



OPEN

## Mycosporine-like amino acid and aromatic amino acid transcriptome response to UV and far-red light in the cyanobacterium *Chlorogloeopsis fritschii* PCC 6912

Carole A. Llewellyn<sup>1</sup>✉, Carolyn Greig<sup>1</sup>, Alla Silkina<sup>1</sup>, Bethan Kultschar<sup>1</sup>, Matthew D. Hitchings<sup>2</sup> & Garry Farnham<sup>3</sup>

The “UV sunscreen” compounds, the mycosporine-like amino acids (MAAs) are widely reported in cyanobacteria and are known to be induced under ultra-violet (UV) light. However, the impact of far red (FR) light on MAA biosynthesis has not been studied. We report results from two experiments measuring transcriptional regulation of MAA and aromatic amino acid pathways in the filamentous cyanobacterium *Chlorogloeopsis fritschii* PCC 6912. The first experiment, comparing UV with white light, shows the expected upregulation of the characteristic MAA *mys* gene cluster. The second experiment, comparing FR with white light, shows that three genes of the four *mys* gene cluster encoding up to mycosporine-glycine are also upregulated under FR light. This is a new discovery. We observed corresponding increases in MAAs under FR light using HPLC analysis. The tryptophan pathway was upregulated under UV, with no change under FR. The tyrosine and phenylalanine pathways were unaltered under both conditions. However, nitrate ABC transporter genes were upregulated under UV and FR light indicating increased nitrogen requirement under both light conditions. The discovery that MAAs are upregulated under FR light supports MAAs playing a role in photon dissipation and thermoregulation with a possible role in contributing to Earth surface temperature regulation.

Cyanobacteria, the first evolutionary extant photosynthetic organisms, have evolved to adapt to a wide range of variable light environments and can survive exposure to both the ultraviolet (UV) and far-red (FR) ends of the visible light spectrum<sup>1,2</sup>. Exposure to natural levels of UV in cyanobacteria has been shown to have a wide range of physiological and biochemical effects, triggering changes in the photosystems, antioxidants and radical scavenging, and the induction of mycosporine amino acids (MAAs)<sup>3,4</sup>. Exposure to light at the FR end of the light spectrum has also been shown to have an extensive acclimatory affect on the genome of cyanobacteria<sup>2</sup>.

Commonly recognised for their UV suncreening and photoprotection ability are a group of > 30 related low molecular weight aromatic amino acid derivatives called the MAAs. MAAs are a widely distributed in high-light marine, freshwater and terrestrial environments<sup>5,6</sup>. They are found in highest abundance within prokaryotic cyanobacteria and within the eukaryotic rhodophytes and dinoflagellates with highest concentrations often found in cells exposed to UV light<sup>5,7,8</sup>. Their absorption maxima range across the UV-A and UV-B regions from 310–360 nm with correspondingly high absorption coefficients ( $\epsilon = 28,100\text{--}50,000\text{ M}^{-1}\text{ cm}^{-1}$ )<sup>5,9</sup>. Their strong UV absorption and photostability is not only important ecologically, but has led to an interest in MAAs as natural sunscreens to replace or supplement existing synthetic based sunscreens<sup>10,11</sup>. Other functions assigned to MAAs aside that of suncreening includes antioxidant activity and osmoregulation; these have been reported for the

<sup>1</sup>Faculty of Science and Engineering, Swansea University, Swansea SA2 8PP, UK. <sup>2</sup>Medical School, Swansea University, Swansea SA2 8PP, UK. <sup>3</sup>Faculty of Medicine and Dentistry, Plymouth University, Plymouth PL4 8AA, UK. ✉email: c.a.llewellyn@swansea.ac.uk

filamentous cyanobacteria species *Lyngbya* sp. CU2555<sup>12</sup>, *Nostoc* sp. R76DM<sup>13</sup> and *Chlorogloeopsis fritschii* PCC 6912<sup>6,14</sup>.

The properties of mycosporine-glycine as an oxy-mycosporine amino acid are reported to be different to those of imino-mycosporines such as shinorine. In a study on *C. fritschii* PCC 6912 comparing UV and osmotic stress, UV-B was found to strongly enhance the accumulation of shinorine, whereas osmotic stress had a more pronounced effect on mycosporine-glycine<sup>14</sup>. In a study on the halotolerant cyanobacterium *Aphanothece halophytica*, mycosporine-2-glycine was found to be induced more under high salinity conditions than under UV-B<sup>15</sup>. In addition, mycosporine-glycine has been found to have strong antioxidant properties whereas MAAs substituted with two amino acids such as shinorine do not<sup>5</sup>. The ability of MAAs to absorb energy without producing reactive oxygen species also suggests that MAAs could have a role in photon-dissipation<sup>16</sup>.

As aromatic amino acid derivatives, MAAs have a core cyclohexenone or cyclohexenimine structure. These have a methoxy group at C-2 position and substitution at C-3 with an amino acid or an imino alcohol giving the oxy-mycosporines such as mycosporine-glycine; also a substitution at C-1 with a second imino group leads to a wider group of imino-mycosporines including the commonly occurring shinorine. Understanding the biosynthetic pathways to the MAAs has been explored using the growing availability of genome sequences across many species of cyanobacteria, including for the more genetically complex filamentous species<sup>8</sup>. Comparative genome analysis on available sequenced genomes reveals that a four gene *mys* cluster (*mysA*, *mysB*, *mysC* and *mysD* or *E*), involving four steps leading to the production of MAAs is commonly present. The *mys* cluster has been found to occur commonly across filamentous cyanobacteria<sup>17</sup>.

The first biosynthetic step to MAAs is the production of desmethyl 4-deoxygadusol (DDG) and 4-deoxygadusol (4-DG). This can either occur via the shikimate pathway or more directly from sedoheptulose 7-phosphate<sup>18,19</sup>. In the shikimate pathway DHQ synthase (DHQS) catalyses conversion of DHQ to DDG, and in pentose phosphate pathway, desmethyl-4-deoxygadusol synthase (DDGS, a DHQS homolog) catalyses the production of DDG<sup>17–19</sup>. In the filamentous strain *Anabaena variabilis* ATCC 29413, both the pentose phosphate and shikimate pathways have been found to be linked to MAA biosynthesis<sup>19</sup>.

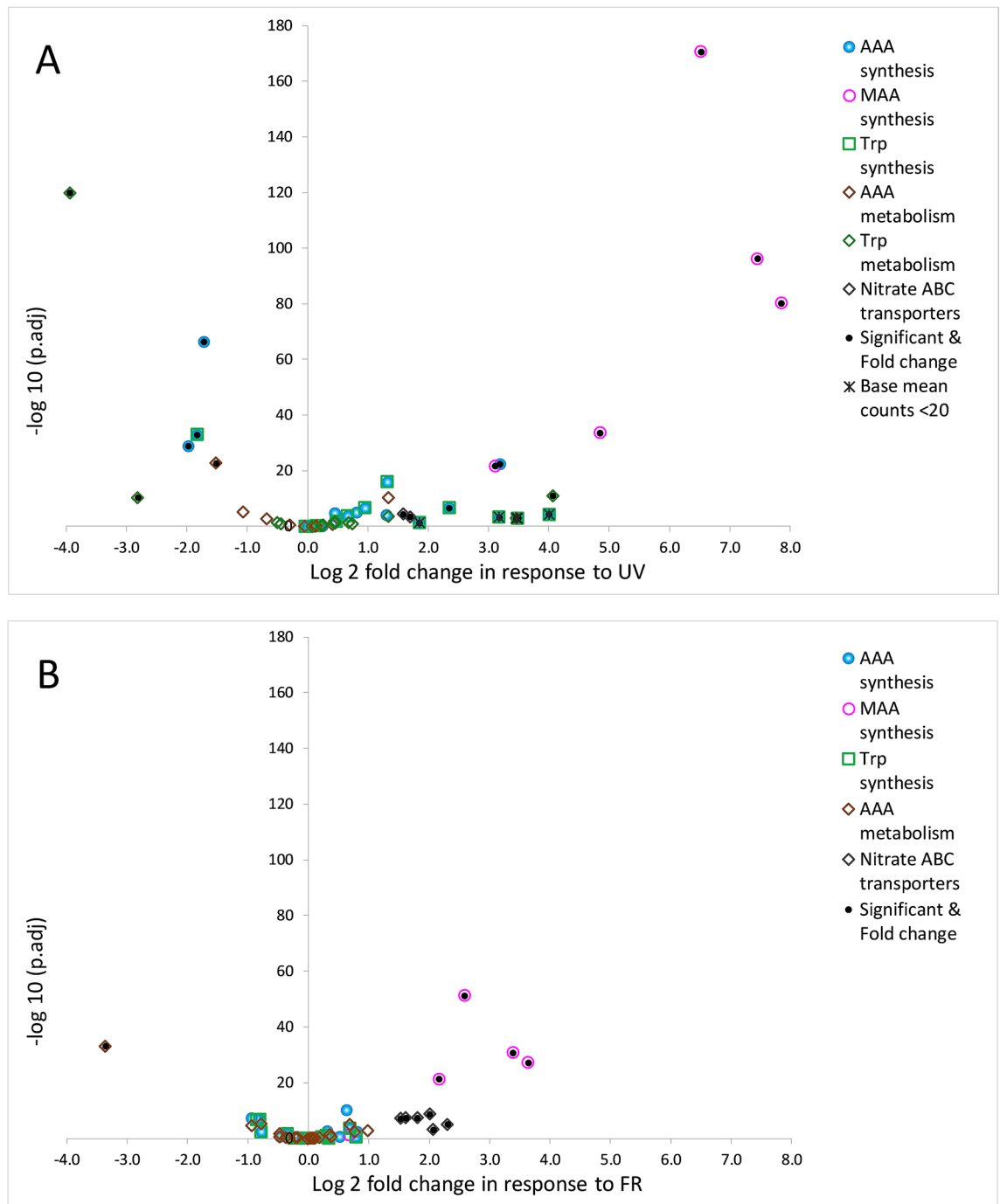
In the second step, the O-methyltransferase (*MysB*) converts DDG to 4-DG. In the third step, 4-DG incorporates the first nitrogen using an ATG grasp (*MysC*) to produce mycosporine-glycine, which is the branch point for all other MAAs. In the fourth step, mycosporine-glycine is converted to the imino-mycosporine, shinorine in *C. fritschii* PCC 6912 and in *A. variabilis* ATCC 29413, with a non-ribosomal peptide synthetase (NRPS)-like enzyme (*MysE*)<sup>18,20,21</sup>, which contains adenylation, thiolation and thioesterase domains. The amino acid adenylation domain has the role of attaching the second amino group to the C1 position of mycosporine-glycine to produce the imino-mycosporine. In *Nostoc punctiforme* ATCC29133 and *Lyngbya* sp. PCC8106 a D-alanine D-alanine ligase homolog (*MysD*) replaces the NRPS-like *MysE* enzyme<sup>22</sup>.

