimation process in photosynthesis', Science Advances, 6(6), p. eaay6415. doi: 10.1126/sciadv.aay6415.
	Chlorogloeopsis sp. PCC 9212	Ho, M. Y. et al. (2016) 'Light-dependent chlorophyll f synthase is a highly divergent paralog of PsbA of photosystem II', Science, 353(6302). doi: 10.1126/science.aaf9178.
	Leptolyngbya sp. JSC-1	Gan, F. et al. (2014) 'Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light', Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsaA2	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) 'Far-red light photoacclimation (FaRLIP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression', Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) 'Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions', BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Fischerella thermalis PCC 7521	Gisriel, C. et al. (2020) 'The structure of Photosystem I acclimated to far-red light illuminates an ecologically important acclimation process in photosynthesis', Science Advances, 6(6), p. eaay6415. doi: 10.1126/sciadv.aay6415.
	Chlorogloeopsis sp. PCC 9212	Ho, M. Y. et al. (2016) 'Light-dependent chlorophyll f synthase is a highly divergent paralog of PsbA of photosystem II', Science, 353(6302). doi: 10.1126/science.aaf9178.
	Leptolyngbya sp. JSC-1	Gan, F. et al. (2014) 'Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light', Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsaB1	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) 'Far-red light photoacclimation (FaRLIP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression', Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Fischerella thermalis PCC 7521	Gisriel, C. et al. (2020) 'The structure of Photosystem I acclimated to far-red light illuminates an ecologically important acclimation process in photosynthesis', Science Advances, 6(6), p. eaay6415. doi: 10.1126/sciadv.aay6415.
	Chlorogloeopsis sp. PCC 9212	Ho, M. Y. et al. (2016) 'Light-dependent chlorophyll f synthase is a highly divergent paralog of PsbA of photosystem II', Science, 353(6302). doi: 10.1126/science.aaf9178.
	Leptolyngbya sp. JSC-1	Gan, F. et al. (2014) 'Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light', Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsaB2	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) 'Far-red light photoacclimation (FaRLIP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression', Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) 'Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions', BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Fischerella thermalis PCC 7521	Gisriel, C. et al. (2020) 'The structure of Photosystem I acclimated to far-red light illuminates an ecologically important acclimation process in photosynthesis', Science Advances, 6(6), p. eaay6415. doi: 10.1126/sciadv.aay6415.
	Chlorogloeopsis sp. PCC 9212	Ho, M. Y. et al. (2016) 'Light-dependent chlorophyll f synthase is a highly divergent paralog of PsbA of photosystem II', Science, 353(6302). doi: 10.1126/science.aaf9178.
	Leptolyngbya sp. JSC-1	Gan, F. et al. (2014) 'Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light', Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsaB3	*Chlorogloeopsis sp. PCC 9212*	Ho, M. Y. et al. (2016) 'Light-dependent chlorophyll f synthase is a highly divergent paralog of PsbA of photosystem II', Science, 353(6302). doi: 10.1126/science.aaf9178.
	Leptolyngbya sp. JSC-1	Gan, F. et al. (2014) 'Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light', Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsaC	*Halomicronema hongdechloris C2206*	Chen, M. et al. (2019) 'Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions', BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Fischerella thermalis PCC 7521	Gisriel, C. et al. (2020) 'The structure of Photosystem I acclimated to far-red light illuminates an ecologically important acclimation process in photosynthesis', Science Advances, 6(6), p. eaay6415. doi: 10.1126/sciadv.aay6415.
	Chlorogloeopsis sp. PCC 9212	Ho, M. Y. et al. (2016) 'Light-dependent chlorophyll f synthase is a highly divergent paralog of PsbA of photosystem II', Science, 353(6302). doi: 10.1126/science.aaf9178.
	Leptolyngbya sp. JSC-1	Gan, F. et al. (2014) 'Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light', Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsaD	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) 'Far-red light photoacclimation (FaRLIP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression', Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) 'Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions', BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Fischerella thermalis PCC 7521	Gisriel, C. et al. (2020) 'The structure of Photosystem I acclimated to far-red light illuminates an ecologically important acclimation process in photosynthesis', Science Advances, 6(6), p. eaay6415. doi: 10.1126/sciadv.aay6415.
	Chlorogloeopsis sp. PCC 9212	Ho, M. Y. et al. (2016) 'Light-dependent chlorophyll f synthase is a highly divergent paralog of PsbA of photosystem II', Science, 353(6302). doi: 10.1126/science.aaf9178.
	Leptolyngbya sp. JSC-1	Gan, F. et al. (2014) 'Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light', Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsaE	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) 'Far-red light photoacclimation (FaRLIP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression', Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) 'Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions', BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Fischerella thermalis PCC 7521	Gisriel, C. et al. (2020) 'The structure of Photosystem I acclimated to far-red light illuminates an ecologically important acclimation process in photosynthesis', Science Advances, 6(6), p. eaay6415. doi: 10.1126/sciadv.aay6415.
	Chlorogloeopsis sp. PCC 9212	Ho, M. Y. et al. (2016) 'Light-dependent chlorophyll f synthase is a highly divergent paralog of PsbA of photosystem II', Science, 353(6302). doi: 10.1126/science.aaf9178.
	Leptolyngbya sp. JSC-1	Gan, F. et al. (2014) 'Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light', Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsaF1	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) 'Far-red light photoacclimation (FaRLIP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression', Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) 'Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions', BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Fischerella thermalis PCC 7521	Gisriel, C. et al. (2020) 'The structure of Photosystem I acclimated to far-red light illuminates an ecologically important acclimation process in photosynthesis', Science Advances, 6(6), p. eaay6415. doi: 10.1126/sciadv.aay6415.
	Chlorogloeopsis sp. PCC 9212	Ho, M. Y. et al. (2016) 'Light-dependent chlorophyll f synthase is a highly divergent paralog of PsbA of photosystem II', Science, 353(6302). doi: 10.1126/science.aaf9178.
	Leptolyngbya sp. JSC-1	Gan, F. et al. (2014) 'Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light', Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsaF2	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) 'Far-red light photoacclimation (FaRLIP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression', Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) 'Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions', BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Fischerella thermalis PCC 7521	Gisriel, C. et al. (2020) 'The structure of Photosystem I acclimated to far-red light illuminates an ecologically important acclimation process in photosynthesis', Science Advances, 6(6), p. eaay6415. doi: 10.1126/sciadv.aay6415.
	Chlorogloeopsis sp. PCC 9212	Ho, M. Y. et al. (2016) 'Light-dependent chlorophyll f synthase is a highly divergent paralog of PsbA of photosystem II', Science, 353(6302). doi: 10.1126/science.aaf9178.
	Leptolyngbya sp. JSC-1	Gan, F. et al. (2014) 'Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light', Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsaI1	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) 'Far-red light photoacclimation (FaRLIP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression', Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Fischerella thermalis PCC 7521	Gisriel, C. et al. (2020) 'The structure of Photosystem I acclimated to far-red light illuminates an ecologically important acclimation process in photosynthesis', Science Advances, 6(6), p. eaay6415. doi: 10.1126/sciadv.aay6415.
	Leptolyngbya sp. JSC-1	Gan, F. et al. (2014) 'Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light', Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsaI2	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) 'Far-red light photoacclimation (FaRLIP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression', Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) 'Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions', BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Fischerella thermalis PCC 7521	Gisriel, C. et al. (2020) 'The structure of Photosystem I acclimated to far-red light illuminates an ecologically important acclimation process in photosynthesis', Science Advances, 6(6), p. eaay6415. doi: 10.1126/sciadv.aay6415.
	Leptolyngbya sp. JSC-1	Gan, F. et al. (2014) 'Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light', Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsaJ1	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) 'Far-red light photoacclimation (FaRLIP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression', Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) 'Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions', BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Fischerella thermalis PCC 7521	Gisriel, C. et al. (2020) 'The structure of Photosystem I acclimated to far-red light illuminates an ecologically important acclimation process in photosynthesis', Science Advances, 6(6), p. eaay6415. doi: 10.1126/sciadv.aay6415.
	Chlorogloeopsis sp. PCC 9212	Ho, M. Y. et al. (2016) 'Light-dependent chlorophyll f synthase is a highly divergent paralog of PsbA of photosystem II', Science, 353(6302). doi: 10.1126/science.aaf9178.
	Leptolyngbya sp. JSC-1	Gan, F. et al. (2014) 'Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light', Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsaJ2	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) 'Far-red light photoacclimation (FaRLIP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression', Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) 'Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions', BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Leptolyngbya sp. JSC-1	Gan, F. et al. (2014) 'Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light', Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsaK	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) 'Far-red light photoacclimation (FaRLIP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression', Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) 'Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions', BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.

	Leptolyngbya sp. JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsaL1	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Fischerella thermalis PCC 7521	Gisriel, C. et al. (2020) ‘The structure of Photosystem I acclimated to far-red light illuminates an ecologically important acclimation process in photosynthesis’, Science Advances, 6(6), p. eaay6415. doi: 10.1126/sciadv.aay6415.
	Chlorogloeopsis sp. PCC 9212	Ho, M. Y. et al. (2016) ‘Light-dependent chlorophyll f synthase is a highly divergent paralog of PsbA of photosystem II’, Science, 353(6302). doi: 10.1126/science.aaf9178.
	Leptolyngbya sp. JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsaL2	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) ‘Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions’, BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Fischerella thermalis PCC 7521	Gisriel, C. et al. (2020) ‘The structure of Photosystem I acclimated to far-red light illuminates an ecologically important acclimation process in photosynthesis’, Science Advances, 6(6), p. eaay6415. doi: 10.1126/sciadv.aay6415.
	Chlorogloeopsis sp. PCC 9212	Ho, M. Y. et al. (2016) ‘Light-dependent chlorophyll f synthase is a highly divergent paralog of PsbA of photosystem II’, Science, 353(6302). doi: 10.1126/science.aaf9178.
	Leptolyngbya sp. JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsaL3	*Leptolyngbya sp. JSC-1*	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsaX	*Fischerella thermalis PCC 7521*	Gisriel, C. et al. (2020) ‘The structure of Photosystem I acclimated to far-red light illuminates an ecologically important acclimation process in photosynthesis’, Science Advances, 6(6), p. eaay6415. doi: 10.1126/sciadv.aay6415.
	Chlorogloeopsis sp. PCC 9212	Ho, M. Y. et al. (2016) ‘Light-dependent chlorophyll f synthase is a highly divergent paralog of PsbA of photosystem II’, Science, 353(6302). doi: 10.1126/science.aaf9178.
	Leptolyngbya sp. JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsaM	*Chlorogloeopsis sp. PCC 9212*	Ho, M. Y. et al. (2016) ‘Light-dependent chlorophyll f synthase is a highly divergent paralog of PsbA of photosystem II’, Science, 353(6302). doi: 10.1126/science.aaf9178.
	Fischerella thermalis PCC 7521	Gisriel, C. et al. (2020) ‘The structure of Photosystem I acclimated to far-red light illuminates an ecologically important acclimation process in photosynthesis’, Science Advances, 6(6), p. eaay6415. doi: 10.1126/sciadv.aay6415.
IsiA1	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Leptolyngbya sp. JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) ‘Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions’, BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
IsiX	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Leptolyngbya sp. JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.

