# CYANIDIALES ECOLOGY: BIODIVERSITY AND TRANSCRIPTOMICS IN YELLOWSTONE NATIONAL PARK

by

Dana Jean Skorupa

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Microbiology

MONTANA STATE UNIVERSITY
Bozeman, Montana

April, 2013

APPROVAL

of a dissertation submitted by

Dana Jean Skorupa

This dissertation has been read by each member of the dissertation committee and has been found to be satisfactory regarding content, English usage, format, citation, bibliographic style, and consistency and is ready for submission to The Graduate School.

Dr. Timothy McDermott

Approved for the Department of Microbiology

Dr. Mark Jutila

Approved for The Graduate School

Dr. Ronald W. Larsen

# STATEMENT OF PERMISSION TO USE

In presenting this dissertation in partial fulfillment of the requirements for a doctoral degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library. I further agree that copying of this dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for extensive copying or reproduction of this dissertation should be referred to ProQuest Information and Learning, 300 North Zeeb Road, Ann Arbor, Michigan 48106, to whom I have granted "the exclusive right to reproduce and distribute my dissertation in and from microform along with the non-exclusive right to reproduce and distribute my abstract in any format in whole or in part."

Dana Jean Skorupa

April, 2013

# ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Dr. Timothy McDermott for all the support, knowledge, and guidance he has provided me throughout my graduate program. I am grateful for the abounding level of time and effort he has contributed to both my research and my education. Thank you to the additional members of my committee Drs. Richard Castenholz, Matthew Fields, Gill Geesey, and Tim McMahon for all your insights and contributions to my graduate career. Thanks to all members of the McDermott lab, both past and present, for making our lab such an enjoyable place to work.

Most importantly, and to whom this thesis is dedicated to, I would like to thank my parents, Ralph and Linda, who have sacrificed much to better my life, love, and education. They are my greatest teachers, and any achievement I may have is a credit to them. Equally, I would like to offer my deepest gratitude to Dave for always believing in me, supporting me, and for all he has done over the past seven years to make my life so wonderful. Additional thanks to my dog Rumi, may I someday be as good a person as you already think I am.

Thanks to the National Science Foundation, the Thermal Biology Institute, and Montana Institute on Ecosystems for financial support throughout my graduate career.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTINUED)

TABLE OF CONTENTS (CONTINUED)

# LIST OF TABLES

# LIST OF FIGURES

LIST OF FIGURES (CONTINUED)

# ABSTRACT

The work presented in this thesis investigated the ecology of a unicellular, and asexual red algae belonging to the order Cyanidiales which are the only phototrophs occurring in low pH (0.5-3.5) and high temperature (38-56°C) geothermal environments. These algae exhibit a dynamic seasonal biomass fluctuation referred to as 'mat decline,' where viability drastically decreases as seasonal ultra-violet (UV) and visible (VIS) irradiance intensify. Temporal experiments coupling UV irradiance filtering with whole-community transcription profiling revealed significant cyanidial gene expression changes occurring as a result of exposure to UV, and that patterns of response adjusted across low and high irradiance time periods. Separate analyses examined genes responding to either increasing seasonal UV or VIS intensity, or by the combined effects of both irradiance wavelengths (UV and VIS). Results not only corroborated known physiological changes to solar irradiance, but also suggested the strategies employed to deal with excess VIS and UV intensity may be highly integrated. Comparative analysis determined that environmental microarrays are highly sensitive in their detection of transcript diversity, and *in situ* gene expression changes. Proteomic work identified several dominant cyanidial proteins which were also abundantly transcribed, suggesting there is good correspondence between highly abundant proteins and gene transcriptional activity for these algae.

Additionally, a study is presented which examined the biodiversity and distribution of the Cyanidiales in Yellowstone National Park (YNP). Phylogenetic reconstruction using the *rbcL* gene sequence identified two well-supported YNP lineages: *Cyanidioschyzon* and *Galdieria*-A, and showed the separation of taxa based on ecophysiological conditions. *Galdieria*-A phylotypes were found exclusively in soil and endolithic habitats. *Cyanidioschyzon* was the sole phylotype found in aqueous environments, but was also detected in all soil and endolithic habitats investigated. Culturing efforts demonstrated that moisture availability controls cyanidial viability and distribution in soil habitats, while alternative environmental factors influence endolithic populations. In addition, autotrophic and heterotrophic viable cell counts in combination with *rbcL* gene sequencing determined that *Galdieria* dominates the species composition in soil environments, while the endolithic habitat contains both *Cyanidioschyzon* and *Galdieria*.

# CHAPTER 1

## THE SCOPE OF THE THESIS

The goal of this thesis was to explore the ecology of an order of unicellular eukaryotic algae known as the Cyanidiales, which are the only photoautotrophs occurring in thermoacidic environments (pH 0.2-4.0, temperature 40-56$^{o}$C). These algae are a dominant component of microbial communities in the cooler reaches of acidic geothermal springs and are especially noticeable in Yellowstone National Park (YNP). The ecology of these algae, however, is poorly understood. Much remains to be learned about the distribution, diversity, and physiology of these organisms in acidic geothermal environments.

Chapter 2 introduces significant findings in field of Cyanidiales biology, and attempts to incorporate recent research advances helping to improve our understanding of these algae. The primary focus of this chapter is to explore the characteristics of the Cyanidiales, as well as their potential use as a model for understanding both the evolution of photosynthetic eukaryotes and life in extreme environments. Evolution and taxonomic position are presented, followed by a summary of the distribution and diversity of these algae in acidic geothermal locations. Additionally, aspects of how these algae cope with various extreme environmental conditions are introduced and discussed. Perhaps most importantly, cyanidial mat decline, a seasonally defined reduction in cyanidial biomass, is introduced in this chapter as means of exploring how UV-VIS irradiance is contributing to the health and activity of the Cyanidiales in their natural habitat.