The aromatic amino acids (AAAs) share biosynthetic pathways with the MAAs and are synthesised via the shikimate pathway with a branch point at dehydroquinone (3-DHQ). The shikimate pathway to production of AAAs is one of the most evolutionarily ancient metabolic pathways, and has been well studied in bacteria, algae and plants. It is estimated that > 30% of photosynthetically fixed carbon in vascular plants is directed from the shikimate pathway into tryptophan, phenylalanine and tyrosine<sup>23</sup>. Chorismate is the branch point for pathways to tryptophan, phenylalanine and tyrosine. AAAs are the building blocks for protein synthesis. They are also precursors to a wide range of essential end products such as pigments, hormones, vitamins and cell wall components, and serve as the precursors for the synthesis of many biologically and neurologically active compounds that are essential for maintaining normal biological functions<sup>23</sup>. For example, chorismate is required for the synthesis of folate (vitamin B9), tocopherols (vitamin E) and vitamin K. As a result AAAs are drug targets for human diseases including neurodegenerative diseases, schizophrenia, and cancers. Many of the downstream metabolites of AAAs are also important as natural products and in human medicine and nutrition<sup>23</sup>.

In particular, tryptophan is known to play a number of stress response and UV protective roles within plant and animal cells. Tryptophan is the amino acid precursor to serotonin and melanin. It has also been highlighted for its potential role in protection and cell signalling response regulator towards UV in plant cells. In *Arabidopsis*, tryptophan has been found to be part of a UV-B receptor known as UV-B resistance 8 (UVR8) serving as the UV-B chromophore which triggers a signalling pathway for UV protection<sup>24,25</sup>. Also involved in light and nutrient stress are the tryptophan-rich sensory proteins (TSPO); these are membrane proteins found across a range of organisms from bacteria to humans including filamentous cyanobacteria<sup>26</sup>. However, the regulation of tryptophan and other AAAs is still poorly understood.

The adaption of cyanobacteria to regions of the light spectrum aside from UV has also been studied. Red and green light adaption has been widely studied in terms of complimentary chromatic adaption (CCA). In CCA cells are able to rapidly adapt their photosystems (PSI, PSII) and phycobilisome (PBS) to red or green light by altering their complement of phycobiliproteins with green light absorbing phycoerythrin or red light absorbing phycocyanin<sup>2</sup>. A major photo-acclimative response including the synthesis of FR absorbing pigments chlorophylls *d* and *f* have been observed in the filamentous cyanobacteria *Leptolyngbya* sp strain JSC-1 when transferred from white to red including FR light (> 700 nm)<sup>27</sup>. The FR absorbing chlorophylls *d* and *f* have also been observed in *C. fritschii*<sup>28,29</sup>. More widely, FR has been shown to induce non-photochemical quenching in the macroalgae *Ulva prolifera*<sup>30</sup>, and to protect against PSI photoinhibition in *Arabidopsis thaliana*<sup>31</sup>. However, we still know relatively little about how cyanobacteria regulate response to changing light conditions. More specifically there have been no studies to assess the response of MAA and AAA pathways under FR light in cyanobacteria.

Here we report two separate transcription profiling experiments focussing on the regulation of the MAA and AAA pathways in *C. fritschii* PCC 6912. *C. fritschii* PCC 6912 (Class Nostocaceae) is a Section-V filamentous strain originally isolated from soil in Allahabad, India<sup>32</sup>. *C. fritschii* is known to contain two MAAs, mycosporine-glycine and shinorine and these are inducible under UV<sup>14</sup>. In the first experiment, we explore the effects of supplementary low levels of UV-B (representing natural levels rather than higher and more damaging levels) on



**Figure 1.** Volcano plots showing significance versus log<sub>2</sub> fold change of expression in genes in *C. fritschii* PCC 6912 associated with exposure to (A) Low level UV-B and (B) FR Filled circles indicate significance with  $-1.3 > \log_2 \text{fold change} > 1.3$   $p_{\text{adj}} < 0.05$ .

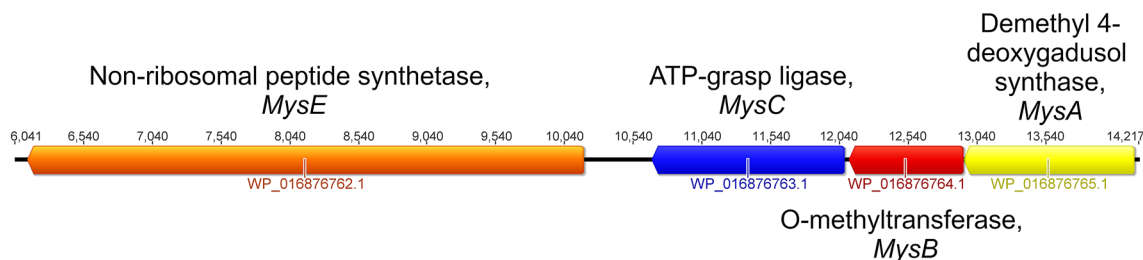
the MAA and AAA pathways. In the second experiment, we explore the effects of exposure to FR light on the same pathways. Both experiments use white light as the control. The experiments were carried out using different sequencing and bioinformatic techniques due to experiments being undertaken at different times and places. Whilst the results from the two experiments are not directly comparable they each provide unique insights into MAA and AAA pathways and the results from the FR light experiment reveals novel and unexpected findings.

## Results

The overall change in expression of genes in *C. fritschii* PCC 6912 exposed to UV-B and FR is shown in Fig. 1; this highlights the significant changes found associated with the MAA and AAA pathways. Associated details on the identified and significantly up and downregulated gene homologs associated with MAAs, AAAs and nitrogen transport are listed in Table 1.