Psb

PsbA	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) ‘Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions’, BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
PsbA1	*Leptolyngbya JSC-1*	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsbA2	*Leptolyngbya JSC-1*	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsbA3	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Leptolyngbya JSC-1	
	Chlorogloeopsis sp. PCC 9212	Gan, F., Shen, G. and Bryant, D. (2014) ‘Occurrence of Far-Red Light Photoacclimation (FaRLiP) in Diverse Cyanobacteria’, Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Chroococcidiopsis thermalis PCC 7203	Gan, F., Shen, G. and Bryant, D. (2014) ‘Occurrence of Far-Red Light Photoacclimation (FaRLiP) in Diverse Cyanobacteria’, Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Calothrix sp. PCC 7507	Gan, F., Shen, G. and Bryant, D. (2014) ‘Occurrence of Far-Red Light Photoacclimation (FaRLiP) in Diverse Cyanobacteria’, Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
	Fischerella thermalis PCC 7521	Gan, F., Shen, G. and Bryant, D. (2014) ‘Occurrence of Far-Red Light Photoacclimation (FaRLiP) in Diverse Cyanobacteria’, Life, 5(1), pp. 4–24. doi: 10.3390/life5010004.
PsbA4/Chl F Synthase	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) ‘Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions’, BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Leptolyngbya JSC-1	
	Chlorogloeopsis sp. PCC 9212	Ho, M. Y. et al. (2016) ‘Light-dependent chlorophyll f synthase is a highly divergent paralog of PsbA of photosystem II’, Science, 353(6302). doi: 10.1126/science.aaf9178.
	Chroococcidiopsis thermalis PCC 7203	Ho, M. Y. et al. (2016) ‘Light-dependent chlorophyll f synthase is a highly divergent paralog of PsbA of photosystem II’, Science, 353(6302). doi: 10.1126/science.aaf9178.
	Calothrix sp. PCC 7507	Ho, M. Y. et al. (2016) ‘Light-dependent chlorophyll f synthase is a highly divergent paralog of PsbA of photosystem II’, Science, 353(6302). doi: 10.1126/science.aaf9178.
PsbB1 / CP47	*Halomicronema hongdechloris C2206*	Chen, M. et al. (2019) ‘Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions’, BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Leptolyngbya JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsbB2 / CP47	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) ‘Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions’, BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Leptolyngbya JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsbC1 / CP43	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) ‘Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions’, BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Leptolyngbya JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsbC2 / CP43	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) ‘Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions’, BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Leptolyngbya JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsbD1	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
PsbD2	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
PsbD3	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Leptolyngbya JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsbE	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) ‘Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions’, BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Leptolyngbya JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsbF / b559	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) ‘Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions’, BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
PsbH1	*Halomicronema hongdechloris C2206*	Chen, M. et al. (2019) ‘Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions’, BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Leptolyngbya JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsbH2	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
PsbI	*Halomicronema hongdechloris C2206*	Chen, M. et al. (2019) ‘Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions’, BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Leptolyngbya JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
	Synechococcus sp. PCC 7335	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
PsbJ	*Halomicronema hongdechloris C2206*	Chen, M. et al. (2019) ‘Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions’, BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Synechococcus sp. PCC 7335	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
PsbK	*Halomicronema hongdechloris C2206*	Chen, M. et al. (2019) ‘Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions’, BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Leptolyngbya JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
	Synechococcus sp. PCC 7335	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
PsbL	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
PsbM	*Halomicronema hongdechloris C2206*	Chen, M. et al. (2019) ‘Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions’, BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Leptolyngbya JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsbN	*Halomicronema hongdechloris C2206*	Chen, M. et al. (2019) ‘Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions’, BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Leptolyngbya JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
	Synechococcus sp. PCC 7335	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
PsbO1 / OEE1	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) ‘Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions’, BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Leptolyngbya JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsbO2 / OEE1	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) ‘Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions’, BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
PsbP / OEE2	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) ‘Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions’, BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Leptolyngbya JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsbQ / OEE3	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) ‘Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions’, BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Leptolyngbya JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsbU	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Leptolyngbya JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsbV1	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) ‘Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions’, BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Leptolyngbya JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsbV2	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) ‘Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions’, BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Leptolyngbya JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsbW	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Leptolyngbya JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsbX	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) ‘Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions’, BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
PsbY	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) ‘Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions’, BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Leptolyngbya JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
PsbZ	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) ‘Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions’, BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Leptolyngbya JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
Psb27	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Halomicronema hongdechloris C2206	Chen, M. et al. (2019) ‘Genome and proteome of the chlorophyll f-producing cyanobacterium Halomicronema hongdechloris: Adaptative proteomic shifts under different light conditions’, BMC Genomics, 20(1), p. 207. doi: 10.1186/s12864-019-5587-3.
	Leptolyngbya JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.

Rfp

RfpA	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Leptolyngbya JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
RfpB	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Leptolyngbya JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.
RfpC	*Synechococcus sp. PCC 7335*	Ho, M. Y. et al. (2017) ‘Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression’, Photosynthesis Research, 131(2), pp. 173–186. doi: 10.1007/s11120-016-0309-z.
	Leptolyngbya JSC-1	Gan, F. et al. (2014) ‘Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light’, Science, 345(6202), pp. 1312–1317. doi: 10.1126/science.1256963.

Supplementary Table S3: Reference astrobiology-relevant genes
Table shows reference proteins used in BLAST analysis to identify the presence or absence of such proteins in identified LLAC

Carbon fixation

Protein	Strain	NCBI Protein ID
Icfa / CcaA	*Acaryochloris sp. RCC1774*	PZD72082.1
	Halomicronema hongdechloris C2206	ASC74335.1
	Leptolyngbya sp. O-77	BAU43522.1
	Microcystis aeruginosa NIES-1211	GBL14433.1
	Microcystis aeruginosa NIES-4264	GCA87517.1
	Synechococcus elongatus PCC 6301	BAD78300.1
	Synechocystis sp. PCC 6803	AGF52563.1

Nitrogen fixation

Protein	Strain	NCBI Protein ID
NifD	*Chroococcidiopsis sp. PCC 6712*	ABU63069.1
	Chroococcidiopsis thermalis PCC 7203	ABU63073.1
	Leptolyngbya sp. PCC 7375	ABU63080.1
	Nodosilinea nodulosa PCC 7104	ABU63072.1
	Pleurocapsa sp. PCC 7327	ABU63079.1
	Pleurocapsa sp. PCC 7516	ABU63084.1
	Pseudanabaena tenuis PCC 7409	ABU63081.1
	Symploca atlantica PCC 8002	ABU63085.1
	Synechococcus sp. PCC 7335	ABU63076.1
	Xenococcus sp. PCC 7305	ABU63075.1
NifH	*Calothrix desertica PCC 7102*	RUS96653.1
	Chlorogloeopsis fritschii PCC 6912	RUR86127.1
	Chroococcidiopsis cubana SAG 39.79	RUT13332.1
	Leptolyngbya sp. Heron Island J	ESA34603.1
	Mastigocoleus testarum BC008	KST63417.1
	Nostoc sp. PCC 6720	CAA83510.1
	Nostoc sp. PCC 7120	CAA24729.1
	Planktothrix sp. FACHB-1365	WP_199314413.1
	Pleurocapsa minor	WP_015145031.1
	Trichodesmium erythraeum IMS101	AAF82637.1
NifK	*Anabaena sp. 90*	AFW96195.1
	Calothrix desertica PCC 7102	ACA61792.1
	Calothrix sp. PCC 7507	ACA61791.1
	Chlorogloeopsis fritschii PCC 6912	ACA61793.1
	Fischerella sp. PCC 7603	ACA61797.1
	Nostoc flagelliforme NX-09	ANQ45517.1
	Nodularia sphaerocarpa PCC 7804	ACA61799.1
	Nodosilinea sp. P-1105	WP_169614083.1
	Synechococcus sp.	AAA64845.1
	Trichodesmium erythraeum IMS101	AAF82639.1

Sucrose Synthesis

Protein	Strain	NCBI Protein ID
Spp	*Anabaena sp. 90*	WP_015080793.1
	Calothrix sp. NIES-3974	WP_096624638.1
	Calothrix sp. PCC 7507	WP_015128786.1
	Chroococcidiopsis sp. TS-821	PPS41211.1
	Fischerella sp. PCC 9431	WP_026719263.1
	Leptolyngbya sp.	PZV11806.1
	Mastigocoleus testarum BC008	KST67811.1
	Oscillatoriales cyanobacterium	TAH17444.1
	Phormidesmis priestleyi	PZO49346.1
	Synechocystis sp. PCC 6714	WP_028948339.1
Sps	*Anabaena sp. 90*	AFW95585.1
	Calothrix desertica PCC 7102	TWH51421.1
	Calothrix sp. NIES-4101	BAZ39260.1
	Calothrix sp. NIES-4071	BAZ09128.1
	Candidatus Synechococcus spongiarum	CZB11805.1
	Geminocystis sp. NIES-3709	BAQ65062.1
	Prochlorococcus marinus str. MIT 9313	CAE22442.1
	Thermosynechococcus elongatus BP-1	BAC08134.1
	Synechococcus elongatus PCC 6301	BAD78920.1
	Synechocystis sp. PCC 6803	BAK51567.1

Trehalose synthesis

Protein	Strain	NCBI Protein ID
TreY	*Anabaena sp. YBS01*	QFZ15363.1
	Chroococcidiopsis sp. CCALA 051	WP_106547162.1
	Chroococcidiopsis sp. PCC 6712	WP_169242252.1
	Cyanobacteria bacterium J055	RMG15117.1
	Nostoc flagelliforme CCNUN1	AUB39850.1
	Nostoc sphaeroides CCNUC1	QFS49479.1
	Oscillatoriales cyanobacterium	TAH15870.1
	Pleurocapsa sp. CCALA 161	WP_106240389.1
	Tolypothrix sp. PCC 7910	QIR35316.1
	Chroococcidiopsis sp. PCC 6712	WP_169242253.1
TreZ	*Fischerella thermalis*	WP_102150420.1
	Leptolyngbya sp.	PZV19733.1
	Nostoc flagelliforme	WP_100903099.1
	Oscillatoriales cyanobacterium	TAH21126.1
	Nostoc sp. TCL26-01	QLE56348.1
	Phormidesmis priestleyi	PZO60856.1
	Synechococcus sp. JA-3-3Ab	ABC98411.1
	Cyanobium sp.	GDX72945.1
TreS	*Cyanosarcina cf. burmensis CCALA 770*	PSB42547.1
	Chroococcidiopsis cubana SAG 39.79	RUT02304.1
	Chlorogloeopsis fritschii PCC 6912	RUR84209.1
	Fischerella muscicola CCMEE 5323	PLZ93418.1
	Leptolyngbya sp.	PZV05162.1
	Phormidesmis priestleyi	PZO51166.1
	Synechococcus sp. MIT S9509	KZR88300.1
	Synechococcus sp. WH 8020	AKN59779.1

Desaturase

Protein	Strain	NCBI Protein ID
DesA	*Calothrix desertica PCC 7102*	RUT01077.1
	Chroococcidiopsis cubana SAG 39.79	RUT12819.1
	Chlorogloeopsis fritschii PCC 6912	RUR87044.1
	Gloeobacter violaceus PCC 7421	BAC90564.1
	Nostoc sp. PCC 7120	RUR75019.1
	Planktothrix agardhii NIVA-CYA 126/8	KEI69195.1
	Prochlorococcus sp.	GIR04531.1
	Synechococcus sp. NIES-970	BAW97752.1
DesB	*Nostoc sp. 36*	CAF18425.1
	Planktothrix agardhii NIVA-CYA 126/8	KEI67496.1
	Synechococcus sp. NIES-970	BAW95305.1
	Synechocystis sp. PCC 6803	BAK50475.1
DesC	*Calothrix sp. NIES-4071*	BAZ17072.1
	Chlorogloeopsis fritschii PCC 6912	RUR75289.1
	Nostoc sp. PCC 7120	BAB77965.1
	Planktothrix agardhii NIVA-CYA 126/8	KEI67645.1
	Phormidium sp. OSCR	KPQ38035.1
	Synechococcus sp. PROS-7-1	ATV95783.1
	Thermosynechococcus sp. NK55	AHB89548.1
DesD	*Arthrospira platensis 540*	AFU92434.1
	Calothrix sp. NIES-3974	BAZ05022.1
	Calothrix sp. NIES-4105	BAZ63650.1
	Microcystis aeruginosa PCC 7806	CAO89155.1
	Nodularia sp. NIES-3585	GAX35149.1
	Synechocystis sp. PCC 6803	BAK50679.1

Hydrogen Production

Protein	Strain	NCBI Protein ID
HoxF	*Arthrospira platensis FACHB-341*	ABC26907.1
	Calothrix sp. NIES-2100	BAY27684.1
	Microcystis aeruginosa NIES-3807	GCL59623.1
	Nostoc sp. NIES-3756	BAT55215.1
	Planktothrix agardhii NIVA-CYA 126/8	KEI67824.1