The primary objective of the work presented in chapter 3 was to explore in detail the effects of *in situ* ultraviolet (UV) and visible (VIS) irradiance on cyanidial gene expression patterns. Previous monthly measurements of aqueous chemistries, temperature, and UV-VIS irradiance conducted in parallel with estimates of cyanidial population size found that that algal viability drastically decreases as seasonal UV-VIS irradiance increases. To further our understanding of this relationship, temporal experiments coupling UV irradiance manipulations (filtering) with whole-community transcription profiling were performed that showed that seasonal UV irradiance intensity affects *in situ* cyanidial mat populations. Specifically, the data provide evidence that UV exposure induces significant changes in cyanidial gene expression patterns during seasonal periods of the year with relatively low UV dosage, notably causing the significant repression of a large number genes coding for repetitive DNA elements. Careful characterization of UV-exposure response occurring during summer, a period of the year with increased UV-VIS irradiance intensity and a lengthened photoperiod, suggested increased VIS irradiance intensity was masking UV-based gene regulation affects. By comparing cyanidial gene expression during the lowest and highest UV-VIS irradiance periods of year, separate analyses identified genes regulated by UV or VIS separately, or by the collective effects of UV and VIS. VIS-specific effects regulated genes involved in replication, energy, and growth-related metabolisms, while expression changes specific to UV intensity involved the significant repression of genes associated with the photosystems and accessory pigments. Interestingly, apparent combined UV-

VIS response was only evident for genes annotated as repetitive DNA elements, where UV was found to attenuate VIS repression affects.

A second objective of the work presented in chapter 3 was to assess the utility of environmental microarrays in improving our understanding of gene expression patterns in natural microbial communities. A suite of systematic comparisons determined that environmental microarray data sets have a high information yield, capturing a large breath of functional genes expressed in the cyanidial transcriptome. In order to assess the potential connectivity of the Cyanidiales transcriptome to the proteome, limited-proteomics-based experiments were performed to indentify and compare protein occurrences with relative transcript expression levels. One dimensional proteomic analysis in combination with LC-MS/MS identified several dominant cyanidial proteins that were also abundantly transcribed in matching environmental microarrays, suggesting there is good correspondence between highly abundant proteins and gene transcriptional activity for these algae

Chapter 4 is primarily concerned with improving the understanding of Cyanidiales biodiversity and distribution in Yellowstone National Park (YNP). Phylogenetic analysis of YNP *rbcL* gene sequences identified two well-supported YNP lineages: *Galdieria*-A and *Cyanidioschyzon*, showing the separation of phylotypes based on ecophysiological conditions. YNP *Galdieria* environmental clones were localized exclusively to non-aqueous habitats, while *Cyanidioschyzon* clones were detected in all aqueous, soil, and endolithic habitats sampled. To further our understanding of major factors shaping Cyanidiales diversity and abundance in non-aqueous environments,

culture-dependent work was performed that showed that in sampled soil sites, cyanidial abundance and viability increased as gravimetric soil moisture content increased. Viability and abundance rates in sampled endolithic habitats were not found to co-vary with gravimetric moisture content, suggesting other environmental factors influence endolithic cyanidial populations. PCR cloning of the *rbcL* gene sequence from the sampled soil sites implied *Galdieria* was the dominant Cyanidiales species, while the endolithic habitat contained both *Cyanidioschyzon* and *Galdieria*. Finally, cultivation-based work uncovered that the anti-fungal agent, terbinafine, was a fairly reliable tool for discriminating between the Cyanidiales; allowing for the selective culturing of *Galdieria* from environmental samples.

In summary, this dissertation highlights the ability of environmental microarrays to link *in situ* gene expression patterns with measurable environmental variables, thus connecting microbial populations and activities with their environment. In addition, this body of work demonstrates the value of combining culture-independent and dependent work to assess the diversity and distribution of microbial populations in natural habitats.

# CHAPTER 2

# TRANSCRIPTOMICS AND DIVERSITY OF THERMOACIDOPHILIC CYANIDIALES LITERATURE REVIEW

## Introduction to Acidic Geothermal Environments

Acidic geothermal environments formed as a result of sulfur mineral oxidation are common features in the Yellowstone National Park (YNP) (Wyoming, USA) geothermal complex. The acidic conditions (pH<3) found in springs and surrounding non-aqueous habitats form due to the reaction of sub-surface steam with cool surface water and oxygen gas, reacting to create high levels of dissolved sulfur dioxide ($SO_2$), hydrogen sulfide ($H_2S$) and elemental sulfur ($S^0$) (50). Through a series of oxidation or hydrolysis reactions these sulfur products form sulfuric acid ($H_2SO_4$) (50). Extremely acidic environments are also known to increase the solubility of many heavy metals; the type and concentration of which are determined by individual spring geochemistry. There is a tendency, however, for cationic metal solubility to be much higher in low pH environments (36). This solubility originates either from sulfide mineral oxidation or the extreme hot spring acidity speeding up mineral weathering reactions (36). Solubility can subsequently lead to extremely high levels of certain metalloids compounds, which can include: antimony (Sb), arsenic (As), boron (B), and mercury (Hg) (35).

Aqueous acidic springs have defined physical and chemical gradients that establish as thermal water drains down the hot spring outflow channel. Water temperature, commonly observed discharging at > 70$^{o}$C, decreases as it travels across the

surface of the hot spring, forming a defined temperature gradient (34, 50). Along with water temperature, geochemical gradients are created as a result of chemical transformations. Significant changes in water chemistry, including degassing, oxidation, and mineral precipitation reactions form as geothermal water is discharged (50). The environmental conditions created from these gradients allow for the formation of distinct thermophilic microbial communities (42).