Product Accession	Gene name	KEGG Description	Log2 fold change		Code
			FR	UV	
<b>3-DHQ synthesis</b>					
WP_016875297.1	<i>aroA2.1</i>	K03856 AROA2, <i>aroA</i> : DAHP synthase; 3-deoxy-7-phosphoheptulonate synthase [EC:2.5.1.54]		3.15	UV upregulated
WP_016875151.1	<i>aroA2.2</i>	K03856 AROA2, <i>aroA</i> : DAHP synthase; 3-deoxy-7-phosphoheptulonate synthase [EC:2.5.1.54]			Gene identified: NDR
WP_016874070.1	<i>aroA</i>	K03856 AROA2, <i>aroA</i> : DAHP synthase; 3-deoxy-7-phosphoheptulonate synthase [EC:2.5.1.54]		-1.98	UV downregulated
WP_016876414.1	<i>aroB</i>	K01735 <i>aroB</i> : DHOS; 3-dehydroquinate synthase [EC:4.2.3.4]			Gene identified: NDR
<b>MAA synthesis</b>					
WP_016876765.1	<i>mysA</i>	K21342 DDGS; demethyl-4-deoxygadusol synthase [EC:4.2.3.154]		2.51	UV/FR upregulated
WP_016876764.1	<i>mysB</i>	K00588 <i>caffoyl</i> -CoA-O-methyltransferase [EC:2.1.1.104]		1.51	UV/FR upregulated
WP_016876763.1	<i>mysC</i>	ATP-grasp ligase forming mycosporine-glycine, Msc no KO assigned		3.55	UV/FR upregulated
WP_016876762.1	<i>mysE</i>	Nonribosomal peptide synthetase MscE no KO assigned		4.79	UV upregulated
<b>AAA synthesis</b>					
WP_016876992.1	<i>aroQ</i>	K03786 <i>aroQ</i> , <i>qutE</i> : type II 3-dehydroquinate dehydratase [EC:4.2.1.10]		-1.73	UV downregulated
WP_016876990.1	<i>aroE</i>	K00014 <i>aroE</i> : shikimate dehydrogenase [EC:1.1.1.25]		2.06	UV/FR upregulated
WP_016875893.1	<i>aroK</i>	K00891 <i>aroK</i> , <i>anol</i> : shikimate kinase COG0709 [EC:2.7.1.71]		3.09	Gene identified: NDR
WP_016878669.1	<i>ark</i>	Shikimate kinase COG0683 homologous to K00939 adenylate kinase [EC:2.7.4.3]		2.01	UV upregulated
WP_016876217.1	<i>aroA</i>	K00800 <i>aroA</i> , <i>EPSP</i> : 3-phosphoshikimate 1-carboxyvinyltransferase [EC:2.4.2.3]			Gene identified: NDR
WP_016874491.1	<i>aroC</i>	K01736 <i>aroC</i> , <i>chorismate</i> mutase [EC:4.2.3.5]			Gene identified: NDR
WP_016879697.1	<i>aroH</i>	K06208 <i>aroH</i> : chorismate mutase [EC:5.4.99.5]			Gene identified: NDR
WP_016872461.1	<i>pkxA</i>	K04518 <i>pkxA2</i> : prephenate dehydratase [EC:4.2.1.51]			Gene identified: NDR
WP_016875155.1	<i>tyrA</i>	K15226 <i>tyrA</i> : argeninate dehydrogenase (NADP+) [EC:1.3.1.78]			Gene identified: NDR
WP_016875300.1	<i>trpA</i>	K01695 <i>trpA</i> : tryptophan synthase alpha chain [EC:4.2.1.20]		2.97	UV upregulated
WP_016875301.1	<i>trpB</i>	K01696 <i>trpB</i> : tryptophan synthase beta chain [EC:4.2.1.20]		2.43	UV upregulated
WP_016875299.1	<i>trpC</i>	K01699 <i>trpC</i> : indole-3-glycerol phosphate synthase [EC:4.1.1.48]		1.86	UV upregulated
WP_016875302.1	<i>trpD</i>	K00766 <i>trpD</i> : anthranilate phosphoribosyltransferase [EC:2.4.2.18]		3.17	UV upregulated
WP_016875298.1	<i>trpE</i>	K13503 <i>trpE</i> : anthranilate synthase [EC:4.1.3.27]		2.30	UV upregulated
WP_016878877.1	<i>trpA</i>	K01695 <i>trpA</i> : tryptophan synthase alpha chain [EC:4.2.1.20]			Gene identified: NDR
WP_016874755.1	<i>trpB</i>	K01696 <i>trpB</i> : tryptophan synthase beta chain [EC:4.2.1.20]			Gene identified: NDR
WP_016876884.1	<i>trpC</i>	K01699 <i>trpC</i> : indole-3-glycerol phosphate synthase [EC:4.1.1.48]			Gene identified: NDR
WP_016875334.1	<i>trpD</i>	K00766 <i>trpD</i> : anthranilate phosphoribosyltransferase [EC:2.4.2.18]			Gene identified: NDR
WP_016877350.1	<i>trpE</i>	K00766 <i>trpE</i> : anthranilate phosphoribosyltransferase family protein [EC:2.4.2.18]			Gene identified: NDR
WP_016876059.1	<i>trpF</i>	K01857 <i>trpF</i> : phosphoribosylanthranilate isomerase [EC:5.3.1.24]		-1.83	UV downregulated
WP_016873541.1	<i>trpG</i>	K01658 <i>trpG</i> : anthranilate synthase component II [EC:4.1.3.27]		1.31	Gene identified: NDR small UV upregulated
<b>AAA metabolism</b>					
WP_016873881.1		NAD(P) <sup>+</sup> /FAD-dependent oxidoreductase; putative L-amino-acid oxidase YobN; K00274 monoamine oxidase MAO [EC:1.4.3.4]		-3.36	FR downregulated
WP_016875303.1		FAD-dependent oxidoreductase; putative L-amino-acid oxidase YobN; K00274 monoamine oxidase MAO [EC:1.4.3.4]			Gene identified: NDR
WP_016873295.1	<i>hmgA</i>	K04511 HGD, <i>hmgA</i> ; homogentisate 1,2-dioxygenase [EC:1.13.11.15]			Gene identified: NDR
WP_016877715.1	<i>fahA</i>	K01555 <i>fahA</i> , <i>fahA</i> ; fumarylacetoacetase [EC:3.7.1.2]			Gene identified: NDR
WP_016873294.1	<i>hpdD</i>	K00457 <i>hpdD</i> , <i>hpdD</i> ; 4-hydroxyphenylpyruvate dioxygenase [EC:1.13.11.27]		-1.52	UV downregulated
WP_02608744.1	<i>aspC</i>	Pyridoxal phosphate-dependent aminotransferase K00813 <i>aspC</i> ; aspartate aminotransferase [EC:2.6.1.1]			Gene identified: NDR
WP_016876265.1	<i>hsc</i>	K00817 <i>hsc</i> : histidinol-phosphate aminotransferase [EC:2.6.1.9]		1.34	small UV upregulated
WP_016878015.1	<i>hsc</i>	K00817 <i>hsc</i> : histidinol-phosphate aminotransferase [EC:2.6.1.9]			Gene identified: NDR
WP_016873290.1	<i>paok</i>	K01912 <i>paok</i> , <i>phenylacetate-CoA</i> ligase [EC:6.2.1.30]			Gene identified: NDR
WP_016876784.1	<i>katG</i>	K03782 <i>katG</i> ; catalase-peroxidase [EC:1.11.1.21]			Gene identified: NDR
WP_016874465.1	<i>tspo</i>	K07185 tryptophan-rich sensory protein homologous to <i>Nostoc</i> TSPO		-1.66	FR downregulated
WP_016876598.1	<i>tspo</i>	TSPO homologous to <i>Fremyella diplospira</i> <i>tspo2</i>		-2.10	UV downregulated
WP_016876599.1	<i>tspo</i>	TSPO homologous to <i>Fremyella diplospira</i> <i>tspo3</i>			Gene identified: NDR
WP_016877233.1	<i>amiE</i>	K01426 <i>amiE</i> , <i>amidase</i> [EC:3.5.1.4]			Gene identified: NDR
WP_016876601.1	<i>amiE</i>	K01426 <i>amiE</i> , <i>amidase</i> [EC:3.5.1.4]			Gene identified: NDR
WP_016873136.1	<i>aldH</i>	K00128 aldehyde dehydrogenase family protein (NADH) [EC:1.2.1.3]			Gene identified: NDR
WP_016878575.1	<i>aldH</i>	K00128 aldehyde dehydrogenase (NAD+) [EC:1.2.1.3]			Gene identified: NDR
WP_016875337.1	<i>aldH</i>	K00128 aldehyde dehydrogenase (NADH) [EC:1.2.1.3]		-3.94	Gene identified: NDR
WP_016873372.1		Hypothetical protein homologous to K00463 <i>INDO</i> ; indoleamine 2,3-dioxygenase [EC:1.13.11.15]			Gene identified: NDR
WP_016873931.1	<i>lpaC</i>	Thiamine pyrophosphate enzyme K04103 <i>lpaC</i> ; indole pyruvate decarboxylase [EC:4.1.1.74]		3.93	UV upregulated
WP_051030344.1		FAD-dependent monooxygenase homologous to K00486 <i>KMO</i> ; kynurenine 3-monooxygenase [EC:1.14.13.9]			Gene identified: NDR
WP_016875162.1		NAD(P) <sup>+</sup> /FAD-dependent oxidoreductase homologous to K11816 <i>INDOLE-3-pyruvate</i> monooxygenase [EC:1.14.13.168]		3.24	UV upregulated
WP_016874602.1		Hypothetical protein homologous to K00453 <i>TD02</i> ; tryptophan 2,3-dioxygenase [EC:1.13.11.11]		-2.75	UV downregulated
<b>Nitrogen transport</b>					
WP_016877347.1	<i>nirA</i>	K00366 <i>nirA</i> ; ferredoxin-nitrite reductase [EC:1.7.1.1]		1.53	FR upregulated
WP_016877945.1	<i>nirB</i>	K15576 <i>nirB</i> , <i>nasE</i> , <i>cytB</i> ; nitrate/nitrite transport system permease protein		1.62	UV/FR upregulated
WP_016877345.1	<i>nirB</i>	K15576 <i>nirB</i> , <i>nasE</i> , <i>cytB</i> ; nitrate/nitrite transport system permease protein		1.81	UV/FR upregulated
WP_016877344.1	<i>nirD</i>	K15578 <i>nirD</i> , <i>nasD</i> ; nitrate/nitrite transport system ATP-binding protein [EC:3.6.3.-]		2.07	FR upregulated
WP_016877343.1	<i>nirD</i>	K15579 <i>nirD</i> , <i>cytD</i> ; nitrate/nitrite transport system ATP-binding protein		2.01	FR upregulated
WP_016877342.1	<i>nirP</i>	K02375 <i>nirP</i> , <i>nasK</i> , <i>nirP</i> , <i>nasK</i> , <i>MFS</i> transporter, <i>NRP</i> family, nitrate/nitrite transporter		2.31	FR upregulated

**Table 1.** Transcriptional regulation of gene homologs associated with MAAs, AAAs and nitrogen transport in *C. fritschii* PCC 6912 exposed to UV and far-red light. Log2 fold – 1.3 > log2 fold change > 1.3 and *p* adj < 0.05. Pink background UV and FR upregulated, purple UV-B upregulated, red FR upregulated, turquoise UV-B downregulated and green FR downregulated.



**Figure 2.** Gene cluster from *C. fritschii* PCC 6912. Cluster analogous to that found in *Anabaena variabilis* encoding the four-enzyme pathway for MAA biosynthesis. The position on NZ\_AJLN01000152 is shown.

Most strikingly, from across the whole genome was that the MAA *mys* gene homolog cluster (Fig. 2) stood out as being significantly upregulated not only in *C. fritschii* PCC 6912 exposed to UV, but also when exposed to FR (Fig. 1). Other notable features included the gene homologs encoding nitrate ABC transporters, required to transport nitrogen into the cell, which were found to be significantly upregulated under both light conditions, and genes associated with the shikimate pathway to produce tryptophan which were predominately upregulated under UV-B but not affected by FR. Below we report these findings in more detail.