	Pseudanabaena sp. ABRG5-3	BBC26262.1
	Synechococcus elongatus PCC 6301	CAA73873.1
	Synechocystis sp. CACIAM 05	WP_162328144.1
HoxU	*Aphanothece halophytica AICU56*	AHI54355.1
	Chroococcidiopsis	WP_015153233.
	Cyanobacteria	WP_039726747.1
	Microcystis aeruginosa PCC 7806SL	ARI80999.1
	Nodosilinea sp. P-1105	WP_169611455.1
	Nostocaceae	WP_114084459.1
	Planktothrix agardhii NIVA-CYA 126/8	KEI67822.1
	Synechococcus elongatus PCC 6301	CAA66381.1
HoxY	*Anabaena sp. 90*	AFW94652.1
	Aphanothece halophytica AICU56	AHI54356.1
	Calothrix sp. NIES-4101	BAZ39311.1
	Microcystis aeruginosa PCC 7806SL	ARI80996.1
	Nostoc sp. PCC 7120	BAB72721.1
	Synechococcus elongatus PCC 6301	CAA66382.1
HoxH	*Acaryochloris marina MBIC11017*	ABW32682.1
	Aphanothece halophytica AICU56	AHI54357.1
	Calothrix sp. NIES-4101	BAZ39312.1
	Nodularia spumigena UHCC 0039	AVZ31482.1
	Nostoc sp. PCC 7120	BAB72723.1
	Phormidium sp. OSCR	KPQ34926.1
	Synechococcus elongatus PCC 6301	CAA66383.1
	Synechocystis sp. PCC 6714	

Supplementary Table S4: FARLIP- and LOLIP-capable cyanobacteria

Table shows outcome of BLAST analysis using ApcE2 (FALRIP) or ApcB3 (LOLIP) proteins to identify FARLIP cluster-containing and LOLIP-cluster containing cyanobacteria in NCBI and JGI/IMG databases. LOLIP, L; FARLIP, F; LOLIP & FARLIP, LF

No.	Strain	ApcE2 IMG Gene ID	ApcE2 GenBank ID	ApcB3 IMG Gene ID	Location	Habitat	Photosynthetic capability	GenBank Assembly ID
1	*Leptolyngbya sp. SIO1E4*		MBE7379958.1		Marine benthic turfs, Ituau, American Samoa	Marine	F	GCA_010672825.2
2	*Fischerella thermalis CCMEE 5268*	2805779980			Chena Hot Springs, Alaska, USA	Hot Spring	F	GCF_002870565.1
3	*Fischerella thermalis CCMEE 5273*	2805788679			Harrison Hot Springs, British Columbia, Canada	Hot Spring	F	GCF_002870585.1
4	*Leptolyngbya sp. JSC-1*	2022833636			Near LaDuke Hot Springs, Montana, USA	Freshwater	F	GCA_000733415.1
5	*Fischerella thermalis BR2B*	2805783684 / 2805785123			Boiling River, Yellowstone National Park, USA	Hot Spring	F	GCF_002870705.1
5	*Fischerella thermalis CCMEE 5208*	2805866163			Obsidian Pool, Yellowstone National Park, USA	Hot Spring	F	GCF_002870755.1
5	*Fischerella thermalis PCC 7521*	2550718703			Spring Water, Mammoth Sinkhole II, Yellowstone National Park, USA	Hot Spring	F	GCA_000317225.1
5	*Fischerella thermalis WC1110*	2805832130			White Creek, Lower Geyser Basin, Yellowstone National Park, USA	Hot Spring	F	GCF_002870525.1
5	*Fischerella thermalis WC114*	2805838004			White Creek, Lower Geyser Basin, Yellowstone National Park, USA	Hot Spring	F	GCF_002870545.1
5	*Fischerella thermalis WC119*	2805841281			White Creek, Lower Geyser Basin, Yellowstone National Park, USA	Hot Spring	F	GCF_002870265.1
5	*Fischerella thermalis WC157*		WP_102148206.1		White Creek, Lower Geyser Basin, Yellowstone National Park, USA	Hot Spring	F	GCF_002870305.1
5	*Fischerella thermalis WC213*	2806026097			White Creek, Lower Geyser Basin, Yellowstone National Park, USA	Hot Spring	F	GCF_002870255.1
5	*Fischerella thermalis WC245*	2806314614			White Creek, Lower Geyser Basin, Yellowstone National Park, USA	Hot Spring	F	GCF_002870505.1
5	*Fischerella thermalis WC246*	2806091916			White Creek, Lower Geyser Basin, Yellowstone National Park, USA	Hot Spring	F	GCF_002870185.1
5	*Fischerella thermalis WC249*	2805882061			White Creek, Lower Geyser Basin, Yellowstone National Park, USA	Hot Spring	F	GCA_002870465.1
5	*Fischerella thermalis WC341*	2805826471			White Creek, Lower Geyser Basin, Yellowstone National Park, USA	Hot Spring	F	GCF_002870345.1
5	*Fischerella thermalis WC344*		WP_009453699.1		White Creek, Lower Geyser Basin, Yellowstone National Park, USA	Hot Spring	F	GCF_002870445.1
5	*Fischerella thermalis WC439*		WP_009453699.1		White Creek, Lower Geyser Basin, Yellowstone National Park, USA	Hot Spring	F	GCF_002870475.1
5	*Fischerella thermalis WC442*		WP_009453699.1		White Creek, Lower Geyser Basin, Yellowstone National Park, USA	Hot Spring	F	GCF_002870415.1
5	*Fischerella thermalis WC527*	2805898446			White Creek, Lower Geyser Basin, Yellowstone National Park, USA	Hot Spring	F	GCF_002870405.1
5	*Fischerella thermalis WC538*	2805892292			White Creek, Lower Geyser Basin, Yellowstone National Park, USA	Hot Spring	F	GCF_002870325.1
5	*Fischerella thermalis WC542*	2805887111			White Creek, Lower Geyser Basin, Yellowstone National Park, USA	Hot Spring	F	GCA_002870385.1
5	*Fischerella thermalis WC558*	2805849819			White Creek, Lower Geyser Basin, Yellowstone National Park, USA	Hot Spring	F	GCA_002870365.1
5	*Fischerella thermalis WC559*	2805855640			White Creek, Lower Geyser Basin, Yellowstone National Park, USA	Hot Spring	F	GCA_002870195.1
6	*Pleurocapsa minor PCC 7327*	2509573778			Hunter's Hot Spring, Oregon	Hot Spring	F	GCF_000317025.1
7	*Leptolyngbya sp. PCC 6406*			2517690605	California, USA	Freshwater	L	GCA_000332095.2
8	*Xenococcus sp. PCC 7305*			2508648507	Marine aquarium, La Jolla, California, USA	Marine	L	GCA_000332055.1
9	*Synechococcus sp. PCC 7335*	647578828		647578853	Snail shell, intertidal zone, Puerto Penasco, Mexico	Marine	LF	GCF_000155595.1
10	*Candidatus Gloeomargarita lithophora D10*			2719355299	Lake Alchichica, Puebla State, Mexico	Freshwater	L	GCA_001870225.1
11	*Chroococcidiopsis cubana SAG 39.79*	2887014983		2887016677	Dry soil, Pinar del Rio, Cuba	Terrestrial	LF	GCF_003991895.1
12	*Chroococcidiopsis cubana CCALA 043*	2805386709 / 2805387205		2805381211	Mineral spring, Santa Fe, Cuba	Freshwater	LF	GCA_003003835.1
13	*Mastigocoleus testarum BC008*	2724942860			Pelagic marine shell fragment, Puerto Rico	Marine	F	GCF_001456025.1
14	*Fischerella thermalis CCMEE 5318*	2805809184			Hot Spring Water, El Salvador	Hot Spring	F	GCF_002870675.1
15	*Cyanosarcina cf. burmensis CCALA 770*	2789936001		2789933144	Backwater, Rio Coraico, La Paz, Bolivia	Freshwater	LF	GCA_003004015.1
16	*Rivularia sp. T60 A2020 040*		MBF2016180.1		El Tatio Hot Springs, Antofagasta, Chile	Hot Spring	F	GCA_015272215.1
17	*Fischerella thermalis CCMEE 5194*	2805973761			Puyehue Hot Springs, Puyehue, Chile	Hot Spring	F	GCA_002870795.1
18	*Chlorogloeopsis fritschii C42 A2020 084*		MBF2007104.1		Cahuelmo Hot Spring, Los Lagos, Chile	Hot Spring	F	GCA_015272425.1
18	*Elainella sp. C42 A2020 010*		MBF2050151.1		Cahuelmo Hot Spring, Los Lagos, Chile	Hot Spring	F	GCA_015272495.1
18	*Hydrococcus sp. C42 A2020 068*		MBF2022682.1		Cahuelmo Hot Spring, Los Lagos, Chile	Hot Spring	F	GCA_015272405.1
18	*Synechococcales cyanobacterium C42 A2020 086*		MBF2072647.1		Cahuelmo Hot Spring, Los Lagos, Chile	Hot Spring	F	GCA_015272325.1
19	*Fischerella thermalis M48 A2018 028*		MBF2070717.1		Porcelana Hot Spring, Los lago, Chile	Hot Spring	F	GCA_015272315.1
19	*Synechococcales cyanobacterium M58 A2018 015*		MBF2000039.1		Porcelana Hot Spring, Los lago, Chile	Hot Spring	F	GCA_015272295.1
20	*Fischerella muscicola CCMEE 5323*	2808399257			Geysir Springs, Iceland	Hot Spring	F	GCF_002870665.1
21	*Chlorogloeopsis fritschii PCC 9212*	2550828642		2550829381	Thermal Spring Water, Orense, Spain	Hot Spring	LF	GCA_000317265.1
22	*Romeria aff. gracilis LEGE 07310*	2917553957		2917551315	Benthic zone of Minho estuary, Caminha, Portugal	Brackish	LF	GCA_015207255.1
23	*Nodosilinea sp. LEGE 07088*	2914094390		2914094433	Douro estuary, Porto, Portugal	Brackish	LF	GCA_015207395.1
23	*Nodosilinea sp. LEGE 07298*	2914025498 / 2914026750		2914028765	Douro Estuary Shore, Porto, Portugal	Brackish	LF	GCA_015207265.1
24	*Chroococcidiopsidales cyanobacterium LEGE 13417*		MBE9016838.1		Porto metropolitan area, Portugal	Terrestrial	F	GCA_015206905.1
25	*Leptolyngbya cf. ectocarpi LEGE 11479*	2914341546		2914343713	Epilithic Subtidal Sample, Plo Negro, Portugal	Marine	LF	GCA_015207065.1
26	*Lusitaniella coriacea LEGE 07157*			2914041607	Tide pool, Praia de Lavadores, Canidelo, Portugal	Marine	L	GCA_015207425.1
27	*Pleurocapsales cyanobacterium LEGE 10410*		MBE9046795.1	2913838106	Intertidal zone pebble, Vila Nova de Mil Fontes, Portugal	Marine	LF	GCA_015207195.1
28	*Plectonema cf. radiosum LEGE 06105*	2914399705			Intertidal zone of Praia da Luz, Lagos, Portugal	Marine	F	GCA_015207665.1
28	*Gloeocapsopsis crepidinum LEGE 06123*			2914407313	Intertidal zone of Praia da Luz, Lagos, Portugal	Marine	L	GCA_015207655.1
29	*Chroococcidiopsis thermalis PCC 7203*	2503611289		2503614632	Soil, near Greifswald, Germany	Terrestrial	LF	GCF_000317125.1
30	*Chroococcales cyanobacterium IPPAS B-1203*			2882385357	Hot Spring Karlovy Vary, Karlovy Vary, Czech Republic	Hot Spring	L	GCA_002749975.1
31	*Fischerella thermalis CCMEE 5282*	2805796010			Sklene Teplice Spring, Banská Bystrica, Slovakia	Hot Spring	F	GCF_002870615.1
32	*Chroococcidiopsis sp. CCALA 051*		WP_106544528.1		Belianske Tatry, Slovakia	Freshwater	F	GCA_003015105.1
33	*Pleurocapsa sp. CCALA 161*	2790194197		2790191999	Freshwater Lake, Vysoke Tatry, Slovakia	Freshwater	LF	GCA_003003995.1
34	*Calothrix sp. PCC 7507*	2505798509			Sphagnum bog, Switzerland	Freshwater	F	GCF_000316575.1
35	*Fischerella thermalis CCMEE 5328*	2805968023			Spring near Cava Scura, Ischia, Italy	Hot Spring	F	GCA_002870845.1
36	*Fischerella thermalis CCMEE 5205*	2805912058			Hot spring, Oman	Hot Spring	F	GCF_002870745.1
37	*Nodosilinea sp. P-1105*	2887195717		2887196978	Cock Soda lake, Altai region, Russia	Hypersaline	LF	GCF_012911975.1
38	*Leptolyngbya sp. IPPAS B-1204*		RNJ67180.1		Lake Tolbo Nuur, Mongolia	Freshwater	F	GCA_003724315.1
39	*Chlorogloeopsis fritschii PCC 6912*	2512510463		2509834706 / 2551969798 / 2512514088	Soil, Allahabad, India	Terrestrial	LF	GCF_003990575.1
40	*Gloeocapsa sp. PCC 7428*			2503796422	Amparai District, Maha Oya, Sri Lanka	Hot Spring	L	GCA_000317555.1
41	*Scytonema millei VB511283*			2648590929	Stone, West Bengal, India	Terrestrial	L	GCA_000817735.3
42	*Fischerella sp. FACHB-380*	2909453173			China	Freshwater	F	GCA_014697535.1
42	*Leptolyngbya sp. FACHB-60*			2914454961	China	Terrestrial	L	GCA_014695775.1
42	*Phormidium sp. FACHB-77*			2920022192	China	Terrestrial	L	GCA_014695595.1
42	*Chroococcidiopsis sp. FACHB-1243*	2914552355		2914551400	China	Freshwater	LF	GCA_014696895.1