Non-aqueous acidic geothermal environments, sometimes referred to as solfataras, exist either adjacent to aqueous hot springs or in areas where the ground surface is significantly heated. Acidity levels in these soil and rock habitats are a product of geochemical reactions identical to those in aqueous settings; the oxidation of sulfides to $H_2SO_4$. Defined physical gradients are known to exist in these habitats, and include specific shifts in temperature and moisture; soil temperature is known to increase as a function of depth, while moisture levels rise as environmental evaporation rates decrease (58, 66, 71). These environments provide sufficient habitats for thermophilic microorganisms seeking protection from intense ultra violet (UV) or visible (VIS) irradiance, competition from other thermophilic microorganisms, or variable moisture conditions.

Due to the intense physical and chemical constraints created in acidic environments few microorganism are known to thrive in these habitats. At $pH < 2.5$ the growth of only a few bacterial, archaeal, and eukaryotic species is supported (29). Few eukaryotic organisms can live under such extreme conditions, yet an order of asexual unicellular red algae known as the Cyanidiales are a dominant component of microbial

communities in acidic geothermal environments. These unique organisms are the only photoautotrophs known to occur in acidic geothermal environments (pH 0.2-4.0) at temperatures above 38$^{o}$C. This review focuses on the ecology, genetic and physiological diversity, and distribution of the Cyanidiales in acidic geothermal environments in Yellowstone National Park.

## Overview on Cyanidiales Biology

### History

In 1933 on the Sunda islands (Indonesia) Geitler formally described a new member to the class Cyanidiophyceae, naming it *Cyanidium caldarium*, and deeming that this microorganism was a newly defined unicellular alga localized to acidic thermal areas (2, 3, 26, 59). Although first discovered in 1839, the taxonomic position of this alga shifted over time, placed among various algal divisions until the introduction of the *Cyanidium caldarium* binomial nomenclature (44, 63). Unfortunately, confusion over species composition continued when multiple Cyanidiales genera were all referred to as *C. caldarium* (8). Taxonomic positions were later elucidated in 1981 by Merola et al. and three genera were defined under the order Cyanidiales: *Cyanidioschyzon*, *Cyanidium*, and *Galdieria* (45). Despite best efforts to clarify and refine taxonomic differences between the three genera, issues persist. Many pure culture isolates originally classified before the Merola et al. (1981) differentiation were all called *C. caldarium*; therefore research conducted with the morphologically identical *G. sulphuraria* strains was usually accredited to *C. caldarium* (62). Database entries of molecular work from these and other miss-classified culture collections continue the cycle of taxonomic disorder (69, 71, 77).

Currently a single species of the genus *Cyanidium* (*C. caldarium*) and *Cyanidioschyzon* (*C. merolae*) exist, while due to variance in cell morphology, the genus *Galdieria* has been delineated into four separate species: *G. daedala*, *G. maxima, G. partita*, and *G. sulphuraria* (69).

Taxonomic Position

Though always found living in acidic geothermal environments as a mixed community, the Cyanidiales genera have clear morphological, physiological, and biochemical differences which help distinguish them (45). *Cyanidioschyzon* cells are pear-shaped (1-3.5 µm diameter) and reproduce via binary fission, whereas *Galdieria* and *Cyanidium* are larger (3-11 µm diameter) spherical shaped cells, that are surrounded by a thick cell wall, and divide through the production of endospores (n= ~4-32) (2, 69). The ability of *Galdieria* to grow on more than 50 different organic carbon substrates further delineates this genus from the obligate photoautotroph *Cyanidium* (5). Other morphological distinctions distinguishing the three genera include: variations in chloroplast and mitochondrion structure and number (*Galdieria*), presence of a vacuole (*Galdieria*), and the absence of dictyosomes (*Cyanidioschyzon*) (2). Other ecophysiological features of *Cyanidioschyzon* and *Cyanidium* highlight their ability to utilize both ammonium and nitrate, while the *Galdieria* genus is restricted exclusively to using ammonium as a nitrogen source (2).

Origins and Evolution

The Cyanidiales are estimated to be more than 1.3 billion years old; molecular clock methods have placed this algal order near the base of the red algal lineage (76, 78).

Their single plastid membrane positions the Cyanidiales and other deeply branching red algae at the root of secondary endosymbiosis, leading to the divergence of the eukaryotic supergroup, Chromalveolata (15, 59). Genomic comparison as well as phylogenetic reconstruction work suggests that the Cyanidiales ancestor was a cell wall-bearing thermoacidophile (5, 76). Phylogenetic relationships inferred from 16S rDNA and plastid gene analysis place the origin of Cyanidiales divergence within the *Galdieria* genus, followed by a mesophilic *Cyanidium* species, and finally diverging to form *Cyanidioschyzon* and the thermophilic *Cyanidium* genera (76). Based on phylogenetic position of the mesophilic *Cyanidium* species, the ancestral Cyanidiales is predicted to be thermoacidophilic in origin (59). With the onset of genome sequencing, a genome comparison between the fully-sequenced *C. merolae* strain 10D genome and a *G. sulphuraria* expressed sequence tag (EST) library confirmed the ancestral nature of *Galdieria* within the Cyanidiales (5). Results suggest that the evolutionary ancestor to *Cyanidioschyzon* was a cell wall bearing organism whose localization over time to aqueous or moist habitats resulted in the loss of genes associated with cell wall and vacuole synthesis (5, 59). Aquatic habitats would also obviate the need for *Cyanidioschyzon* to maintain metabolic pathways associated with heterotrophy.