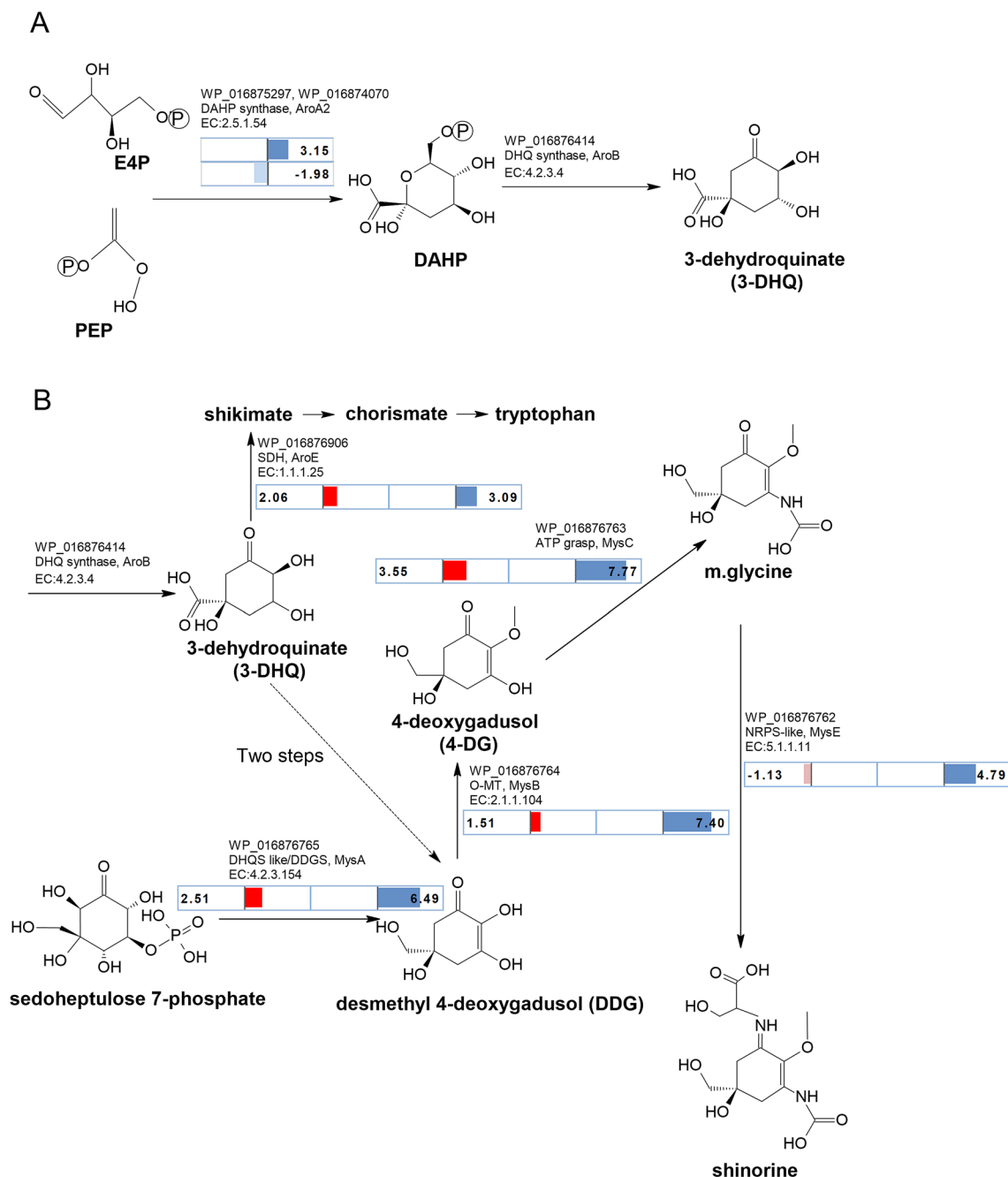
**Pathway to 3-dehydroquinate.** 3-DHQ is the branch point to producing both MAAs and AAAs. 3-DHQ is formed from phosphoenolpyruvate (PEP) and D-erythrose 4-phosphate (E4P), which are intermediates in glycolysis and the pentose phosphate pathway respectively (Fig. 3A). The first step is the production of 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) using DAHP synthase (AroA; E.C 2.5.1.54; Fig. 3A). In the second step, DAHP is converted to 3-DHQ using 3-dehydroquinase (DHQS)<sup>19,20,23</sup>. We found three annotated DAHP synthase gene homologs in the *C. fritschii* PCC 6912 genome (Table 1). One DAHP synthase gene homolog was found within the chorismate-tryptophan cluster and was upregulated under UV light (WP\_016875297: *aroA2.1*; UV log<sub>2</sub> fold change of 3.15). Of the other two gene homologs, expression was downregulated under UV light for one (WP\_016874070: UV log<sub>2</sub> fold – 1.98), and unchanged for the other (Fig. 3A, Table 1). Expression of the single isolated DHQS gene was not affected by either lighting regimen (WP\_016876414; EC 4.2.3.4; *aroB*).

**MAA synthesis.** We found that the gene homologs representing the *mys* cluster (Fig. 2) encoding DDGS (*MysA*), O-methyltransferase (*MysB*) and the ATP grasp (*MysC*) and NRPS (*MysE*) were all upregulated in the UV-B experiment (WP\_016876765 to WP\_016876762; log<sub>2</sub> fold change of between 4.79 and 7.7) (Fig. 3B; Table 1). However, most remarkably, we also found upregulation of *mys A, B* and *C* gene homolog cluster in the FR experiment. Here, however the upregulation was not as high as observed in our UV-B experiment (log<sub>2</sub> fold change of between 1.51 and 3.55) (Fig. 3B). In contrast to the pattern seen under UV-B, under FR the fourth gene homolog in the MAA cluster, *mysE*, which catalyses the conversion of mycosporine-glycine to shinorine was slightly downregulated (WP\_016876762; log<sub>2</sub> fold change – 1.13) although this was not deemed as significant (Fig. 3B). The induction of MAAs under UV light has previously been established in *Gloeocapsa* sp and in *C. fritschii* PCC 6912<sup>14,33</sup>, however this is the first time that it has been observed in cells exposed to FR light.

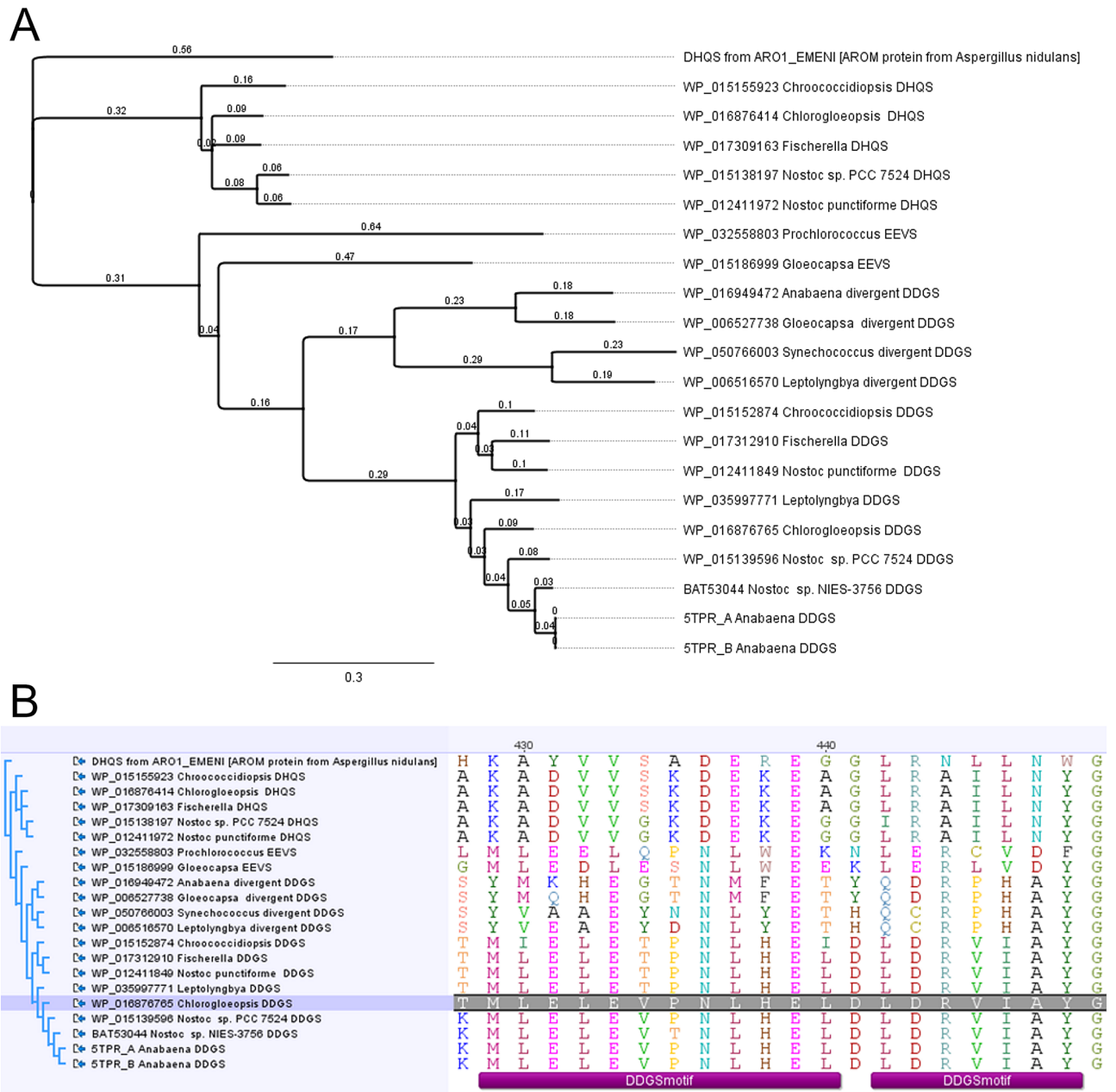
The DHQS-like gene homolog predicted as DDGS associated with the MAA cluster was upregulated under both FR and UV (WP\_016876765; UV/FR log<sub>2</sub> fold change 2.51 and 6.49 respectively; Fig. 3B). This DHQS-like (DDGS) gene homolog has been designated as *mysA*<sup>17</sup>. Two copies of DHQS have also been reported for *Anabaena variabilis* PCC 7937 and *Anabaena* sp. PCC 7120 where they were considered to be catalysing different reactions including one for MAA synthesis and the other for AAA synthesis<sup>34</sup>. We now know that one of these DHQS copies was most likely a DDGS<sup>18</sup>. Annotation of these sugar phosphate cyclase family proteins was confounded by DHQS sharing significant sequence similarity to 2-epi-5-epi-valionone synthase (EEVS) and making current database annotation unreliable<sup>35</sup>. To confirm the identity of the sugar phosphate cyclase family proteins in *C. fritschii* PCC 6912, we used BLAST with representative cyanobacterial proteins, alignments and examination of the published discriminatory motifs<sup>36</sup>. From this we confirmed WP\_016876765 as DDGS and WP\_016876414a as DHQS (Fig. 4). There were no further sugar phosphate cyclase family proteins present.