43	*Cyanobacterium TDX16*		OWY64183.1		Tianjin, China	Freshwater	F	GCA_002213405.1
44	*Leptolyngbya sp. KIOST-1*			2619437334	Pond, Ansan, South Korea	Freshwater	L	GCA_000763385.1
45	*Calothrix parasitica NIES-267*	2776204670			Oshoro Bay, Hokkaido, Japan	Marine	F	GCF_002368095.1
46	*Calothrix sp. NIES-3974*	2776195382			Hot Spring Sediment, Nakabusa Hot Spring, Japan	Hot Spring	F	GCF_002368395.1
47	*Fischerella sp. NIES-3754*	2687481121			Suwa Shrine Hot Spring, Nagano, Japan	Hot Spring	F	GCF_001548455.1
48	*Fischerella sp. NIES-4106*	2776279192			Hagiu Forest, Chiba, Japan	Terrestrial	F	GCF_002368315.1
49	*Fischerella thermalis CCMEE 5330*	2805873729			Hakone Hot Spring, Hakone, Japan	Hot Spring	F	GCA_002870725.1
50	*Fischerella major NIES-592*		WP_073556984.1		Hagiwara Hot Spring, Tomiji, Japan	Hot Spring	F	GCF_001904645.1
50	*Hydrococcus rivularis NIES-593*		WP_073600233.1		Hagiwara Hot Spring, Tomiji, Japan	Hot Spring	F	GCA_001904635.1
51	*Fischerella thermalis CCMEE 5201*	2805905503			Waitangi Springs, Rotoma, New Zealand	Hot Spring	F	GCF_002870785.1
52	*Fischerella thermalis CCMEE 5196*	2805962307			Hot Spring, Ohinemutu, New Zealand	Hot Spring	F	GCA_002870825.1
52	*Fischerella thermalis CCMEE 5198*	2805797952			Hot Spring, Ohinemutu, New Zealand	Hot Spring	F	GCF_002870635.1
53	*Fischerella muscicola PCC 7414*	2550708254			Hot Spring Water, New Zealand	Hot Spring	F	GCA_000317205.1
54	*Halomicronema hongdechloris C2206*	2758436805		2758437816	Stromatolite mat, Shark Bay, Austrialia	Marine	LF	GCF_002075285.3
55	*Cyanobacteria bacterium CRU 2 1*		NJR61196.1		Stromatolite mat, Cape Recife, South Africa	Marine	F	GCA_012034815.1
55	*Cyanobacteria bacterium RU 5 0*		NJO40009.1		Stromatolite mat, Cape Recife, South Africa	Marine	F	GCA_012033255.1
55	*Hydrococcus sp. CRU 1 1*		NJP18818.1		Stromatolite mat, Cape Recife, South Africa	Marine	F	GCA_012034135.1
55	*Hydrococcus sp. RU 2 2*		NJM87016.1 (partial)		Stromatolite mat, Cape Recife, South Africa	Marine	F	GCA_012032735.1
55	*Leptolyngbyaceae cyanobacterium RM2 2 4*		NJO48766.1		Stromatolite mat, Cape Recife, South Africa	Marine	F	GCA_012033305.1
55	*Leptolyngbyaceae cyanobacterium RU 5 1*		NJP09552.1		Stromatolite mat, Cape Recife, South Africa	Marine	F	GCA_012034055.1
55	*Phormidesmis sp. RL 2 1*		NJM97203.1		Stromatolite mat, Cape Recife, South Africa	Marine	F	GCA_012033015.1
55	*Pleurocapsa sp. CRU 1 2*		NJO96835.1		Stromatolite mat, Cape Recife, South Africa	Marine	F	GCA_012033675.1
55	*Richelia sp. RM1 1 1*		NJN07436.1		Stromatolite mat, Cape Recife, South Africa	Marine	F	GCA_012032385.1
56	*Calothrix_sp._CSU_2_0*		NJR18301.1		Stromatolite mat, Nelson Mandela Bay, Schoenmakerskop, South Africa	Marine	F	GCA_012034595.1
56	*Hydrococcus sp. CSU 1 8*		NJQ98908.1 (partial)		Stromatolite mat, Nelson Mandela Bay, Schoenmakerskop, South Africa	Marine	F	GCA_012034605.1
56	*Leptolyngbyaceae cyanobacterium SM1 4 3*		NJL39709.1		Stromatolite mat, Nelson Mandela Bay, Schoenmakerskop, South Africa	Marine	F	GCA_012031415.1
56	*Pleurocapsa sp. SU 5 0*		NJK57233.1		Stromatolite mat, Nelson Mandela Bay, Schoenmakerskop, South Africa	Marine	F	GCA_012030825.1
56	*Richelia sp. SL 2 1*		NJO26766.1		Stromatolite mat, Nelson Mandela Bay, Schoenmakerskop, South Africa	Marine	F	GCA_012033205.1
56	*Richelia sp. SM1 7 0*		NJM17574.1		Stromatolite mat, Nelson Mandela Bay, Schoenmakerskop, South Africa	Marine	F	GCA_012031715.1
	Fischerella thermalis JSC-11	2505770972			Missing	Missing	F	GCA_000231365.2
	Oscillatoriales cyanobacterium JSC-12	2510098318			Missing	Missing	F	GCA_000309945.1
	Fischerella sp. PCC 9605	2516147122		2516148359	Missing	Missing	LF	GCA_000517105.1

Supplementary Table S5: Previously unreported FARLIP cluster-containing cyanobacteria

Table shows strains of cyanobacteria previously unreported to be FARLIP cluster-containing, and sequence variation in the conserved 'VIPEDVT' motif

Strain	Location	Habitat	VIPEDVT
Chroococcidiopsis cubana CCALA 043	Mineral spring, Santa Fe, Cuba	Freshwater	VIPEDVT
Chroococcidiopsis sp. FACHB-1243	Freshwater Sample, China	Freshwater	VIPEDVT
Chroococcidiopsidales cyanobacterium LEGE 13417	Porto metropolitan area, Portugal	Terrestrial	VIPEDVT
Hydrococcus sp. CRU 1 1	Stromatolite mat, Cape Recife, South Africa	Marine	IIPEDVT
Hydrococcus sp. CSU 1 8	Stromatolite mat, Nelson Mandela Bay, Schoenmakerskop, South Africa	Marine	IIPEDVT
Hydrococcus sp. RU 2 2	Cape Recife, South Africa	Marine	IIPEDVT
Leptolyngbya sp. SIO1E4	Marine benthic turfs, Ituau, American Samoa	Marine	VIPEDVT
Leptolyngbya cf. ectocarpi LEGE 11479	Epilithic Subtidal Sample, Plo Negro, Portugal	Marine	VIPEDAT
Leptolyngbya sp. IPPAS B-1204	Lake Tolbo Nuur, Mongolia	Freshwater	VIPEDVT
Leptolyngbyaceae cyanobacterium RU 5 1	Stromatolite mat, Cape Recife, South Africa	Marine	VIPEDVT
Leptolyngbyaceae cyanobacterium RM2 2 4	Stromatolite mat, Cape Recife, South Africa	Marine	IIPEDVT
Leptolyngbyaceae cyanobacterium SM1 4 3	Stromatolite mat, Nelson Mandela Bay, Schoenmakerskop, South Africa	Marine	IIPEDVT
Nodosilinea sp. LEGE 07298	Douro Estuary Shore, Porto, Portugal	Brackish	VIPEDVT
Nodosilinea sp. P-1105	Cock Soda lake, Altai region, Russia	Hypersaline	VIPEDVT
Nodosilinea sp. LEGE 07088	Douro estuary, Porto, Portugal	Marine	VIPEDVT
Phormidesmis sp. RL 2 1	Stromatolite mat, Cape Recife, South Africa	Marine	VIPEDVT
Plectonema cf. radiosum LEGE 06105	Intertidal zone of Praia da Luz, Lagos, Portugal	Marine	VIPEDVT
Pleurocapsa sp. CRU 1 2	Stromatolite mat, Cape Recife, South Africa	Marine	VIPEDVT
Pleurocapsales cyanobacterium LEGE 10410	Intertidal zone pebble, Vila Nova de Mil Fontes, Portugal	Marine	VIPEDVT
Richelia sp. SL 2 1	Stromatolite mat, Nelson Mandela Bay, Schoenmakerskop, South Africa	Marine	VIPEDVT
Romeria aff. gracilis LEGE 07310	Benthic zone of Minho estuary, Caminha, Portugal	Brackish	VIPEDVT
Richelia sp. SM1 7 0	Stromatolite mat, Nelson Mandela Bay, Schoenmakerskop, South Africa	Marine	VIPEDVT
Richelia sp. RM1 1 1	Stromatolite mat, Cape Recife, South Africa	Marine	VIPEDVT
Calothrix sp. CSU 2 0	Stromatolite mat, Nelson Mandela Bay, Schoenmakerskop, South Africa	Marine	VIPEDVT
Chlorogloeopsis fritschii C42 A2020 084	Cahuelmo Hot Spring, Los Lagos, Chile	Hot Spring	VIPEDVT
Hydrococcus sp. C42 A2020 068	Cahuelmo Hot Spring, Los Lagos, Chile	Hot Spring	VIPEDVT
Pleurocapsa sp. SU 5 0	Stromatolite mat, Nelson Mandela Bay, Schoenmakerskop, South Africa	Marine	VIPEDVT
Synechococcales cyanobacterium M58 A2018 015	Porcelana Hot Spring, Los lago, Chile	Hot Spring	VIPEDVT
Synechococcales cyanobacterium C42 A2020 086	Cahuelmo Hot Spring, Los Lagos, Chile	Hot Spring	VIPEDVT

Supplementary Table S6: Previously unreported LOLIP cluster-containing cyanobacteria

Table shows strains of cyanobacteria previously unreported to be LOLIP cluster-containing, and sequence variation in the conserved 'GDITXPGGNMYP' motif