## Characterization of Cyanidiales Ecology

### Habitats

General Distribution. The Cyanidiales display such strong global distribution in acidic geothermal environments that they are theorized to inhabit any environment with

suitable temperature and pH parameters (8). Optimally grown at 45$^{o}$C, the Cyanidiales are detected in aquatic habitats ranging from 35$^{o}$C to 56$^{o}$C; below 35$^{o}$C competition from other eukaryotic algae is too great for survival, while above 56$^{o}$C components of the photosynthetic machinery begin to break down (Fig 2.1A, and 2.1B)(17). The Cyanidiales are also known to colonize non-aqueous geothermal habitats, including: moist soils, rock surfaces (interlithic), and endolithic growth beneath the rock surface (Fig.2.1C and 2.1D) (27, 77). In non-aqueous environments, these algae can grow at temperatures as low as 10$^{o}$C (18). Along with temperature, acidity serves as another defining niche characteristic, with pH ranging from 0.2 to 4.0 (optimal ~pH 2.0–3.0) (18).

Although uncommon, cyanidial populations have been found in habitats which are not acidic and/or thermal in nature. Cave ecosystems in Chili, Israel, France, and southern Italy have been found to contain endolithic *Cyanidium* populations (11, 23, 59, 65). These Cyanidiales members thrive under low temperature ($\geq$ 10$^{o}$C), circum-neutral pH (pH 5-7), and low light intensity (> 1% of daylight) conditions (4). It is thought that the *Cyanidium* populations are specifically adapted to cave environments (11); they show a preference for low light conditions, as well as the cool microclimate provided by cave habitats. Highly acidic environments associated with mining operations are also suitable microhabitats for mesophilic Cyanidiales (1). Due to the lack of geothermal activity, Cyanidiales isolated from mining area have a lower optimal growth temperatures, as well as higher tolerance towards freezing treatments (31).

Ecophysiological Niches. All major acidic geothermal areas are inhabited by a mixture of Cyanidiales genera, with moisture content theorized to drive species composition (11, 55, 77). Historically *Cyanidioschyzon* is thought to localize to wet


Fig. 2.1 Photographs of Cyanidiales occurring in Yellowstone National Park. (a) Narrow green bands along the perimeter of a geothermal spring, Dragon Spring. (b) Dominant green mats at Lemonade Creek. (c) Endolithic communities growing as a layer under a rock surface (d) Sub-surface soil communities.

habitats (i.e. aqueous springs), while *Galdieria* and *Cyanidium* are primarily found in non-aqueous settings. Recent molecular studies further define Cyanidiales habitat preference by showing the separation of cyanidial populations based on prevailing environmental conditions. Cultivation-independent surveys using *rbc*L clone sequences support the preference of two *Galdieria* clades; *Galdieria*-A and *Galdieria*-B, as well as *Cyanidium* for non-aqueous thermoacidic environments (11, 77). Desiccation tolerance experiments further delineate *Cyanidium* and *Galdieria*; the latter is much more resistant to dehydration treatment (55), thus restricting *Cyanidium* populations to more moist non-aqueous environments. Although no morphological variation currently exists to separate

the *Galdieria* clades, a *Galdieria*-B isolate has shown preference for lower temperature and light conditions than its *Galdieria*-A counterpart (55). This data, along with the fact that *Galdieria*-B members have only been localized to low humidity environments, suggests this taxa is adapted to shady and dry habitats (11, 55).

Biodiversity

Small highly acidic geothermal habitats are globally scattered, usually in and around volcanic areas. Due to reduced desiccation tolerance, the Cyanidiales are not thought to undergo globally long-distance migration (28, 66). This constraint means cyanidial populations are sometimes isolated geographically, potentially undergoing significant genetic divergence when faced with environmental selection pressures (30). Evidence of isolated thermophilic microbial populations formed as a result of geographical separation has been shown before within the Yellowstone geothermal complex by members of the bacterial genus *Sulfurihydrogenibium* (68). Distribution patterns were found to correlate with the Yellowstone caldera boundary, indicating that geographic separation contributed significantly to detected sequence diversity levels (68). Evidence also suggests that following geographical separation, ecophysiological forces present in a given environment can subsequently result in local adaptations (73). Indeed adaptation to specific localized environments has been shown in cyanidial populations. Environmental polymerase chain reaction (PCR) surveys for the *rbc*L gene in Phlegrean Fields (Pisciarelli, Italy) showed Cyanidiales clades phylogenetically separating in major clades according to habitat preference, and then diverging within each clade by geographical origin (11).

## Biogeography

Dispersal Mechanisms. Brock (1978) first noted that the infrequent occurrence low pH environments should make widespread dispersal of the Cyanidiales difficult, and that yet these algae are detected in nature almost anytime suitable habitat conditions are provided. Even dispersal across individual acidic hot springs within the same geothermal region is thought to be made difficult by non-optimal ecophysiological conditions (59). While subsurface hydrothermal connections are a possibility for proximate hot springs, thermal waters within a single geothermal complex are most likely not universally linked, or if linked, most likely contain geothermal waters far too hot for microbes to survive while dispersing (6). Along with physical isolation, the Cyanidiales intolerance towards desiccation should severely affect transportation across large geographic areas (31, 55). Obstructions aside, these seemingly large environmental barriers do not appear to prevent algal transport.

Possible methods for localized transport include attachment to traveling animals and/or insects. In YNP, snow-free geothermal basins are prime feeding grounds for animals during winter. Attachment of thermophiles to the fur or feet of these animals and subsequent transport to different geothermal areas is one explanation. Various fly species have also been found to feed on hot spring phototrophic mats; insects could therefore also serve as transport vectors over short distances (74). The most plausible explanation for regional transport appears, however, to be in airborne dispersal. Bonheyo et al. (2005) have detected thermophilic bacterial 16s rDNA sequences in the steam emanating from hot springs, demonstrating the aerosolization and movement of thermophilic bacteria.