High performance liquid chromatography (HPLC) analysis of the FR experimental biomass confirmed increases in MAAs under FR with a significant increase in mycosporine-glycine (Table 2; Supplementary Fig. S1). An increase was also observed for shinorine but because of variability across the three experimental replicate samples, this was not deemed to be significant. HPLC analysis of the UV-B samples revealed levels under the white light control to be below detection limit with low levels of shinorine and mycosporine-glycine induced in the samples corresponding to the transcriptome experiment (Table 2). We also observed that in additional samples collected at 4 h in the UV experiment, MAA levels were much higher at this time point, indicating the dynamic nature in acclimation of cells to the stress environment.

**AAA synthesis.** From 3-DHQ, 3-dehydroquinase dehydratase (DHD, EC:4.2.1.10) and then shikimate dehydrogenase (SDH, EC 1.1.1.25) catalyse the two steps to produce shikimate (Fig. 5). In *C. fritschii* a gene homolog (*aroQ*) encoding DHD was downregulated under UV and unaltered under FR (WP\_016876992; UV log<sub>2</sub> fold change of – 1.73). In contrast a gene homolog encoding shikimate dehydrogenase in *C. fritschii* to pro-



**Figure 3.** Regulation of genes on the MAA biosynthetic pathway in *C. fritschii* PCC 6912 under UV-B and FR light. Blue and red/pink represents significant  $\log_2$  fold change for UV-B light exposure and FR light exposure respectively ( $-1.3 > \log_2 \text{fold} > 1.3$  and  $p \text{ adj} < 0.05$ ). Right and left of mid-line denotes upregulation and downregulation respectively. (A) Early common biosynthetic shikimate pathway to both MAAs and AAAs. Pathway from E4P and PEP to 3-dehydroquininate showing regulation under UV. E4P = D-erythrose 4-phosphate and PEP = phosphoenolpyruvate; both intermediates in glycolysis and the pentose phosphate pathway. DAHP = 3-deoxy-D-arabino-heptulosonate 7-phosphate. 3-DHQ = 3-dehydroquininate. (B) Pathway to MAAs via DDGS with change in regulation of the four *mys* gene cluster (WP\_016876762—WP\_016876765) under FR and UV. DHQS: 3-dehydroquininate synthase, DDGS: demethyl 4-deoxygadusol synthase, *mysA*; O-MT: caffeoyl-CoA O-methyltransferase, O-methyltransferase, *mysB*; ATP gras, predicted ATP-dependent carboxylase. ATP-grasp ligase superfamily forming mycosporine-glycine, *mysC* (has no EC number); NRPS non-ribosomal peptide synthase—Amino acid adenylation domain, *mysE*. The branch point to AAAs from 3-DHQ is also shown.



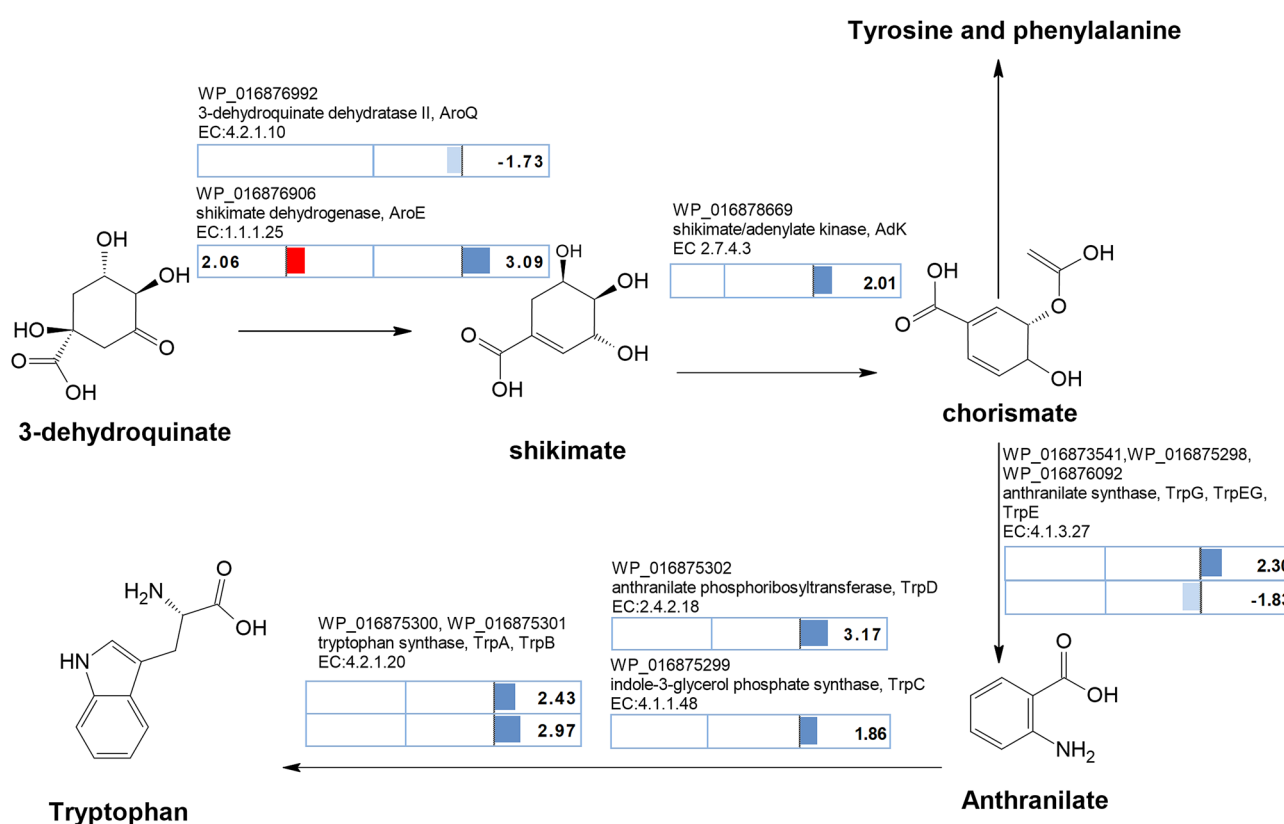
**Figure 4.** Analysis of sugar phosphate cyclase DHQS and DGGs sequences in *C. fritschii* PCC 6912. Comparative analysis with representative cyanobacterial proteins, including examination of the published discriminatory motifs<sup>36</sup>. (A) Neighbour Joining tree (using the Jukes Cantor a bootstrap of 1000) drawn in Geneious version 8.1.9 (<https://www.geneious.com>), from a MUSCLE alignment of cyanobacterial DDGS, EEVS and DHQS using the DHQS from the AROM protein of *Aspergillus nidulans* as an outgroup. (B) Alignment showing the DDGS discriminatory motif. Two proteins were identified: a DGGs [WP\_016876765.1] (highlighted) and a DHQS [WP\_016876414.1].

duce shikimate was significantly upregulated under both UV and FR (WP\_016876906; EC.1.1.1.25; *aroE*; UV/FR log<sub>2</sub> fold change of 3.09 and 2.06 respectively; Table 1).

Shikimate is converted to chorismate in a further 3 steps, via shikimate kinase, EPSP synthase and chorismate synthase. We found a bacterial shikimate kinase which was unaltered (WP\_016875893.1: *AroK*). We also found an adenylate kinase family shikimate kinase, similar to those found in *Nostoc* and *Anabaena*, which was transcriptionally upregulated under UV (WP\_016878669.1: *AdK*; UV log<sub>2</sub> fold change 2.01 (Fig. 5, Table 1). There were no changes in transcriptional regulation under UV or FR for chorismate synthase or for 5-enolpyruvylshikimate-3-phosphate (EPSP) synthase. Chorismate is the branch point to the AAAs tryptophan, tyrosine and phenylalanine, as well as folate (B9) and to phyloquinone (vitamin K) involving multiple steps for each pathway (Supplementary Fig. S2). From chorismate, anthranilate synthase produces anthranilate on the pathway to produce tryptophan and indole or chorismate mutase produces prephenate as the branch point to tyrosine

	Shinorine	Std.Dev	<i>p</i>	m-glycine	Std.Dev	<i>p</i>
<b>UV-B experiment</b>						
White light	0.0	0.0		0.0	0.0	
4 h	65.8	17.4	0.02	57.7	7.0	0.004
4h4d	1.4	1.7	0.3	0.6	1.0	0.42
<b>FR experiment</b>						
White light	21.0	0.3		286.0	29.0	
Far-red	42.0	19.0	0.2	423.0	56.0	0.03

**Table 2.** Estimated average concentrations of MAAs ( $\mu\text{g g}^{-1}$  dry wt): averages determined from three experimental replicates corresponding to replicate transcriptome experiment samples where *C. fritschii* PCC 6912 was exposed to either UV or Far Red (FR) light both with a white light control. In addition to the 4h4d transcriptome samples further HPLC samples were collected in the UV experiment at 4 h. HPLC method for UV experiment was according to Carreto et al.<sup>37</sup> and for FR experiment as described in the text. Concentrations estimated from HPLC peak areas and published extinction coefficients<sup>38</sup>. *p* value significance determined using a two tailed T test. For representative chromatograms for FR samples see Supplementary S2.



**Figure 5.** Shikimate pathway from 3-dehydroquinate to the aromatic amino acid tryptophan. Pathway adapted and simplified from Maeda and Dudareva<sup>23</sup> with overlay of log<sub>2</sub> fold change ( $-1.3 > \log_2 \text{fold} > 1.3$  and  $p \text{ adj} < 0.05$ ) in regulation in *C. fritschii* PCC 6912 under both FR (red denotes upregulation, no downregulation for this pathway) and UV (blue denotes upregulation, pale blue downregulation). No significant change in regulation for other key enzymes including 5-enolpyruvylshikimate 3-phosphate synthase and chorismate synthase involved in conversion of shikimate to chorismate. No change in regulation from chorismate to tyrosine and phenylalanine.

and phenylalanine. We found an upregulation of gene homologs under UV encoding enzymes for the conversion of chorismate to tryptophan; these were found in a gene cluster (WP\_016875298 to WP\_016875302; *tyrA*, *B*, *C* and *D*; UV log<sub>2</sub> fold change 1.86–3.17) (Fig. 5). There was no change in expression in this gene cluster under FR (Fig. 5). It was within this cluster that we found DAHP which was, as indicated above, upregulated only under UV (WP\_016875297; EC:2.5.1.54, *aroA*.2.1; UV log<sub>2</sub> fold change 3.15) (Fig. 3).