Title	Location	Habitat	GDITXPGGNMYP
Candidatus Gloeomargarita lithophora D10	Lake Alchichica, Puebla State, Mexico	Freshwater	GDITLPGGNMYP
Chroococcidiopsis cubana CCALA 043	Mineral spring, Santa Fe, Cuba	Freshwater	GDITCPGGNMYP
Chroococcales cyanobacterium IPPAS B-1203	Spring Karlovy Vary, Karlovy Vary, Czech Republic	Hot Spring	GDITCPGGNMYP
Chroococcidiopsis cubana SAG 39.79	Dry soil, Pinar del Rio, Cuba	Terrestrial	GDITCPGGNMYP
Chroococcidiopsis sp. FACHB-1243	China	Freshwater	GDITCPGGNMYP
Gloeocapsopsis crepidinum LEGE 06123	Intertidal zone of Praia da Luz, Lagos, Portugal	Marine	GDITCPGGNMYP
Fischerella sp. PCC 9605	Missing	Missing	GDITCPGGNMYP
Leptolyngbya sp. KIOST-1	Pond, Ansan, South Korea	Freshwater	GDITCPGGNMYP
Lusitaniella coriacea LEGE 07157	Tide pool, Praia de Lavadores, Canidelo, Portugal	Marine	GDITCPGGNMYP
Nodosilinea sp. P-1105	Cock Soda lake, Altai region, Russia	Hypersaline	GDITLPGGNMYP
Nodosilinea sp. LEGE 07088	Douro estuary, Porto, Portugal	Brackish	GDITLPGGNMYP
Nodosilinea sp. LEGE 07298	Douro Estuary Shore, Porto, Portugal	Brackish	GDITCPGGNMYP
Romeria aff. Gracilis LEGE 07310	Benthic zone of Minho estuary, Caminha, Portugal	Brackish	GDITLPGGNMYP
Scytonema millei VB511283	Stone, West Bengal, India	Terrestrial	GDITCPGGNMYP
Leptolyngbya cf. ectocarpi LEGE 11479	Epilithic Subtidal Sample, Plo Negro, Portugal	Marine	GDITLPGGNMYP

Supplementary Table S7: Genomics of the FARLIP cluster across five genus of cyanobacteria

Strain	Gene #	Gene	JGI Gene ID	DNA length (bp)	Protein length (aa)	NCBI Contig	NCBI Contig Accession	Start	End	Notes
Nodosilinea sp. P-1105	1	Hypothetical	2887195707	414	137	NODE_35	SMDQ01000035.1	37324	37737	
	2	PsbA3	2887195708	1089	362	NODE_35	SMDQ01000035.1	37889	38977	
	3	Hypothetical	2887195709	147	48	Absent	Absent	Absent	Absent	
	4	PsbH	2887195710	114	37	NODE_35	SMDQ01000035.1	39251	39472	
	5	PsbB2	2887195711	1533	510	NODE_35	SMDQ01000035.1	39556	41088	
	6	PsbC2	2887195712	1470	489	NODE_35	SMDQ01000035.1	41180	42607	
	7	Hypothetical	2887195713	187	62	Absent	Absent	Absent	Absent	
	8	PsbD3	2887195714	1059	352	NODE_35	SMDQ01000035.1	42825	43883	
	9	ApcD3	2887195715	531	176	NODE_35	SMDQ01000035.1	44013	44543	
	10	Hypothetical	2887195716	183	60	NODE_35	SMDQ01000035.1	44659	44841	
	11	ApcE2	2887195717	2237	778	NODE_35	SMDQ01000035.1	44838	47174	
	12	ApcD2	2887195718	480	159	NODE_35	SMDQ01000035.1	47330	47809	
	13	ApcB2	2887195719	486	161	NODE_35	SMDQ01000035.1	47920	48405	
	14	ApcA2	2887195720	477	158	NODE_35	SMDQ01000035.1	48490	48966	
	15	PsbO1	2887195721	831	276	NODE_35	SMDQ01000035.1	49673	50503	
	16	PsbF2	2887195722	135	44	NODE_35	SMDQ01000035.1	50672	50806	
	17	PsbV1	2887195723	468	155	NODE_35	SMDQ01000035.1	51103	51570	
	18	PsbA4	2887195724	1218	405	NODE_35	SMDQ01000035.1	52486	53703	
	19	PsaA2	2887195725	1698	565	NODE_35	SMDQ01000035.1	53811	55484	incomplete; missing stop
	20	Hypothetical	2887198450	306	101	NODE_98	SMDQ01000098.1	3782	4021	
	21	RfpC	2887198451	375	124	NODE_98	SMDQ01000098.1	4288	4662	
	22	RfpA	2887198452	2526	841	NODE_98	SMDQ01000098.1	4652	7177	
	23	RfpB	2887198453	1947	648	NODE_98	SMDQ01000098.1	7187	9143	
	24	PsaJ2	2887198454	132	43	NODE_98	SMDQ01000098.1	9457	9588	
	25	PsaF2	2887198455	483	160	NODE_98	SMDQ01000098.1	9621	10103	
	26	PsaI2	2887198456	201	66	NODE_98	SMDQ01000098.1	10153	10353	
	27	PsaL2	2887198457	543	180	NODE_98	SMDQ01000098.1	10632	11174	
	28	PsaB2	2887198459	1818	605	NODE_98	SMDQ01000098.1	11299	13530	
	29	PsaA2	2887198460	627	208	NODE_98	SMDQ01000098.1	13588	14271	incomplete; missing start
Plectonema cf. radiosum LEGE 06105	1	PsaJ2	2914399698	153	50	NODE_199	JADEWL010000089.1	9465	9617	
	2	PsaF2	2914399699	483	160	NODE_199	JADEWL010000089.1	9741	10223	
	3	PsbB1	2914399700	1530	509	NODE_199	JADEWL010000089.1	10331	11860	
	4	PsbC2	2914399701	1401	466	NODE_199	JADEWL010000089.1	12388	13785	
	5	IsiA regulator	2914399702	187						
	6	PsbD3	2914399703	1059	352	NODE_199	JADEWL010000089.1	13939	14997	
	7	ApcD3	2914399704	522	173	NODE_199	JADEWL010000089.1	15082	15603	
	8	ApcE2	2914399705	2343	780	NODE_199	JADEWL010000089.1	15575	17917	
	9	ApcD2	2914399706	477	158	NODE_199	JADEWL010000089.1	17985	18461	
	10	ApcB2	2914399707	486	161	NODE_199	JADEWL010000089.1	18711	19196	
	11	ApcA2	2914399708	477	158	NODE_199	JADEWL010000089.1	19312	19788	
	12	PsbA3	2914399709	1092	363	NODE_199	JADEWL010000089.1	19858	20949	
	13	RfpB	2914399710	1869	622	NODE_199	JADEWL010000089.1	21916	23784	
	14	RfpA	2914399711	2562	853	NODE_199	JADEWL010000089.1	23830	26391	
	15	RfpC	2914399712	369	122	NODE_199	JADEWL010000089.1	26602	26970	
	16	RfpC	2914399713	390	129	NODE_199	JADEWL010000089.1	27142	27531	
	17	PsbA4	2914399714	1101	366	NODE_199	JADEWL010000089.1	27911	29011	
	18	PsaA2	2914398452	2367	788	NODE_134	JADEWL010000061.1	58	2424	
	19	PsaB2	2914398453	2223	740	NODE_134	JADEWL010000061.1	2581	4803	
	20	PsaL2	2914398454	513	170	NODE_134	JADEWL010000061.1	5280	5792	

	21	PsaI2	2914398455	201	66	NODE_134	JADEWL010000061.1	5864	6064	
Nodosilinea sp. LEGE 07088	1	PsbA3	2914094381	1080	359	NODE_120	JADEWX010000057.1	13142	14221	
	2	Hypothetical	2914094382	147	48	NODE_120	JADEWX010000057.1	14365	14511	
	3	PsbH2	2914094383	219	72	NODE_120	JADEWX010000057.1	14572	14790	
	4	PsbB2	2914094384	1536	511	NODE_120	JADEWX010000057.1	14965	16500	
	5	PsbC2	2914094385	1464	487	NODE_120	JADEWX010000057.1	16664	18085	
	6	isiA regulator	2914094386	187	No data	Absent	Absent	Absent	Absent	
	7	PsbD3	2914094387	1059	352	NODE_120	JADEWX010000057.1	18215	19273	
	8	ApcD3	2914094388	534	177	NODE_120	JADEWX010000057.1	19474	20007	
	9	Hypothetical	2914094389	156	51	NODE_120	JADEWX010000057.1	20078	20233	
	10	ApcE2	2914094390	2388	765	NODE_120	JADEWX010000057.1	20230	22617	
	11	ApcD2	2914094391	480	159	NODE_120	JADEWX010000057.1	22680	23159	
	12	ApcB2	2914094392	486	161	NODE_120	JADEWX010000057.1	23404	23889	
	13	ApcA2	2914094393	477	158	NODE_120	JADEWX010000057.1	24112	24588	
	14	PsbO1	2914094394	843	280	NODE_120	JADEWX010000057.1	25175	26017	
	15	PsbF2	2914094395	135	44	NODE_120	JADEWX010000057.1	26279	26413	
	16	PsbV1	2914094396	516	171	NODE_120	JADEWX010000057.1	26504	27019	
	17	PsbA4	2914094397	1206	401	NODE_120	JADEWX010000057.1	27586	28791	
	18	PsaA2	2914094398	2355	784	NODE_120	JADEWX010000057.1	28880	31234	
	19	PsaB2	2914094423	2232	743	NODE_124	JADEWX010000059.1	121	2352	
	20	PsaL2	2914094424	570	189	NODE_124	JADEWX010000059.1	2526	3095	
	21	PsaI2	2914094425	183	60	NODE_124	JADEWX010000059.1	3218	3400	
	22	PsaF2	2914094426	483	160	NODE_124	JADEWX010000059.1	3516	3998	
	23	PsaJ2	2914094427	141	46	NODE_124	JADEWX010000059.1	4017	4157	
	24	RfpB	2914094428	1947	648	NODE_124	JADEWX010000059.1	4710	6656	
	25	RfpA	2914094429	2601	866	NODE_124	JADEWX010000059.1	6756	9281	
	26	RfpC	2914094430	375	124	NODE_124	JADEWX010000059.1	9538	9912	
	27	Hypothetical	2914094431	420	139	NODE_124	JADEWX010000059.1	10245	10664	
Nodosilinea sp. LEGE 07298	1	PsaL2	2914025487	543	180	NODE_183	JADEXE010000183.1	86	628	
	2	PsaB2	2914025488	2232	743	NODE_183	JADEXE010000183.1	771	3002	
	3	PsaA2	2914025489	2355	784	NODE_183	JADEXE010000183.1	3055	5409	
	4	PsbA4	2914025490	1218	405	NODE_183	JADEXE010000183.1	5481	6698	
	5	PsbV1	2914025491	516	171	NODE_183	JADEXE010000183.1	7529	8044	
	6	PsbF	2914025492	135	44	NODE_183	JADEXE010000183.1	8229	8363	
	7	PsbO	2914025493	831	276	NODE_183	JADEXE010000183.1	8698	9528	
	8	Hypothtetical	2914025494	174	57	NODE_183	JADEXE010000183.1	9569	9742	
	9	ApcA2	2914025495	477	158	NODE_183	JADEXE010000183.1	10264	10740	
	10	ApcB2	2914025496	486	161	NODE_183	JADEXE010000183.1	11013	11498	
	11	ApcD2	2914025497	480	159	NODE_183	JADEXE010000183.1	11623	12102	
	12	ApcE2	2914025498	1455	485	NODE_183	JADEXE010000183.1	11216	13672	incomplete; missing end
	12	ApcE2	2914026750	870	289	NODE_268	JADEXE010000264.1	1	871	incomplete; missing start
	13	Hypothtetical	2914026751	156	51	NODE_268	JADEXE010000264.1	868	1023	
	14	ApcD3	2914026752	531	176	NODE_268	JADEXE010000264.1	1053	1583	
	15	PsbD3	2914026753	1059	352	NODE_268	JADEXE010000264.1	1800	2858	
	16	PsbC2	2914026754	1470	489	NODE_268	JADEXE010000264.1	3055	4482	
	17	isiA regulator	2914026755		No data	Absent	Absent	Absent	Absent	
	18	PsbB2	2914026756	1533	510	NODE_268	JADEXE010000264.1	4749	6281	
	19	PsbH2	2914026757	222	73	NODE_268	JADEXE010000264.1	6366	6587	
	20	Hypothtetical	2914026758	147	48	NODE_268	JADEXE010000264.1	6605	6751	
	21	PsbA3	2914026759	1080	359	NODE_268	JADEXE010000264.1	6978	8057	