Research thus far does not, however, address how airborne thermophiles can maintain viability when traveling over large distances.

Brock (1967) provides an alternative explanation to dispersal mechanisms, suggesting that large geothermal complexes were most likely widespread on early Earth, and that during this time thermophilic microorganisms only had to dissipate over short distances because of the close proximity of major geothermal areas. This theory however, does not explain how the Cyanidiales came to colonize geothermal areas on isolated islands (i.e. Iceland) which have never contacted major continents (30, 59). It is more likely that major volcanic events resulting in the discharge of large quantities of volcanic gas are what allow for the long-distance travel of cyanidial cells (59). It has been noted that dissipation of any microbial population into a new habitat is usually unsuccessful and occurs only under ideal conditions, but that given enough evolutionary time, transport will eventually succeed (22).

Site Colonization. The ability of different environments to support cyanidial colonization is influenced by numerous ecophysiological factors. Physical factors such as temperature, pH, and chemical energy gradients are all known to control the distribution of aqueous mat populations. Studies have looked at the restoration of thermophilic microbial populations following severe physical disruption of a geothermal outflow channel, and noted significant re-colonization of thermophilic microbes in as little as two days, cyanidial colonization rates are expected to be on par with these findings (41, 59). Much less is known about the establishment of cyanidial populations in soil and endolithic environments. Laboratory experiments using *Galdieria* showed that when

cultivated on geothermal rocks supplemented with silica, the algae began colonizing the rock surface after only a few days, and that over time the silica mineral encrusted these algae within rock layers (27). Continual thickening of the silica layer in natural habitats is thought to significantly reduce ambient light levels (27); however, heterotrophic *Galdieria* populations would remain unaffected, and could continue to persist.

Tolerances

Proton Concentration. To maintain integrity of cellular components, organisms living in acidic environments are adapted to handle large proton ($H^+$) concentrations, developing complex mechanisms for sustaining a near-neutral intracellular pH. Through the use of $H^+$ efflux pumps, the Cyanidiales are able to survive in the presence of a steep proton gradient approaching nearly 1:1 million (29, 57). Using a *C. caldarium* pure culture, Enami et al. (1986) showed that when cultivated in the dark and inhibited for respiration the external solution surrounding the *C. caldarium* cells rapidly increased in pH, indicating the presence of passive proton leakage into the cytosol when there is no energy supply available, and highlighting the importance of $H^+$ efflux pumps. Interestingly, light-activated $H^+$ export has also been shown in *C. caldarium*, and is thought to function by activation of the plasma membrane ATPase (38). Combining these pieces of evidence, it appears cyanidial cells maintain a neutral internal pH through light or oxygen induced ATPase proton efflux. Additional characterization of the *C. caldarium* and *C. merolae* $H^+$-ATPase revealed that unlike higher plants, the C-terminal region of the $H^+$-ATPase in these algae lacks an important auto-inhibitory domain, which would normally function to regulate the $H^+$-ATPase activity (20).

Carbon Dioxide Availability. Acidic environments are limited with regards to the level of dissolved inorganic carbon (DIC) biologically available to primary producers. In low pH habitats (pH<4) the DIC reaction equilibrium severely reduces bioavailable bicarbonate ($HCO_3^-$) pools (29), resulting in a limited amount of $CO_2$ for photosynthesis. Loss of $HCO_3^-$ is compensated for in low pH systems by the high rate of $CO_2$ uptake from the atmosphere; however, this does not aid aqueous microbial populations due to the fact that $CO_2$ diffusion rates in water remain extremely slow (29). This, coupled to the fact that many phototrophs encoded for a ribulose bis-phosphate carboxylase-oxygenase (rubisco) carboxylation enzyme with a low $CO_2$ affinity, mean primary production in aqueous acidic environments is still a difficult task (10, 29). Some photosynthetic organism have compensated for these limitations by developing a $CO_2$-Concentrating Mechanism (CCM) to stockpile $CO_2$ near the rubisco enzyme (16). Interestingly, work has shown that a CCM is actively used in *C. merolae* but is not in *C. caldarium* (16, 57, 80). This discrepancy can be explained by the fact that the *C. caldarium* rubisco enzyme has a high affinity for $CO_2$, making a concentration mechanism unnecessary (70). Contrastingly, in *C. merolae*, the rubisco enzyme has a low carboxylation efficiency, indicating a CCM is necessary in order to promote sufficient autotrophic metabolism (80). There appears to be an evolutionary trade-off between the presence of a CCM and rubisco specificity levels; high rubisco carboxylation rates obviate the need for concentrating $CO_2$ intracellularly (70). The localization of *Cyanidioschyzon* to aqueous habitats (where $CO_2$ diffusion rates are low) along with its inability to grow heterotrophically help explain the reason for utilizing a CCM.

Heavy Metals. Due to either the oxidation of sulfide minerals, or accelerated mineral weathering at low pH, acidic environments tend to have increased concentrations of heavy metals such as: As, Sb, B, and Hg (54, 67). The maintenance of a low pH helps maintain metal solubility. Toxicity of trace elements varies depending on speciation; anions are typically more toxic (64), as their protonation under acidic conditions allow for easy absorption to the cell surface, while cationic metals have lower absorption rates which prevent them from gaining access to the cell interior. Many acidophilic microorganisms have developed mechanisms of metal resistance, detoxification, or metabolism in order to survive in these environments. The Cyanidiales are primarily thought to utilize resistance methods in order to endure such harsh environmental surroundings, though tolerance levels across genera are known to vary (49). In general, *Galdieria* is thought to be more resistant to toxic concentrations of heavy metals than *Cyanidioschyzon* and *Cyanidium* (59). Qin et al. (2009) recently demonstrated the influence of a *Cyanidioschyzon* sp. on the biogeochemical cycling of Arsenic (As), highlighting its ability to oxidize Arsenite [As (III)], reduce Arsenate [As (V)], and methylate As (III). Meanwhile *G. sulphuraria* has been shown to be involved in the reduction of mercury (Hg) (37); this alga was shown to sequentially convert inorganic Hg (II) to the less toxic elemental Hg (0). Results suggest that along with detoxification, the Cyanidiales are also contributing to the biogeochemical cycling of heavy metals in acidic environments.