**AAA metabolism.** For the tyrosine and phenylalanine metabolic pathways, there was little change in regulation. A putative monoamine oxidase potentially involved in tyrosine and tryptophan metabolism was downregulated in FR (WP\_016873881.1; EC: 1.4.3.4; FR log<sub>2</sub> fold change – 3.36). Additionally, 4-hydroxyphenylpyruvate dioxygenase in the tyrosine and phenylalanine metabolism pathways was slightly downregulated under UV (WP\_016873294.1, EC:1.13.11.27; *hppD*, UV log<sub>2</sub> fold change – 1.52) (Supplementary Figs. S2, S3). A reduction in tyrosine and phenylalanine in *C. fritschii* PCC 6912 exposed to UV-B has previously been observed<sup>39</sup>.

Next we investigated the tryptophan pathways. We first looked for a change in regulation associated with any tryptophan-rich sensory translocator protein (TSPOs). TSPOs are known to be associated with stress adaptation and have recently been found to be upregulated by green light and in response to nutrient deficiency and stress in the cyanobacteria *Fremyella diplosiphon*<sup>26,40</sup>. We found two similar adjacent TSPOs in *C. fritschii* with significant homology (> 75%) to TSPOs in the cyanobacteria *Fremyella diplosiphon*. Expression was not changed for the gene similar to *Fremyella tspo2*, but the *tspo3* homolog was moderately downregulated under UV-B and not affected by FR (WP\_016876598; UV log<sub>2</sub> fold change – 2.10). A third, less related TSPO homologous to proteins found in *Nostoc* sp. PCC7120 and *Anabaena variabilis* was downregulated under FR and unchanged under UV (WP\_016874465; FR log<sub>2</sub> fold change – 1.66). These TSPOs were thus differentially modulated under different light regimens.

Next we investigated if UVR8. Tryptophan has been associated with UVR8 protein serving as the UV-B chromophore that triggers a signalling pathway for UV protection<sup>24</sup>. We found no homolog to UVR8 in *C. fritschii*. Finally, we were interested to determine if there were changes in the tryptophan related indole acetic acid (IAA). There is some indication that the phytohormone related IAA is modulated by tryptophan and light<sup>41</sup>. One of the pathways to IAA is via indolepyruvate. Conversion of indolepyruvate to indoleacetaldehyde by indolepyruvate decarboxylase (IpDC) is the rate-limiting step in this pathway to IAA synthesis. We found indolepyruvate decarboxylase was upregulated under UV-B but not FR (WP\_016877391; E.C 4.1.1.74; IpDC; UV log<sub>2</sub> fold change 3.93) (Table 1; Supplementary Fig. S4). We also found upregulation under UV of a gene homolog involved in the production of indole-3-pyruvate monooxygenase which is also involved in IAA synthesis (WP\_016875162; E.C.1.14.13.168; UV log<sub>2</sub> fold change 3.24). In contrast, two gene homologs involved in tryptophan metabolism and formation of kynurenine and niacin using tryptophan 2,3 dioxygenase were downregulated under UV (WP\_016874813; EC: 1.2.13; UV log<sub>2</sub> fold change – 3.94 and WP\_016874602; EC 1.13.11.11; UV log<sub>2</sub> fold change – 2.75 respectively).

**The role of nitrogen.** Nitrogen, with its major source being nitrate in cyanobacteria, is essential to the biosynthesis of MAAs and AAAs. We found that in addition to the *mys* cluster standing out for upregulation under FR, that the nitrate transport genes were also prominent for their upregulation under FR. Nitrate transporters activate transport of nitrate into the cell for glutamine production<sup>42</sup>. For AAAs, nitrogen is introduced with glutamine converting chorismate to anthranilate (Fig. 5). Glutamine, required for anthranilate production, utilises a source of nitrogen which is primarily sourced through nitrate in the growth medium. Glutamine is required to introduce nitrogen into the aromatic chorismate ring structure converting it using anthranilate synthase into anthranilate on the pathway to tryptophan (Fig. 5). For MAAs, nitrogen is added to 4-deoxygadusol using an ATP grasp (Fig. 2). Tyrosine and phenylalanine require additional conversion before nitrogen is added via aminotransferases<sup>23</sup>. Nitrate transporter genes, *nrtA* and *nrtB* were found within a nitrogen regulation cluster and were upregulated under both UV and FR (WP\_016877346-WP\_016877345; log<sub>2</sub> fold change range 1.59–1.81). Associated *nirA*, *nrtC*, *nrtD* and *nrtP* were upregulated only under FR (WP\_016877347-WP\_016877342 FR log<sub>2</sub> fold change 1.53–2.31) (Table 1). Our results showing upregulation of genes associated with nitrate assimilation indicates that increased nitrogen is required under both UV and FR light conditions.

## Conclusion

MAAs and AAAs are two groups of low molecular weight aromatic compounds containing nitrogen that are related in their biosynthetic pathways. MAAs are widely recognised as being important in photoprotection and AAAs are important in the synthesis of proteins and other essential end products. MAAs in addition to their important physiological role and evolutionary and ecological relevance are increasingly sought after for their use in sunscreen and cosmetic products. Likewise, in addition to their important physiological role, AAAs have a wide application in human health and more widely in human medicine and nutrition.

Our results provide new understanding on the regulation of MAA and AAA biosynthetic pathways under low level UV and FR light. Most importantly, we have shown for the first time that the *mys* gene homolog cluster in *C. fritschii* PCC 6912 is associated with MAA synthesis up to mycosporine-glycine is upregulated when exposed to UV and to FR light. Increases in MAAs measured using HPLC was found in extracts of the corresponding samples of *C. fritschii* PCC 6912 exposed to FR.

Our finding that there is upregulation of the MAA pathway under FR suggests that MAAs may have a role in the photon dissipation of light and thermodynamic optimisation. Recently scytonemin, a related UV-A sunscreen compound, was shown to increase soil temperature in cyanobacterial biocrust communities, decreasing soil albedo significantly with the potential for impacting biosphere feedback and affecting the climate<sup>43</sup>. The ability of FR to influence production of MAAs supports the role of these compounds as photon dissipators opening up new possibilities on the importance of these compounds in heat regulation on Earth.

We also found a cluster of genes associated with the pathway to tryptophan upregulated under UV but not under FR. Pathways to tyrosine and phenylalanine biosynthesis were unaltered under both UV and FR. Nitrate transporters were also found to be upregulated under both UV and FR light with some only being upregulated under FR indicating the requirement for additional nitrogen. Our results highlight the complex finely tuned

interconnectivity between AAAs and MAAs and their response to a changing light environment under both UV and FR.

## Methods

**UV-B light exposure experiment.** Experimental set up for the low level UV-B light exposure was as previously reported<sup>44</sup>. *Chlorogloeopsis fritschii* (Mitra) PCC 6912 was inoculated at 1:50 dilution from a master culture and cultivated in 5 L Erlenmeyer flasks containing 2 L BG11 media with 10 mM HEPES buffer at pH 7.5. The culture was perfused with 1% CO<sub>2</sub> and was maintained at 38 °C using white light (410–750 nm) with an intensity of 60 μmol photons m<sup>-2</sup> s<sup>-1</sup> (Grolux fluorescent tubes). At exponential growth phase, cells were harvested and transferred to nine quartz Erlenmeyer 500 mL flasks containing 200 mL of fresh BG11, and 10 mM HEPES at pH 7.5 to give a concentration of 0.44 g L<sup>-1</sup> wet weight (approximating to 0.04 g L<sup>-1</sup> dry weight). All 9 flasks were exposed to the same white light as for the stock culture for 4 days and 4 h (100 h in total). Three flasks were exposed to white light with no UV-B light acting as the control (white) for 100 h, three flasks were exposed to white light supplemented with UV-B light for the final 4 h of the experiment (samples referred to as; 4 h) and, three flasks were exposed to white light supplemented with 4 h of UV-B light each day for 4 days (samples referred to as; 4h4d). Flasks exposed to UV-B light were placed 10 cm from UV-B tubes (Philips) supplying 3 μmol m<sup>-2</sup> s<sup>-1</sup> at wavelength range 300–310 nm (Supplementary S1).

At the end of the experiment, all 9 flasks were placed on ice and were centrifuged (3000 g) at 4 °C. The pelleted biomass was snap frozen in liquid nitrogen before storage at –80 °C. Transcriptomics to determine differential gene regulation was undertaken on the white light (control) and 4h4d samples. HPLC analysis to determine MAA content was undertaken on the white, 4 h and 4h4d samples.

**RNA preparation and sequencing.** RNA was extracted with Trizol followed by terminator exonuclease digestion to enrich for mRNA and subsequently cleaned using a Qiagen RNeasy column. RNA sequencing was conducted at the Centre for Genomic Research, Institute of Integrative Biology at the University of Liverpool, UK, L69 7ZB, using the Life Technologies SOLiD sequencing platform. For each sample, at least 49,034,856 sequences were obtained (50 bp, min average quality 20 as per manufacturer specifications; per sample average sequence number: 57,516,996.44).

Alignment of reads was carried out using the *C. fritschii* PCC 6912 genome as reference. The sequences obtained for each sample were aligned on to the reference using Bowtie version 0.12.7, using the colour space option. Prior the alignment step, the sequences required the conversion to a pseudo-FASTQ file required as input for Bowtie. For each sequence, only the best alignment was reported by Bowtie, or one was randomly chosen if many were equally best. The average percentage of unambiguously aligned sequences was 47.67%, with a minimum of 35.9% and the maximum equal to 51.9. Considering the known 6968 genes, an average of 83.11% of these were identified as expressed across all the samples (ranging between 73.95% and 87.15%). All the obtained alignment files were processed using HTSeq-count<sup>45</sup>, and reads aligning to the reference genome sequences were counted according to the gene features that they mapped to, as defined in the GTF files.