	22	RfpC	2914025771	375	124	NODE_19	JADEXE010000019.1	36327	36701	
	23	RfpA	2914025772	2610	869	NODE_19	JADEXE010000019.1	36691	39300	
	24	RfpB	2914025773	1980	659	NODE_19	JADEXE010000019.1	39313	41292	
	25	PsaJ2	2914025774	132	43	NODE_19	JADEXE010000019.1	42270	42401	
	26	PsaF2	2914025775	483	160	NODE_19	JADEXE010000019.1	42444	42926	
Romeria aff. Gracilis LEGE 07310	1	PsbA3	2917553946	975	324	NODE_81	JADEXG010000043.1	13833	14912	
	2	roRNA MIR1846	2917553947	81	No data	Absent	Absent	Absent	Absent	
	3	Hypothetical	2917553948	387	128	NODE_81	JADEXG010000043.1	15637	16023	
	4	Hypothetical	2917553949	147	48	NODE_81	JADEXG010000043.1	16074	16220	
	5	PsbH2	2917553950	231	76	NODE_81	JADEXG010000043.1	16300	16530	
	6	PsbB2	2917553951	1536	511	NODE_81	JADEXG010000043.1	16638	16173	
	7	PsbC2	2917553952	1449	482	NODE_81	JADEXG010000043.1	18229	19677	
	8	isiA regulator	2917553953	187	No data	Absent	Absent	Absent	Absent	
	9	PsbD3	2917553954	1059	352	NODE_81	JADEXG010000043.1	19742	20800	
	10	ApcD3	2917553955	531	176	NODE_81	JADEXG010000043.1	20853	21383	
	11	Hypothetical	2917553956	207	68	NODE_81	JADEXG010000043.1	21370	21576	
	12	ApcE2	2917553957	2400	799	NODE_81	JADEXG010000043.1	21616	24015	
	13	ApcD2	2917553958	480	159	NODE_81	JADEXG010000043.1	24082	24561	
	14	ApcB2	2917553959	161	486	NODE_81	JADEXG010000043.1	24695	25090	
	15	ApcA2	2917553960	477	158	NODE_81	JADEXG010000043.1	25119	25595	
	16	Hypothetical	2917553961	300	99	NODE_81	JADEXG010000043.1	25697	25996	
	17	PsbO1	2917553962	672	223	NODE_81	JADEXG010000043.1	25947	26774	
	18	f18 family protein	2917553963	159	52	NODE_81	JADEXG010000043.1	26839	26997	
	19	PsbV1	2917553964	462	153	NODE_81	JADEXG010000043.1	27098	27610	
	20	PsbA4	2917553965	1167	388	NODE_81	JADEXG010000043.1	28065	29231	
	21	PsaA2	2917553966	2358	785	NODE_81	JADEXG010000043.1	29312	31669	
	22	PsaB2	2917553967	2232	743	NODE_81	JADEXG010000043.1	31712	33943	
	23	PsaL2	2917553968	570	189	NODE_81	JADEXG010000043.1	34083	34652	
	24	PsaI2	2917553969	246	81	NODE_81	JADEXG010000043.1	34703	34948	
	25	PsaF2	2917553970	492	163	NODE_81	JADEXG010000043.1	35038	35529	
	26	PsaJ2	2917553971	141	46	NODE_81	JADEXG010000043.1	35567	35707	
	27	RfpB	2917553972	1968	655	NODE_81	JADEXG010000043.1	36138	38105	
	28	RfpA	2917553973	2562	853	NODE_81	JADEXG010000043.1	38186	40747	
	29	RfpC	2917553974	372	123	NODE_81	JADEXG010000043.1	40776	41147	
Richelia sp. SL 2 1	1	PsaI2	NJO26755.1	200	66	NODE_32	JAAUSO010000010.1	18305	18505	
	2	PsaL2	NJO26756.1	512	170	NODE_32	JAAUSO010000010.1	18577	19089	
	3	PsaB2	NJO26757.1	2222	740	NODE_32	JAAUSO010000010.1	19582	21804	
	4	PsaA2	No data	2362	No data	NODE_32	JAAUSO010000010.1	22104	24466	Frameshifted
	5	PsbA4	NJO26758.1	1106	368	NODE_32	JAAUSO010000010.1	24584	25690	
	6	RfpC	NJO26759.1	389	129	NODE_32	JAAUSO010000010.1	26181	26570	
	7	RfpC	No data	367	No data	NODE_32	JAAUSO010000010.1	26721	27088	Frameshifted
	8	RfpA	NJO26760.1	2561	853	NODE_32	JAAUSO010000010.1	27319	29880	
	9	RfpB	NJO26761.1	1868	622	NODE_32	JAAUSO010000010.1	29933	31801	
	10	PsbA3	NJO26762.1	1091	363	NODE_32	JAAUSO010000010.1	32532	33623	
	11	ApcA2	NJO26763.1	476	158	NODE_32	JAAUSO010000010.1	33686	34162	
	12	ApcB2	NJO26764.1	485	161	NODE_32	JAAUSO010000010.1	34293	34778	
	13	ApcD2	NJO26765.1	476	158	NODE_32	JAAUSO010000010.1	34930	35406	
	14	ApcE2	NJO26766.1	2342	780	NODE_32	JAAUSO010000010.1	35565	37907	
	15	ApcD3	NJO26767.1	521	173	NODE_32	JAAUSO010000010.1	37879	38400	
	16	PsbD3	NJO26768.1	1058	352	NODE_32	JAAUSO010000010.1	38439	39497	

	17	PsbC2	No data	1396	No data	NODE_32	JAAUSO010000010.1	39641	41037	Frameshifted
	18	PsbB1	NJO26769.1	1529	509	NODE_32	JAAUSO010000010.1	41577	43106	
	19	PsaF2	NJO26770.1	482	160	NODE_32	JAAUSO010000010.1	43315	43797	
	20	PsaJ	NJO26771.1	149	49	NODE_32	JAAUSO010000010.1	43928	44077	
Phormidesmis sp. RL 2 1	1	PsbC2	No data	1239	No data	NODE_689	JAAUPS010000151.1	1	1240	incomplete, missing N-terminus
	2	PsbC2	No data	1153	No data	NODE_218	JAAUPS010000053.1	1	1154	incomplete, missing C-terminus
	3	PsbD3	NJM97201.1	1058	352	NODE_218	JAAUPS010000053.1	1206	2264	
	4	ApcD3	NJM97202.1	551	183	NODE_218	JAAUPS010000053.1	2333	2884	
	5	ApcE2	NJM97203.1	2249	749	NODE_218	JAAUPS010000053.1	3163	5412	
	6	ApcD2	No data	478	No data	NODE_218	JAAUPS010000053.1	5502	5980	Frameshifted
	7	ApcB2	NJM97204.1	485	161	NODE_218	JAAUPS010000053.1	6038	6523	
	8	ApcA2	NJM97205.1	476	158	NODE_218	JAAUPS010000053.1	6632	7108	
	9	PsbO1	NJM97206.1	854	248	NODE_218	JAAUPS010000053.1	7387	8241	
	10	PsbF2	No data	133	No data	NODE_218	JAAUPS010000053.1	8329	8462	Frameshifted
	11	NblA	No data	221	73	NODE_218	JAAUPS010000053.1	8490	8711	Frameshifted
	12	PsbV1	NJM97208.1	530	176	NODE_218	JAAUPS010000053.1	8817	9347	
	13	PsbA3	NJM97209.1	1079	359	NODE_218	JAAUPS010000053.1	9908	10987	
	14	RfpC	NJM97210.1	371	123	NODE_218	JAAUPS010000053.1	11071	11442	
	15	RfpA	No data	2581	No data	NODE_218	JAAUPS010000053.1	11644	14225	Frameshifted
	16	RfpB	NJM97211.1	2054	684	NODE_218	JAAUPS010000053.1	14222	16276	
	17	PsaJ2	NJM97212.1	140	46	NODE_218	JAAUPS010000053.1	16702	16842	
	18	PsaF2	NJM97213.1	482	160	NODE_218	JAAUPS010000053.1	16887	17369	
	19	PsaI2	NJM97214.1	173	57	NODE_218	JAAUPS010000053.1	17416	17589	
	20	PsaL2	NJM97215.1	572	190	NODE_218	JAAUPS010000053.1	17650	18222	
	21	PsaB2	No data	2232	No data	NODE_218	JAAUPS010000053.1	18460	20692	Frameshifted
	22	PsaA2	No data	2349	No data	NODE_218	JAAUPS010000053.1	20900	23249	Frameshifted
	23	PsbA4	No data	1180	No data	NODE_218	JAAUPS010000053.1	23266	24446	Frameshifted
Richelia sp. SM1_7_0	1	PsaI2	NJM17561.1	200	66	NODE_36	JAAUTV010000007.1	27568	27768	
	2	PsaL2	NJM17562.1	512	170	NODE_36	JAAUTV010000007.1	27840	28352	
	3	PsaB2	NJM17563.1	2222	740	NODE_36	JAAUTV010000007.1	28845	31067	
	4	PsaA2	NJM17564.1	2366	788	NODE_36	JAAUTV010000007.1	31367	33733	
	5	PsbA3	NJM17565.1	1106	368	NODE_36	JAAUTV010000007.1	33851	34957	
	6	RfpC	NJM17566.1	389	129	NODE_36	JAAUTV010000007.1	35448	35837	
	7	RfpC	NJM17567.1	368	122	NODE_36	JAAUTV010000007.1	35988	36356	
	8	RfpA	NJM17568.1	2561	853	NODE_36	JAAUTV010000007.1	36587	39148	
	9	RfpB	NJM17569.1	1868	622	NODE_36	JAAUTV010000007.1	39201	41069	
	10	PsbA4	NJM17570.1	1091	363	NODE_36	JAAUTV010000007.1	41800	42891	
	11	ApcA2	NJM17571.1	476	158	NODE_36	JAAUTV010000007.1	42954	43430	
	12	ApcB2	NJM17572.1	485	161	NODE_36	JAAUTV010000007.1	43561	44046	
	13	ApcD2	NJM17573.1	476	158	NODE_36	JAAUTV010000007.1	44198	44674	
	14	ApcE2	NJM17574.1	2342	780	NODE_36	JAAUTV010000007.1	44833	47175	
	15	ApcD3	NJM17575.1	521	173	NODE_36	JAAUTV010000007.1	47147	47668	
	16	PsbD3	No data	1057	No data	NODE_36	JAAUTV010000007.1	47707	48764	Frameshifted
	17	PsbC2	NJM17576.1	1397	465	NODE_36	JAAUTV010000007.1	48894	50291	
	18	PsbB1	NJM17577.1	1529	509	NODE_36	JAAUTV010000007.1	50831	52360	
	19	PsaF2	NJM17578.1	482	160	NODE_36	JAAUTV010000007.1	52569	53051	
	20	PsaJ2	NJM17579.1	149	49	NODE_36	JAAUTV010000007.1	53182	53331	

Richelia s. RM1_1_1	1	PsaL2	NJN07430.1	149	49	NODE_50	JAAUQH010000009.1	55068	55217	
	2	PsaF2	NJN07431.1	482	160	NODE_50	JAAUQH010000009.1	55348	55830	
	3	PsbB1	NJN07432.1	1529	509	NODE_50	JAAUQH010000009.1	56039	57568	
	4	PsbC2	NJN07433.1	1397	465	NODE_50	JAAUQH010000009.1	58105	59502	
	5	PsbD3	NJN07434.1	1058	352	NODE_50	JAAUQH010000009.1	59625	60683	
	6	ApcD3	NJN07435.1	521	173	NODE_50	JAAUQH010000009.1	60722	61243	
	7	ApcE2	NJN07436.1	2342	780	NODE_50	JAAUQH010000009.1	61215	63557	
	8	ApcD2	NJN07437.1	476	158	NODE_50	JAAUQH010000009.1	63716	64192	
	9	ApcB2	NJN07438.1	485	161	NODE_50	JAAUQH010000009.1	64330	64815	
	10	ApcA2	NJN07439.1	476	158	NODE_50	JAAUQH010000009.1	64946	65422	
	11	PsbA3	no data	1090	No data	NODE_50	JAAUQH010000009.1	65485	66575	Frameshifted
	12	RfpB	NJN07440.1	1868	622	NODE_50	JAAUQH010000009.1	67306	69174	
	13	RfpA	NJN07441.1	2561	853	NODE_50	JAAUQH010000009.1	69227	71788	
	14	RfpC	NJN07442.1	368	122	NODE_50	JAAUQH010000009.1	72019	72387	
	15	RfpC	NJN07443.1	389	129	NODE_50	JAAUQH010000009.1	72538	72927	
	16	PsbA4	NJN07444.1	1106	368	NODE_50	JAAUQH010000009.1	73418	74524	
	17	PsaA2	NJN07445.1	-67634	788	NODE_50	JAAUQH010000009.1	74642	7008	
	18	PsaB2	NJN07446.1	2222	740	NODE_50	JAAUQH010000009.1	77308	79530	
	19	PsaJ2	NJN07447.1	512	170	NODE_50	JAAUQH010000009.1	80022	80534	
	20	PsaI2	NJN07448.1	200	66	NODE_50	JAAUQH010000009.1	80606	80806	