Associations with Non-Photosynthetic Thermoacidophiles

Source water temperatures of many acidic geothermal springs are often far above the upper temperature limit for the Cyanidiales (60-84$^{o}$C) (34); here acidophilic bacterial and archaeal microorganisms dominate the biota. Primary production in these upper temperatures ranges is dominated by chemolithotrophic prokaryotes (41). Molecular and geochemical analyzes have shown that that alterations to the physical and biochemical parameters of the hot spring outflow channel result in shifts in the thermophilic microbial community (35). Phylogenetic analysis of 16s rDNA from an acid-sulfate-chloride (ASC) hot spring system in YNP suggested *Acidimicrobium*, *Hydrogenobaculum*, *Desulferella*, and *Meiothermus*-like populations are important bacterial members, while *Caldococcus*, *Metallosphaera*, *Stygiolobus*, and *Thermocladium*-like organisms are associated archaeal populations (34). Each of these acidophilic organisms is uniquely adapted to a specific suite of environmental niche conditions, which act to control distribution and activity. As the acidic hot spring reaches moderate temperatures, shifts in the microbial population occur again, and the community becomes dominated by the eukaryotic Cyanidiales. Below temperatures of ~38$^{o}$C competition from other photosynthetic *Chlorella*-like and *Paradoxia* algae, along with the protist *Euglena* is too great for the Cyanidiales to persist (21, 36). Much less is known about bacterial and archaeal communities inhabiting terrestrial acidic environments. Using a cultivation-independent approach, Walker et al. (2005) analyzed an acidic geothermal endolithic community in YNP, demonstrating that *Galdieria* dominates the endolithic community along with a *Mycobacterium* species (71). Numerous other bacterial and archaeal sequence signatures were detected, with

chemolithotrophy thought to be the primary energy production mechanism for these acidophilic microorganisms (71).

## Cyanidiales as a Model Organism

### Cell Biology and Metabolism

Their simplistic structure and unique metabolic capabilities have resulted in the use of several Cyanidiales species as biological models in the areas of physiology and cellular architecture. The existence of high selection pressures in the acidic hot spring environment is thought to aid these algae in retaining very ancient structural attributes; *C. caldarium* and *C. merolae* cells contain only a single nucleus, mitochondrion, and chloroplast (39, 51). Many other primitive eukaryotes have organelles that vary widely in shape and number, making it difficult to study the division and evolution of cellular compartments within the eukaryotic cell (39). Here, the Cyanidiales serve as an important intermediate to elucidate the evolutionary conversion of cyanobacteria from endosymbionts to eukaryotic organelles. The photosynthetic machinery of *C. merolae* is also considered to be relatively simplistic, encoding for a limited number chlorophyll and carotenoid pigments, making it a good model system for exploring the evolution and ancestry of enzymes involved in eukaryotic photosynthesis (13).

### Genome Sequencing

Recent progress in the field of genomics is improving knowledge about these algae at the molecular level. Analysis of the fully-sequenced *C. merolae* strain 10D genome uncovered that this alga has one of the smallest nuclear genomes (16.5 Mbp)

among all photosynthetic eukaryotes, as well as a relatively low complexity level; of the 5,331 nuclear genes encoded for on 20 linear chromosomes, only .5% contained introns (43). This contrasts genomic data from other simplistic eukaryotes where introns were prevalent in ~5-80% of all open reading frames (ORF) (43). Results suggest that all but 26 *C. merolae* genes can be directly translated into protein sequence, making them very similar to eubacterial systems. These simplistic genomic features of are peaking interests in using the *C. merolae* as a model organism to help address questions about the origin, evolution, and essential metabolic processes in photosynthetic eukaryotes. Additional amino acid determination and structural information from *C. merolae* might also function as an important tool in understanding the thermo-stability properties of cyanidial proteins (5).

Additional sequencing and annotation of ~3,000 expressed sequence tags (EST) from a *G. sulphuraria* cDNA library has been used to detect the presence of several unique metabolic enzymes either not seen before in a eukaryote, or that are more evolutionarily related to prokaryotic proteins (72). A comparison of the *G. sulphuraria* EST library to the completed *Cyanidioschyzon* genome found that almost 50% of all tested *G. sulphuraria* genes contained introns and sequence signatures of a spliceosome (5). A direct comparison of *G. sulphuraria* and *C. merolae* sequence data found numerous variations at the molecular-level (5), helping resolve the physiological difference known to exist between the two genera. The lack of solute transporters encoded for in the *C. merolae* genome helps explain its obligate autotrophic metabolism (5). As well, several enzymes functionally annotated to be involved in cell wall

biosynthesis, modification, and degradation were missing in *C. merolae,* explaining the physical absence of a cell wall in this genus (59). Genes putatively associated with adaptation to osmotic stress also varied across the two genomes, helping elucidate the habitat variation observed for these genera *in situ* (59).