The differential expression analyses between white and the 4h4d samples were performed using R (version 2.14) and edgeR package. The gene-counts were normalised using “loess smooth” method from the ‘limma’ package. The “GLM” model was applied to the normalised data (with EdgeR package), and the dispersion related to each gene (genewise dispersion) and the pairwise group comparisons were performed to identify differentially expressed genes for each of the three possible group comparisons. For each contrast, each gene with a *p* value below 0.05 (after adjusting for multiple testing effect using the False Discovery Rate approach<sup>46</sup>) were selected as differentially expressed for that contrast. Significant changes in regulation were defined as log<sub>2</sub> fold change ≤ –1.3 or ≥ 1.3 and *p* adj < 0.05.

**Far-red light exposure experiment.** Three conical 1 L flasks containing 800 mL *C. fritschii* culture with an initial optical density (OD) at 750 nm of 0.1–0.2 Absorbance Units (AU) equating to 0.02–0.04 g L<sup>-1</sup> dry weight were cultivated under white LED light with an intensity of 100 μmol photons m<sup>-2</sup> s<sup>-1</sup> for 6 days to an OD at 750 nm of 0.4 AU approximating to 0.08 g L<sup>-1</sup> dry weight. 50 mL aliquots were taken from three flasks and transferred to 50 mL centrifuge tubes (Falcon) to provide the white light control samples. The flasks were then placed under FR LED light with emission centred at 710 nm (Alibaba.com; Supplementary S1), providing ~18 μmol photons m<sup>-2</sup> s<sup>-1</sup>, and grown for a further 24 h. After 24 h, 50 mL aliquots were taken to provide the FR light samples (OD 750 nm of 0.5 approximating to 0.1 g L<sup>-1</sup> dry weight).

**RNA preparation and sequencing.** Samples in centrifuge tubes were immediately cooled on ice then pellets for RNA extraction were prepared by centrifugation for 15 min at 4 °C at 3500 rpm, then further concentrated in a reweighed microfuge tubes at 4 °C at 5000 rpm. After supernatant removal, tubes were weighed on ice to calculate wet weights and the pellets were flash frozen in liquid nitrogen and stored at –80 °C. Pellet wet weights were between 30 and 50 mg. For extraction, pellets were resuspended in 1 mL cold Trizol reagent (Thermo Fisher Scientific) and homogenised using 0.5 mm glass beads (VK05) in a Precellys 24 homogenizer (Bertin) at 6500 rpm for 2 × 20 s with a 10 s break. After 5 min incubation, the sample was extracted with 0.2 mL chloroform followed by centrifugation at 12,000 × g for 15 min at 4 °C. The upper aqueous phase was mixed with an equal volume of 70% ethanol and applied to a PureLink RNA Mini Kit spin cartridge (Thermo Fisher Scientific). The sample was washed and treated on-column with PureLink DNase (Thermo Fisher Scientific) according to the manufacturer’s instructions except for an extra wash step, before drying and elution in 100 μL RNase free H<sub>2</sub>O.

RNA concentrations were determined using a NanoDrop ND-1000 Spectrophotometer (Thermo Scientific). DNA and RNA concentrations were also measured separately using a Qubit 3.0 Fluorometer (Thermo Fisher

Scientific). Residual DNA contamination was removed by two 30 min treatments at 37 °C using 1.5 µL TURBO DNase (Thermo Fisher Scientific) followed by enzyme removal using the inactivation reagent supplied in a TURBO DNA-free Kit (Thermo Fisher Scientific) according to the instructions. To monitor the presence of DNA, primers were designed to SecA, a protein translocase subunit suitable as a reference gene in the heterocystous cyanobacteria *Nostoc* sp. PCC 7120<sup>47</sup> and realtime PCR was performed before and after treatment using the CFX96 Touch Real-Time PCR Detection System (Bio-Rad Laboratories) with PerfeCTa SYBR Green FastMix (Quantabio). This confirmed complete removal of amplifiable DNA from all samples. RNA purity and quality was confirmed by running on an Agilent 2100 Bioanalyzer using a RNA 6000 Nano Kit and by cDNA synthesis using qScript cDNA SuperMix (Quantabio) followed by realtime PCR amplification as above. Prior to analyses, adapter sequences were removed from sample reads using the cutadapt tool.

Total RNA was used as starting material for the generation of sequence ready libraries. Briefly, bacterial ribosomal RNA was removed from the samples with use of a RiboZero bacteria kit. To preserve strand specificity, rRNA free RNA was subjected to TruSeq stranded mRNA sample preparation. Final libraries were normalised to 4 nM prior to pooling. A final library concentration of 20 pM was used to sequence the libraries on a MiSeq platform at the Swansea University Sequencing Facility, generating a total of 99,448,410 high quality reads between all samples using multiple V3 2×75 bp PE sequencing runs<sup>48</sup>.

Due to an incomplete genome assembly and annotation of the *C. fritschii* genome, analysis of gene expression was carried out in two phases. Initially transcripts were de novo assembled using Rockhopper 2<sup>49,50</sup> in order to identify novel transcripts absent from the current genome annotation due to the genome's draft status (gene truncation due to contig boundary interruption of coding sequence causing absence of annotations). The de novo gene list was added to the current annotation. Reads for each sample were mapped to the draft genome using the Subread aligner and read summarization carried out by feature Counts using our improved annotation<sup>51,52</sup>.

Un-normalised Read count summarisation was fed into DeSeq2 R package<sup>53</sup> for differential gene expression analysis measuring the effect of the two conditions, white light and red light, on gene expression levels. Log<sub>2</sub> fold change along with Wald test *p* values and adjusted *p* values were generated from the DeSeq2 normalised dataset. Significant changes in regulation were defined as log<sub>2</sub> fold change ≤ -1.3 or ≥ 1.3 and *p* adj < 0.05.

**MAA extraction and analysis for UV\_B and FR samples.** Each pellet of a known weight was re-suspended in 100% HPLC grade MeOH (1 mL) and left in the dark at 4 °C overnight (24 h). After centrifugation (5 min at 12,000 rpm), the supernatant was removed and evaporated to dryness using a rotary vacuum concentrator. The dried extract was re-dissolved in 600 µL of deionised water and transferred to autosample vials for HPLC analysis<sup>13</sup>.

For the UV experiment, MAA HPLC analysis was according to Carreto et al.<sup>37</sup> using a Thermo Accela Series HPLC system with a binary pump, chilled autosampler (4 °C) and photodiode array detector. Separation was with two columns connected series thermostated at 35 °C (1) a polymeric double-endcapped C18 column (5 µm, 4.6 mm i.d. × 150 mm length; Alltima, Alltech) and (2) a polymer-coated silica reversed-phase C18 column (5 µm, 4.6 mm i.d. × 250 mm length; CapCell Pak UG, Shiseido) protected with a guard column cartridge (4.6 mm i.d. × 20 mm length; Alltima, Alltech).

For the FR experiment, MAA HPLC analysis was performed using an Agilent 1100 system with a binary pump, an autosampler injector and diode array. The stationary phase was an Alltima Altech C18, 4.6 × 150 mm, 5 µm column thermostated at 35 °C. The re-suspended extracts were injected (100 µL) using an auto-sampler. The mobile phases consisted of; Eluent A: Water (0.01% TFA, v/v) and Eluent B: 70% methanol (0.054% TFA, v/v) with a gradient of; 99% A for 10 min, to 80% A over 5 min, to 1% A over 5 min, held for 3 min and increased to 99% A over 2 min. MAAs were monitored at 320 nm with spectral scanning of HPLC separated peaks from 250 to 400 nm. MAAs were identified based on spectral matching of the two main characterised MAAs (shinorine, λ<sub>max</sub> = 334 nm; m-gly, λ<sub>max</sub> = 310 nm) known to be present in *C. fritschii*<sup>14</sup>. Concentrations expressed as µg g<sup>-1</sup> dry weight were estimated from HPLC peak areas and published extinction coefficients using the Beer-Lambert law using ε = extinction coefficient (shinorine, ε = 44,700 M<sup>-1</sup> cm<sup>-1</sup>; m-gly, ε = 28,790 M<sup>-1</sup> cm<sup>-1</sup>)<sup>38</sup>, *p* value significance was determined using a two-tailed t-test.

## Data availability

The datasets generated for this study can be found within an NCBI BioProject with Accession Number PRJNA545395.

Received: 18 May 2020; Accepted: 26 October 2020

Published online: 26 November 2020

## References

- Singh, S. P., Häder, D.-P. & Sinha, R. P. Cyanobacteria and ultraviolet radiation (UVR) stress: mitigation strategies. *Ageing Res. Rev.* **9**, 79–90 (2010).
- Montgomery, B. L. Seeing new light: recent insights into the occurrence and regulation of chromatic acclimation in cyanobacteria. *Curr. Opin. Plant Biol.* **37**, 18–23 (2017).
- Huang, L., McCluskey, M. P., Ni, H. & LaRossa, R. A. Global gene expression profiles of the cyanobacterium *Synechocystis* sp. strain PCC 6803 in response to irradiation with UV-B and white light. *J. Bacteriol.* **184**, 6845–6858 (2002).
- Hideg, É., Jansen, M. K. & Strid, Å. UV-B exposure, ROS, and stress: inseparable companions or loosely linked associates? *Trends Plant Sci.* **18**, 107–115 (2013).
- Shick, J. M. & Dunlap, W. C. Mycosporine-like amino acids and related gadusols: biosynthesis, accumulation, and UV-protective functions in aquatic organisms. *Annu. Rev. Physiol.* **64**, 223–262 (2002).
- Oren, A. & Gunde-Cimerman, N. Mycosporines and mycosporine-like amino acids: UV protectants or multipurpose secondary metabolites? *Fems. Microbiol. Lett.* **269**, 1–10 (2007).