Supplementary Table S8: Genomics of LOLIP cluster across identified LLAC

Strain	Gene	Strand	JGI Gene ID
Nodosilinea sp. P-1105	LHCB	_	2887196976
	IsiX	+	2887196977
	ApcB3	_	2887196978
	ApcD4	_	2887196979
Nodosilinea sp. LEGE 07088	ApcD4	+	2914094432
	ApcB3	+	2914094433
	LHCB	-	2914094434
	IsiX	-	2914094435
Nodosilinea sp. LEGE 07298	ApcA4	+	2914028764
	ApcB3	+	2914028765
	LHCB	+	2914028766
	IsiX	-	2914028767
Romeria aff. Gracilis LEGE 07310	ApcA4	+	2917551314
	ApcB3	+	2917551315
	LHCB	+	2917551316
	IsiX	-	2917551317
Chroococcales cyanobacterium IPPAS B-1203	ApcD4	+	2882385356
	ApcB3	+	2882385357
	LHCB	-	2882385358
	IsiX	+	2882385359
Halomicronema hongdechloris C2206	LHCB	-	2758437815
	ApcB3	-	2758437816
	ApcD4	-	2758437817
Synechococcus sp. PCC 7335	ApcD4	+	647578852
	ApcB3	+	647578853
	LHCB	+	647578854
Leptolyngbya sp. PCC 6406	LHCB	-	2517690603
	IsiX	+	2517690604
	ApcB3	-	2517690605
	ApcD4	-	2517690606
Fischerella sp. PCC 9605	ApcD4	+	2516148358

	ApcB3	+	2516148359
	LHCB	-	2516148360
Xenococcus sp. PCC 7305	LHCB		2508648506
	ApcB3		2508648507
	ApcD4		2508648508
Gloeocapsopsis crepidinum LEGE 06123	ApcD4	+	2914407312
	ApcB3	+	2914407313
	LHCB	-	2914407314
	IsiX	+	2914407315
Gloeocapsa sp. PCC 7428	LHCB	+	2503796421
	ApcB3	-	2503796422
	ApcD4	-	2503796423
Chroococcidiopsis thermalis PCC 7203	LHCB	+	2503614631
	ApcB3	-	2503614632
	ApcD4	-	2503614633
Chroococcidiopsis cubana SAG 39.79	LHCB	+	2887016675
	IsiX	-	2887016676
	ApcB3	-	2887016677
	ApcD4	-	2887016678
Candidatus Gloeomargarita lithophora D10	ApcD4	+	2719355298
	ApcB3	+	2719355299
	LHCB	-	2719355300
	IsiX	+	2719355301
Chroococcidiopsis cubana CCALA 043	LHCB	+	2805381209
	IsiX	-	2805381210
	ApcB3	-	2805381211
	ApcD4	-	2805381212
Leptolyngbya sp. KIOST-1	LHCB	+	2619437333
	ApcB3	+	2619437334
	ApcD4	+	2619437335
Lusitaniella coriacea LEGE 07157	ApcD4	+	2914041606

	ApcB3	+	2914041607
	LHCB	+	2914041608
	IsiX	-	2914041609
Scytonema millei VB511283	ApcD4	+	2648590928
	ApcB3	+	2648590929
	LHCB	-	2648590930
	IsiX	+	2648590931
Chroococcidiopsis sp. FACHB-1243	ApcD4	+	2914551399
	ApcB3	+	2914551400
	LHCB	-	2914551401
	IsiX	+	2914551402
Chlorogloeopsis fritschii PCC 6912	LHCB	+	2509834705
	ApcB3	-	2509834706
	ApcD4	-	2509834707
	LHCB	+	2512514087
	ApcB3	-	2512514088
	ApcD4	-	2512514089
	ApcD4	+	2551969797
	ApcB3	+	2551969798
	LHCB	-	2551969799
	IsiX	+	2551969800
Chlorogloeopsis fritschii PCC 9212	LHCB	+	2550829379
	IsiX	-	2550829380
	ApcB3	-	2550829381
	ApcD4	-	2550829382
Leptolyngbya sp. FACHB-60	ApcD4	+	2914454960
	ApcB3	+	2914454961
	LHCB	+	2914454962
	IsiX	-	2914454961
Phormidium sp. FACHB-77	ApcD4	+	2920022191
	ApcB3	+	2920022192
	LHCB	+	2920022193
	IsiX	-	2920022194

Cyanosarcina cf. burmensis CCALA 770	ApcB3	-	2789933144
	ApcD4	-	2789933145
Pleurocapsales cyanobacterium LEGE 10410	ApcB3	-	2913838106
	ApcD4	-	2913838107
Pleurocapsa sp. CCALA 161	ApcB3	-	2790191999
	ApcD4	-	2790192000
Leptolyngbya cf. ectocarpi LEGE 11479	ApcB3	-	2914343713
	ApcD4	-	2914343714
	LHCB	-	2914343711
	IsiX	+	2914343712

Supplementary Table S9: Astrobiology-relevant genes in FARLIP and LOLIP capable strains of cyanobacteria

Table shows presence or absence of selected genes, as indicated by a protein ID or blank cell. In the case of hits, JGI/IMG Gene ID or NCBI Protein ID have been recorded from blastp hits on the respective databases. LOLIP, L; FARLIP, F; LOLIP & FARLIP, LF