*Cyanidioschyzon* Physiology

Taxonomic classification methods have traditionally distinguished the three formally recognized Cyanidiales genera largely based on cell morphology and reproductive characteristics. However, with the onset of molecular analyses, phylogenetic relationships formally defined using traditional classification methods are often times in direct contrast to molecular sequence data collected. Conflicting classification such as this is known to exist within the Cyanidiales order. An 18S rDNA Cyanidiales environmental survey conducted by Ferris et al. (2005) in an acidic geothermal stream, Nymph Creek (YNP, WY), recovered a single species, *C. merolae* (21). Interestingly, when cultivated, the *C. merolae* isolates were morphologically circular in shape, and reproduced through the use of endospores (21). All cultivated strains had morphologies and cell division characteristics similar to those of *Galdieria* and *Cyanidium*. A recent study by Toplin et al. (2008) in YNP confirmed these unique findings; the majority of an extensive YNP park-wide isolate collection (~120 isolates) were made up of a cyanidial subtype termed type IA, which had more than 99% *rbc*L and 18s rDNA sequence identity to the fully-sequenced *C. merolae* strain 10D genome (69). However, type IA isolates (82 isolates) were again morphologically and reproductively similar to *Cyanidium* and *Galdieria* (69). This incongruity is not unprecedented for members of the Cyanidiales; *G.*

*maxima* are morphologically distinct from *C. merolae* as well as capable of heterotrophic growth, yet these taxa show a close phylogenetic relationship (11, 69, 77). Thus it appears that physical characteristics such as morphology are not informative enough to establish a robust taxonomy among these algae, and that through the reliance on traditional classification methods we most likely underestimate the total genetic diversity of these algae in nature (11).

## Population Diversity and Dynamics

### UV and Light Inhibition

Irradiance, which includes both ultraviolet (UV) radiation and visible (VIS) light, is known to be a major stress factor affecting photosynthetic organisms in natural environments. Phototrophs must simultaneously harvest photosynthetically active radiation (PAR) (400-700 nm) for cellular energy, while minimizing photoinhibition effects triggered as a result of exposure to high concentrations of PAR and UV irradiance. With regards to UV, both UV-A (315-400 nm) and UV-B (280-315 nm) are known to detrimentally effect photosynthetic organisms. UV-B is generally considered more damaging due to increased levels of energetic radiation (12). Inhibitory cellular effects from high levels of PAR and UV are thought to induce species-specific cellular responses. An examination of cellular responses induced as a result of high UV irradiance conditions in cyanobacterial populations uncovered the activation of a variety of protective and/or counteractive mechanisms (53), including: avoidance, DNA repair, protein synthesis, and the production of protective chemicals. Within the Cyanidiales, microarray analysis of a *C. merolae* pure culture assessed the effects of VIS intensity on

gene expression response, uncovering the specific activation of a plastid-encoded sigma factor under high light conditions; suggesting a possible VIS-specific adaptive response in these algae (47). Additional analysis of VIS effects on a *G. sulphuraria* pure culture isolate found low photosynthetic productivity under increased VIS intensity, again hinting at a possible VIS-induced toxicity
response (52).

Cyanidial Mat Decline

Several *in situ* studies have examined cyanidial response to UV and visible light irradiance levels at select geothermal features in YNP. At Nymph Creek (YNP, WY), variation in cyanidial photosynthetic rates as a result of ambient UV exposure found primary productivity was unaltered by UV radiation (9). Castenholz et al. (2010) point out, however, that the selection of a heavily shaded site location and sampling of intact mat cores may have disrupted actual cyanidial response to UV irradiance effects. In another Yellowstone study, temporally intensive work conducted at Dragon Spring (Norris Geyser Basin, YNP) was initiated after cyclical color changes were observed to occur in several cyanidial mat communities in YNP (Fig. 2.2) (40). The mats, normally dark green and lush, were found to deteriorate drastically during the summer months. Measurement of potentially important physical factors along with monthly viable cell counts suggested UV irradiance, VIS irradiance, and water temperature were all important physical factors influencing seasonal changes in cyanidial population densities. This ecological event, referred to as ‘mat decline’ was directly linked to environmental

UV irradiance levels, and results suggested UV and VIS irradiance were likely keystone environmental factors controlling the health and activity of the Cyanidiales *in situ* (40).

Physical Environment Factors and Solar Irradiance Effects. Chemical and physical factors measured on a monthly basis showed that several ecophysiological factors were an important environmental selector (40), controlling Cyanidiales viability and contributing to mat decline. Chemical measurements, including pH, showed low levels of variation when examined monthly over the course of a year (40), suggesting chemical conditions were almost static-like. Contrastingly, variations in the most probable number (MPN) viable cell counts were directly linked to seasonal changes in water temperature and UV-VIS irradiance (Fig. 2.3). Water temperature correlated well with viable cell counts through the beginning of the year, but diverged sharply in summer (Fig. 2.3A) (40). In geothermal springs, water temperature is often thought of as chemostat-like, though changes in geothermal activity, air temperature, and intermittent flow levels are known to cause inconsistencies (40). Interestingly, MPN counts deviated from water temperature just as maximal UV and VIS irradiance measurements peaked (Fig. 2.3B) (40). Statistical analysis confirmed the significance of the results; 59% of the seasonal MPN variability could be linked to seasonal variation in water temperature and UV-VIS irradiance (40).