7. Llewellyn, C. A. & Airs, R. L. Distribution and abundance of MAAs in 33 species of microalgae across 13 classes. *Marine Drugs* **8**, 1273–1291 (2010).
8. Garcia-Pichel, F. & Castenholz, R. W. Occurrence of UV-absorbing, Mycosporine-like compounds among cyanobacterial isolates and estimates of their screening capacity. *Appl. Environ. Microbiol.* **59**, 163–169 (1993).
9. Schmid, D., Cornelia, S. & Fred, Z. UV-A sunscreen from red algae for protection against premature skin aging. *Cosmetics* **2004**, 139–143 (2004).
10. Liddell, P. A. *et al.* Mimicry of carotenoid function in photosynthesis: synthesis and photophysical properties of a carotenopyropheophorbide. *Photochem. Photobiol.* **36**, 641–645 (1982).
11. Derikvand, P., Llewellyn, C. A. & Purton, S. Cyanobacterial metabolites as a source of sunscreens and moisturizers: a comparison with current synthetic compounds. *Eur. J. Phycol.* **52**, 43–56 (2016).
12. Rastogi, R. P. *et al.* Ultraviolet radiation and cyanobacteria. *J. Photochem. Photobiol. B* **141**, 154–169 (2014).
13. Rastogi, R. P., Madamwar, D. & Incharoensakdi, A. Sun-screening bioactive compounds mycosporine-like amino acids in naturally occurring cyanobacterial biofilms: role in photoprotection. *J. Appl. Microbiol.* **119**, 753–762 (2015).
14. Portwich, A. & Garcia-Pichel, F. Ultraviolet and osmotic stresses induce and regulate the synthesis of mycosporines in the cyanobacterium *Chlorogloeopsis* PCC 6912. *Arch. Microbiol.* **172**, 187–192 (1999).
15. Waditee-Sirisattha, R., Kageyama, H., Fukaya, M., Rai, V. & Takabe, T. Nitrate and amino acid availability affects glycine betaine and mycosporine-2-glycine in response to changes of salinity in a halotolerant cyanobacterium *Aphanothece halophytica*. *Fems Microbiol. Lett.* **362**, fnv198 (2015).
16. Conde, F. R., Churio, M. S. & Previtali, C. M. The deactivation pathways of the excited-states of the mycosporine-like amino acids shinorine and porphyrin-334 in aqueous solution. *Photochem. Photobiol. Sci.* **3**, 960–967 (2004).
17. Gao, Q. & Garcia-Pichel, F. Microbial ultraviolet sunscreens. *Nat. Rev. Microbiol.* **9**, 791–802 (2011).
18. Balskus, E. P. & Walsh, C. T. The genetic and molecular basis for sunscreen biosynthesis in cyanobacteria. *Science* **329**, 1653–1656 (2010).
19. Pope, M. A. *et al.* O-Methyltransferase is shared between the pentose phosphate and shikimate pathways and is essential for mycosporine-like amino acid biosynthesis in *Anabaena variabilis* ATCC 29413. *ChemBioChem* **16**, 320–327 (2015).
20. Portwich, A. & Garcia-Pichel, F. Biosynthetic pathway of mycosporines (mycosporine-like amino acids) in the cyanobacterium *Chlorogloeopsis* sp strain PCC 6912. *Phycologia* **42**, 384–392 (2003).
21. Challis, G. L. & Naismith, J. H. Structural aspects of non-ribosomal peptide biosynthesis. *Curr. Opin. Struct. Biol.* **14**, 748–756 (2004).
22. Gao, Q. & Garcia-Pichel, F. An ATP-grasp ligase involved in the last biosynthetic step of the iminomycosporine shinorine in *Nostoc punctiforme* ATCC 29133. *J. Bacteriol.* **193**, 5923–5928 (2011).
23. Maeda, H. & Dudareva, N. The shikimate pathway and aromatic amino acid biosynthesis in plants. *Annu. Rev. Plant Biol.* **63**, 73–105 (2012).
24. Wu, D. *et al.* Structural basis of ultraviolet-B perception by UVR8. *Nature* **484**, 214–219 (2012).
25. Rizzini, L. *et al.* Perception of UV-B by the *Arabidopsis* UVR8 protein. *Science* **332**, 103–106 (2011).
26. Busch, A. W. U. & Montgomery, B. L. Distinct light-, stress-, and nutrient-dependent regulation of multiple tryptophan-rich sensory protein (TSPO) genes in the cyanobacterium *Fremyella diplosiphon*. *Plant Signal Behav.* **12**, e1293221 (2017).
27. Gan, F. *et al.* Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light. *Science* **345**, 1312–1317 (2014).
28. Ho, M. Y. & Bryant, D. A. Global transcriptional profiling of the cyanobacterium *Chlorogloeopsis fritschii* PCC 9212 in far-red light: insights into the regulation of chlorophyll d synthesis. *Front. Microbiol.* **10**, 465 (2019).
29. Airs, R. L. *et al.* Chlorophyll f and chlorophyll d are produced in the cyanobacterium *Chlorogloeopsis fritschii* when cultured under natural light and near-infrared radiation. *FEBS Lett.* **588**, 3770–3777 (2014).
30. Zheng, Z., Gao, S. & Wang, G. Far red light induces the expression of LHCSR to trigger nonphotochemical quenching in the intertidal green macroalgae *Ulva prolifera*. *Algal Res.* **40**, 101512 (2019).
31. Kono, M., Yamori, W., Suzuki, Y. & Terashima, I. Photoprotection of PSI by far-red light against the fluctuating light-induced photoinhibition in *Arabidopsis thaliana* and field-grown plants. *Plant Cell Physiol.* **58**, 35–45 (2017).
32. Ma, K. Two new algae from Indian Soil. *Ann. Bot.* **14**, 457–464 (1950).
33. Garcia-Pichel, F., Wingard, C. E. & Castenholz, R. W. Evidence regarding the UV sunscreen role of a mycosporine-like compound in the cyanobacterium *Gloeocapsa* sp. *Appl. Environ. Microbiol.* **59**, 170–176 (1993).
34. Singh, S. P., Klisch, M., Sinha, R. P. & Hader, D. P. Genome mining of mycosporine-like amino acid (MAA) synthesizing and non-synthesizing cyanobacteria: a bioinformatics study. *Genomics* **95**, 120–128 (2010).
35. Osborn, A. R. *et al.* De novo synthesis of a sunscreen compound in vertebrates. *Elife* **4**, e05919 (2015).
36. Osborn, A. R. *et al.* Evolution and distribution of C7-cyclitol synthases in prokaryotes and eukaryotes. *ACS Chem. Biol.* **12**, 979–988 (2017).
37. Carreto, J. I., Carignan, M. O. & Montoya, N. G. A high-resolution reverse-phase liquid chromatography method for the analysis of mycosporine-like amino acids (MAAs) in marine organisms. *Mar. Biol.* **146**, 237–252 (2005).
38. Bandaranayake, W. M. Mycosporines: are they nature's sunscreens?. *Nat. Prod. Rep.* **15**, 159–172 (1998).
39. Kultschar, B., Dudley, E., Wilson, S. & Llewellyn, C. A. Intracellular and extracellular metabolites from the cyanobacterium *Chlorogloeopsis fritschii*, PCC 6912, during 48 hours of UV-B exposure. *Metabolites* **9**, 74 (2019).
40. Busch, A. W. U., WareJoncas, Z. & Montgomery, B. L. Tryptophan-rich sensory protein/translocator protein (TSPO) from cyanobacterium *Fremyella diplosiphon* binds a broad range of functionally relevant tetrapyrroles. *Biochemistry* **56**, 73–84 (2017).
41. Prasanna, R. *et al.* Rediscovering cyanobacteria as valuable sources of bioactive compounds. *Prikl. Biokhim. Mikrobiol.* **46**, 133–147 (2010).
42. Kobayashi, M., Rodriguez, R., Lara, C. & Omata, T. Involvement of the C-terminal domain of an ATP-binding subunit in the regulation of the ABC-type nitrate/nitrite transporter of the cyanobacterium *Synechococcus* sp. strain PCC 7942. *J. Biol. Chem.* **272**, 27197–27201 (1997).
43. Couradeau, E. *et al.* Bacteria increase arid-land soil surface temperature through the production of sunscreens. *Nat. Commun.* **7**, 10373 (2016).
44. Llewellyn, C. A., Airs, R. L., Farnham, G. & Greig, C. Synthesis, regulation and degradation of carotenoids under low level UV-B radiation in the filamentous cyanobacterium *Chlorogloeopsis fritschii* PCC 6912. *Front. Microbiol.* **11**, 163 (2020).
45. Anders, S., Pyl, P. T. & Huber, W. HTSeq: a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
46. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **57**, 289–300 (1995).
47. Pinto, F., Pacheco, C. C., Ferreira, D., Moradas-Ferreira, P. & Tamagnini, P. Selection of suitable reference genes for RT-qPCR analyses in cyanobacteria. *PLoS ONE* **7**, e34983 (2012).
48. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal* **17**, 10–12 (2011).
49. Tjaden, B. De novo assembly of bacterial transcriptomes from RNA-seq data. *Genome Biol.* **16**, 1 (2015).
50. McClure, R. *et al.* Computational analysis of bacterial RNA-Seq data. *Nucleic Acids Res.* **41**, e140 (2013).
51. Liao, Y., Smyth, G. K. & Shi, W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.* **41**, e108 (2013).

52. Liao, Y., Smyth, G. K. & Shi, W. FeatureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
53. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

### Acknowledgements

The far-red experiment work was supported by a BBSRC-NIBB: Phyconet Proof of Concept Grant. UV experimental work was supported by BBSRC Grant BB/E018998/1, Interreg-NEW—EnAlgae and NERC NBAF Liverpool Sequencing Centre (NBAF527). We thank Dr Dan White for the HPLC analysis of MAAs in the UV experiment.

### Author contributions

C.A.L. conceived ideas, oversaw the work and wrote the manuscript, C.G. did RNA extraction for FR experiment, UV and FR bioinformatic data analysis and contributed to manuscript writing, A.S. cultivated initial stock of *C. fritschii* and undertook FR experiments, provided biomass for RNA extraction for the FR experiment, B.K. undertook MAA HPLC analysis of FR samples, M.H. undertook FR sequencing and bioinformatics, G.F. undertook UV experiments and secured NBAF funding for UV sample sequencing. All authors reviewed the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-77402-6>.

**Correspondence** and requests for materials should be addressed to C.A.L.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020