Strain	Photosynthetic capability	Genes																						
		Nitrogen Fixation			Carbon Fixation		Sucrose Synthesis			Trehalose Synthesis			Desaturases				Hydrogen Production						
		NifH	NifD	NifK	Carbonic Anhydrase	Sus	SPP	TreY	TreZ	TreS	DesA	DesB	DesC	DesD	HoxF	HoxU	HoxY	HoxH				
Calothrix parasitica NIES-267	F	2776204593	2776204598	2776204639		2776202022	2776204720	2776198996	2776202488	2776202489	2776202492		2776201583	2776201584	2776201582	2776201586								
Calothrix sp. NIES-3974	F	2776197526	2776197527	2776197538		2776195633	2776195493	2776195493					2776195994	2776195992	2776196246	2776195993	2776197346	2776197349	2776197350	2776197351				
Calothrix sp. PCC 7507	F	2505798331	2505798332	2505798362	4; 2505799366;	2505801578;	2505799663	2505803787	2505800818				2505800862	2505800861	2505800863									
Chlorogloeopsis fritschii PCC 6912	LF	2509828926	2509828925	2509828914	3; 2509830236;	2509833759;	2509834714	2509831345	2509831309	2509830720	2509830721	2509830394		2509828526	2509828527	2509833888		2509830734	2509830731	2509830730	2509830727			
Chlorogloeopsis fritschii PCC 9212	LF	2550827568	2550827567	2550827556	5; 2550829389;	2550832398;	2882382998	2550830650	2550833529	2550832381	2550832383	2550828467		2550827909	2550827910	2550831991		2550830167	2550830164	2550830163	2550830160			
Chroococcales cyanobacterium IPPAS B-1203	L					2882382998	2882383401	2882387089								2882385156		2882386445		2882388533				
Chroococcidiopsidales cyanobacterium LEGE 13417	F	!BE9019454.1	!BE9019453.1	!BE9019946.1		MBE9018252.1	MBE9016505.1	!BE9018275.1	!BE9018406.1	!BE9018405.1	!BE9020727.1		MBE9019783.1	MBE9019908.1	MBE9016777.1			MBE9017230.1	MBE9017244.1	MBE9017243.1				
Chroococcidiopsis cubana CCALA 043	LF	2805386811		2805387202	2; 2805383319;	2805387005;	2805384519	2805384143	2805383196	2805384164	2805384165	2805383125	2805382536	2805382250	2805383632		2805386211	2805386210	2805384675	2805384674				
Chroococcidiopsis cubana SAG 39.79	LF	2887015171	2887015169	2887015168	3; 2887020303;	2887020089;	2887017310	2887014248	2887017053	2887014658	2887014659	2887020836	2887015895	2805382250	2887016167		2887021364	2887021365	2887020390	2887020391				
Chroococcidiopsis sp. CCALA 051	F	015156515.1	106548296.1	106548297.1		WP_106544156.1	PSM46832.1	106547831.1	106547162.1	106547163.1	106546821.1	WP_106548325.1	WP_106547572.1	WP_106546357.1		WP_106546073.1	WP_106546074.1	WP_106544562.1	WP_106544561.1					
Chroococcidiopsis sp. FACHB-1243	LF				4; 2914554149;	2914549670;	2914550629	2914551012	2914555416	2914552649	2914552650	2914551700	2914550087	2914549720	2914551870		2914553664	2914553662	2914552450	2914552449				
Chroococcidiopsis thermalis PCC 7203	LF	2503614865	2503614864	2503614863	2; 2503615300;	2503613073;	2503613567	2503615614	2503615718	2503611768	2503611769	2503611125	2503610357	2503613025	2503614413		2503611411	2503611410	2503612091	2503612092				
Cyanobacterium TDX16	F					OWY69404.1	OWY69233.1	OWY67811.1	OWY67530.1	OWY67531.1	OWY67546.1		OWY66668.1	OWY69411.1	OWY64669.1	OWY70506.1								
Cyanosarcina cf. burmensis CCALA 770	F	2789931878	2789931879	2789931880	4; 2789936905;	2789931523;	2789935139	2789935540	2789933956	2789934123	2789934122	2789935981	2789937289	2789933316	2789935922		2789933415	2789933416	2789932176	2789932175				
Fischerella major NIES-592	F	OKH13016.1	0735562131	0735562141		WP_062249732.1	WP_073556720.1	062245134.1	073554643.1	073554644.1	073554648.1				WP_062245706.1		WP_073555610.1	WP_073555609.1	WP_073555608.1	WP_073555567.1				
Fischerella muscicola CCMEE 5323	F	2808401232	2808401233	2808401234	9; 2808403542;	2808401330;	2808402992	2808401189	2808401087	2808398490	2808398489	2808398603			2808399177									
Fischerella muscicola PCC 7414	F	2550705375	2550705376	2550705377	0; 2550709304;	2550706682;	2550705453	2550708994	2550709302	2550709470	2550709469	2550709464			2550709117									
Fischerella sp. FACHB-380	F	2909456297	2909456296	2909456295	1; 2909458073;	2909452581;	2909455614	2909456633	2909455616	2909455432	2909455433	2909455440			2909455099		2909458313							
Fischerella sp. NIES-3754	F	2687481206	2687481207	2687481208	2687483352;	2687485726;	2687481145	2687483457	2687485075	2687483155	2687483154	2687483151			2687483568		2687482469	2687482471	2687482472	2687485375				
Fischerella sp. NIES-4106	F	2776277546	2776277545	2776277544	3; 2776274447;	2776273175;	2776279056	2776275772	2776275497	2776276366	2776276367	2776275257	2776274699		2776276370									
Fischerella sp. PCC 9605	F	2516149990	2516149991	2516149992	5; 2516144491;	2516143996;	2516150016	2516148239	2516143894	2516146868	2516146869	2516147286			2516146902		2516148201	2516148199	2516148198	2516148195				
*Fischerella thermalis BR2B**	F	2805784456	2805784455	2805784454		2805783659;	2805782779	2805783890	2805785549	2805783802	2805783803	2805783805			2805783558		2805783105	2805783106	2805783107	2805783507				
Fischerella thermalis CCMEE 5194	F	2805971820	2805971819	2805971818	2805973975;	2805971558;	2805973786	2805971339	2805970600	2805972425	2805972426	2805972486			2805970313		2805973477	2805973476	2805973475	2805972312				
Fischerella thermalis CCMEE 5196	F	2805963960	2805964921	2805964922		2805961252;	2805961937	2805965543	2805961840	2805963415	2805963479	2805965429			2805963345		2805961542	2805965343	2805965344	2805963971				
Fischerella thermalis CCMEE 5198	F	2805799346	2805799345	2805799344		2805798718;	2805799614	2805801348	2805799265	2805798944	2805798945	2805798750			2805800824		2805800182	2805800927	2805800926	2805798804				
Fischerella thermalis CCMEE 5201	F	2805905247	2805905246	2805905245	5; 2805907451;	2805905572;	2805909289	2805904652	2805904641	2805905805	2805905806	2805906841			2805905862		2805905214	2805905656	2805905657	2805905870				
Fischerella thermalis CCMEE 5205	F	2805910228	2805910229	2805910230	3; 2805913221;	2805912034;	2805913666	2805914209	2805914090	2805909765	2805909766	2805911751			2805914162		2805914670	2805914671	2805914672	2805914167				
Fischerella thermalis CCMEE 5208	F	2805865572	2805865573	2805865574		2805868240;	2805866985	2805868776	2805865719	2805868067	2805868068	2805868070			2805868785		2805867492	2805867493	2805867494	2805865118				
Fischerella thermalis CCMEE 5268	F	2805779772	2805779773	2805779774	2805779955;	2805781512;	2805778015	2516148239	2805778535	2805778873	2805778872	2805778243			2805779139		2805777493	2805777494	2805777495	2805779133				
Fischerella thermalis CCMEE 5273	F	2805790419	2805790420	2805790421	2805788955;	2805789467;	2805790744	2805788349	2805792006	2805788053	2805788054	2805788089			2805789697		2805791957	2805791957	2805791958	2805789692				
Fischerella thermalis CCMEE 5282	F	2805794038	2805794039	2805794040	2805795998;	2805793866;	2805796760	2805794020	2805794623	2805793900	2805793901	2805793903			2805792295		2805794418	2805794419	2805794420	2805792300				
Fischerella thermalis CCMEE 5318	F	2805811328	2805811329	2805811330	2805811400;	2805809198;	2805967521	2805811860	2805807434	2805811881	2805811882	2805809375			2805812394		2805811930	2805808120	2805808119	2805812385				
Fischerella thermalis CCMEE 5328	F	2805965733	2805965734	2805965735		2805966463	2805966998	2805968134	2805967896	2805968756	2805967601				2805967250		2805966127	2805966128	2805966129	2805967291				
Fischerella thermalis CCMEE 5330	F	2805871432	2805871433	2805871434	2805873508;	2805870636;	2805874472	2805873635	2805870528	2805870762	2805872643	2805871557			2805869984		2805872024	2805872025	2805874029	2805869977				
Fischerella thermalis JSC-11	F	2505767548	2505767547	2505767546		2505767162;	2505770947	2505767057	2505768235	2505770639	2505770640	2505770642			2505768423		2505767250	2505767251	2505767252	2505768962				
Fischerella thermalis PCC 7521	F	2550718611	2550718610	2550718609		2550715280;	2550718678;	2550717258	2550716352	2550715471	2550715472	2550715474			2550719897		2550716112	2550716113	2550716114	2550719417				
Fischerella thermalis WC1110	F	2805829624	2805829623	2805829622		2805831695;	2805830573	2805830286	2805829875	2805832907	2805832908	2805832910			2805829797		2805829242	2805829243	2805829244	2805831463				
Fischerella thermalis WC114	F	2805835751	2805835752	2805835753		2805836122;	2805838029	2805835700	2805834101	2805833787	2805833786	2805834357			2805837755		2805833988	2805833987	2805833986	2805837750				
Fischerella thermalis WC119	F	2805841134	2805841133	2805841132		2805840958;	2805841306	2805840371	2805839742	2805838279	2805838278	2805838321			2805840060		2805838348	2805838349	2805838350	2805838527				
Fischerella thermalis WC157	F	172451471.1	009459215.1	009459214.1		WP_009453673.1	WP_102146430.1	009457872.1	009455287.1	009455288.1	009455290.1						WP_102147914.1	WP_009459822.1	WP_009459823.1	WP_009756785.1				
Fischerella thermalis WC213	F	2806023873	2806023874	2806023875		2805841306;	2806024412	2806022634	2806025012	2806023581	2806023582	2806023584			2806022981		2806025884	2806025883	2806025882	2806025603				
Fischerella thermalis WC245	F	2806312656	2806312657	2806312658		2806314589;	2806314266	2806314922	2806312398	2806312117	2806312118	2806310591			2806310991		2806310634	2806310635	2806310636	2806310996				
Fischerella thermalis WC246	F	2806089169	2806089170	2806089171		2806091891;	2806089335	2806087732	2806088196	2806089572	2806089573	2806087768			2806091674		2806090181	2806090180	2806090179	2806091669				
Fischerella thermalis WC249	F	2805882207	2805882206	2805882205		2805880712;	2805882121	2805879917	2805883063	2805880141	2805880142	2805882403			2805883045		2805880378	2805880377	2805880376	2805883040				
Fischerella thermalis WC341	F	2805825396	2805825397	2805825398		2805826065;	2805824312	2805825359	2805825435	2805827749	2805827748	2805827746			2805824588		2805824366	2805824365	2805824364	2805824753				
Fischerella thermalis WC344	F	172451471.1	009459215.1	009459214.1		WP_009453673.1	WP_102147247.1	009457872.1	009455287.1	009455288.1	009455290.1						WP_102146203.1	WP_009459822.1	WP_009459823.1	WP_009756785.1				
Fischerella thermalis WC439	F	172451471.1	009459215.1	009459214.1		WP_009459673.1	WP_009459985.1	009457872.1	009455287.1	009455288.1	016870765.1						WP_102146203.1	WP_009459822.1	AII19891.1	WP_009756785.1				
Fischerella thermalis WC442	F	172451471.1	009459215.1	009459214.1		WP_009459673.1	WP_102146430.1	009457872.1	009455287.1	009455288.1	016870765.1						WP_180963282.1	WP_009459822.1	WP_009459823.1	WP_009756785.1				
Fischerella thermalis WC527	F	2805895265	2805895264	2805895263		2805898913;	2805896720	2805894509	2805894263	2805896266	2805896265	2805896263			2805895018		2805894618	2805894617	2805894616	2805895013				
Fischerella thermalis WC538	F	2805891433	2805891434	2805891435		2805890414;	2805894037	2805889691	2805889704	2805890425	2805890424	2805891579			2805894189		2805892455	2805892454	2805892453	2805894194				
Fischerella thermalis WC542	F	2805887202	2805887201	2805887200		2805887132;	2805886058	2805884811	2805887998	2805885000	2805884999	2805887068			2805887244		2805885559	2805885558	2805885556	2805887249				
Fischerella thermalis WC558	F	2805851384	2805851383	2805851382		2805851293;	2805850603	2805849125	2805852491	2805849359	2805849358	2805851698			2805849010		2805849085	2805849084	2805849083	2805849016				
Fischerella thermalis WC559	F	2805855732	2805855731	2805855730		2805855659;	2805858156	2805853725	2805856540	2805856090	2805856091	2805856093			2805856800		2805854891	2805854890	2805854889	2805856805				
Gloeocapsa sp. PCC 7428	L					2503795313;	2503795818	2503792959	2503793328			2503794705	2503794200		2503792593									
Gloeocapsopsis crepidinum LEGE 06123	L	2914407434	2914407433	2914407432		2914409264;	2914406894	2914409405	2914408935			2914407459	2914408800		2914405718		2914409911	2914409916	2914409917	2914409918				
Halomicronema hongdechloris C2206	LF					2758435702				2758435385	2758435386	2758437140			2758440042									
Hydrococcus rivularis NIES-593	F	073598931.1	073599038.1	073598932.1		WP_073599500.1	WP_073601003.1	073599884.1					WP_073600969.1		WP_073597934.1		WP_073598128.1	WP_073598127.1	WP_073598126.1	WP_073598125.1				
Leptolyngbya cf. ectocarpi LEGE 11479	F					2914340102							2914339776	2914340495	2914342682		2914342429	2914344042	2914344043	2914343991				
Leptolyngbya sp. FACHB-60	L					2914458465;	2914456182			2914458800	2914458801	2914456076	2914458990	2914458793	2914457948									
Leptolyngbya sp. IPPAS B-1204	F	RNJ66749.1	RNJ66748.1	NJ66747.1		RNJ64888.1	RNJ69173.1		RNJ68677.1	RNJ64832.1	RNJ66987.1	RNJ68672.1												
Leptolyngbya sp. JSC-1	F	2022832730	2022832731	2022832732		2022831142;	2022830730		2022833364	2022831713	2022832905	2022833370			2022832171									
Leptolyngbya sp. KIOST-1	L	2619437837	2619437838	2619437839	2619442700;	2619437530;	2619441605		2619441045	2619441046	2619441760	2619439923	2619442217		2619438455		2619438460		2619438461	2619438463				
Leptolyngbya sp. PCC 6406	L	2517693975	2517693976	2517693977		2517692023						2517691213			2517691215									
Lusitaniella coriacea LEGE 07157	L					2914042576	2914040656	2914044369				2914042422	2914042422		2914044269									
Mastigocoleus testarum BC008	F	2724943698	2724943697	2724942722	2724944600;	2724937102;	2724944642	2724942571	2724942528	2724937242	2724937243	2724937079	2724943001	2724943000	2724943002									
Nodosilinea sp. LEGE 07088	LF	2914094847	2914094848	2914095918		2914099109				2914097680	2914097681		2914094154	2914097673	2914098833		2914098669	2914098666	2914098665	2914098658				
Nodosilinea sp. LEGE 07298	LF	2914028723	2914028724	2914028725		2914025621;	2914027716			2914025160	2914025159	2914025840	2914026009	2887193727	2914028119									
Nodosilinea sp. P-1105	LF	2887195446	2887195448	2887195449		2887198354;	2887198297			2887195144	2887195143	2887193359	2887196741	2887193727	2887194132		2887193013	2887193010	2887193009	2887193007				
Oscillatoriales cyanobacterium JSC-12	F	2510098635	2510098636	2510098637		2510097408;	2510096363	2510096732	2510098095				2510096834		2510095651									
Phormidium sp. FACHB-77	L					2920022893;	2920024388			2920025257	2920025258	2920025541	2920023671	2920025250	2920022465									
Plectonema cf. radiosum LEGE 06105	F	2517693975	2914402970	2517693977		2913840883	2914399434	2914400993	2914400660	2914400659	2914401410		2914400638	2914400639	2914402138	2914400640								
Pleurocapsa sp. CCALA 161	F				2790194682;	2790196040;	2790196854	2790191935	2790194730	2790195711	2790195712		2790193646		2790192302									
Pleurocapsa minor PCC 7327	F	2509575869	2509575869	2509575868		2509572431	2509572554	2509575124					2509575014		2509571943		2509574563	2509574565	2509574567	2509574568				
Pleurocapsales cyanobacterium LEGE 10410	F	2913843793	2913843794	2913843795		2913840883	2790191935	2913840046					2913843498		2913838192		2913841233	2913841906	2913841905	2913844285				
Romeria aff. gracilis LEGE 07310	LF					2917549276			2917551142	2917551143	2917551001		2917554029		2917552691									
Scytonema millei VB511283	L				2; 2648594441;	2648589857;	2648589769	2648590192	2648592645	2648593322	2648593321	2648590003	2648591692	2648594385	2648593820		2648592768	2648592770	2648593009	2648593008				
Synechococcus sp. PCC 7335	LF	647576434	647576433	647576432		647576804			647580233	647580234	647580830		647579295	647580248	647579347		647580720	647580790	647580789	647580787				
Xenococcus sp. PCC 7305	L	2508648322	2508648323	2508648324		2508653162		2508650263			2508650821		2508651115		2508648698		2508648401	2508648372	2508648371	2508648369				