As a means of assessing which individual irradiance component was associated with decreased cyanidial cell viability, photosynthetic activity was measured as a function of irradiance components that were directly manipulated through the use of

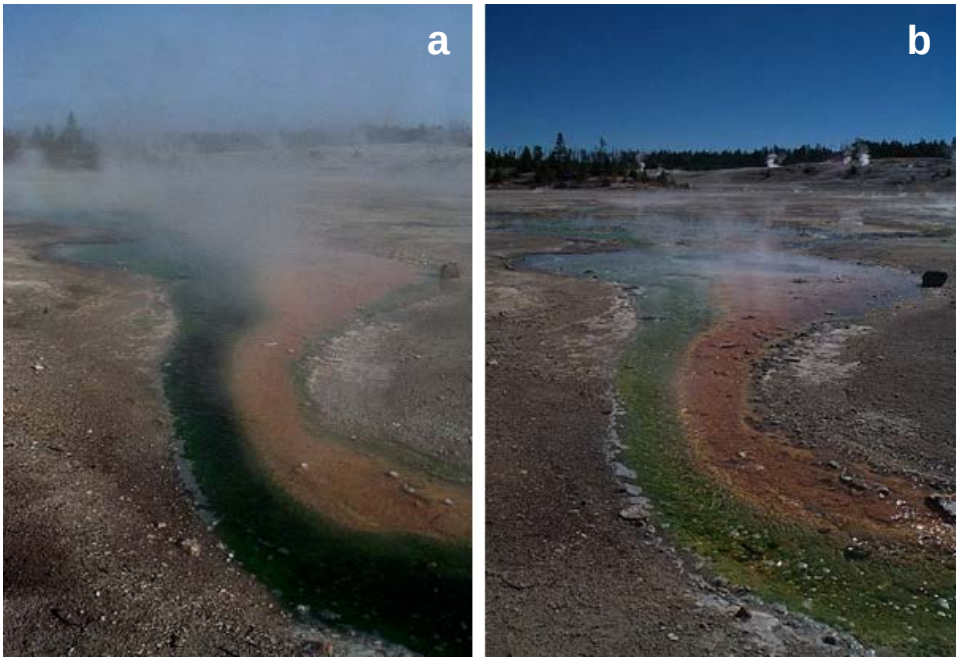


Fig. 2.2 Photographs of Cyanidiales documenting mat decline during summer 2005. Panels are of Pinwheel Geyser, a geothermal spring located near the boardwalk in Norris Geyser Basin, YNP in (a) January and (b) July 2005. Adapted from Lehr et al., 2007.

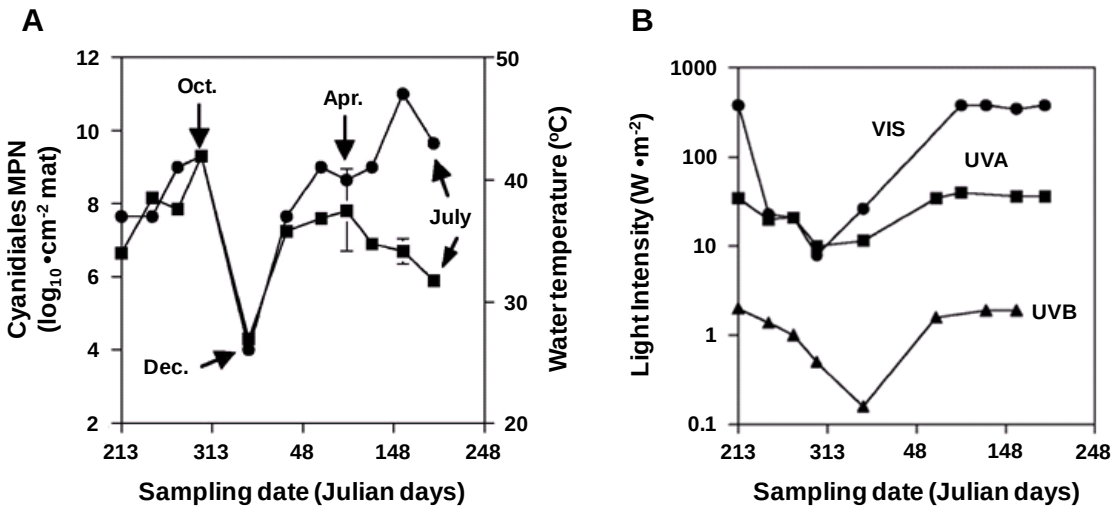


Fig. 2.3 Most probable number (MPN) counts of Dragon Spring Cyanidiales as compared with water temperature and UV–visible (VIS) intensity. (A) MPN counts (■) and temperature (●) as a function of time. Error bars depict the range of the MPN counts for each sampling time. (B) UVA (315–400 nm, ■), UVB (280–315 nm, ▲), and VIS light (400–700 nm, ●). Adapted from Lehr et al., 2007.

selective filters (40). Results provided evidence that VIS and UV irradiance both contribute to the inhibition of cyanidial photosynthesis; significant differences in photosynthetic activity were noted across all filter types in both morning as well as sampling conducted after the algal mats were exposed to several hours of ambient irradiance (Fig. 2.4) (40). Data showed that photosynthetic rates increased by roughly 30-45% when both UVA and UVB were screened out (depending on the time of day) (40).

Population Diversity. A comparison of microsatellite sequences derived from both YNP pure culture isolates and *in situ* cyanidial mat populations tracked algal diversity and dynamics over the course of the seasonal mat decline period. Seasonal *in situ* algal amplicons were matched electrophoretically to pure cultures isolates to assess whether the dramatic differences observed in the viable cell counts could be linked to shifts in cyanidial population abundance and diversity (40). Indeed, profiles showed that mat populations varied seasonal in both their dominance within the community (i.e. band intensity) as well as population diversity (i.e. number of bands) (Fig 2.5) (40). Population-level specialization occurring as a result of select physical and chemical conditions has been shown before in YNP using the thermophilic cyanobacteria *Mastigocladus*. Here, genetically divergent populations formed specialized phenotypes, suggesting that differences genetic diversity were of significant ecological importance (46).

## Transcriptomics

Large-scale analysis of global gene expression patterns is currently used in the

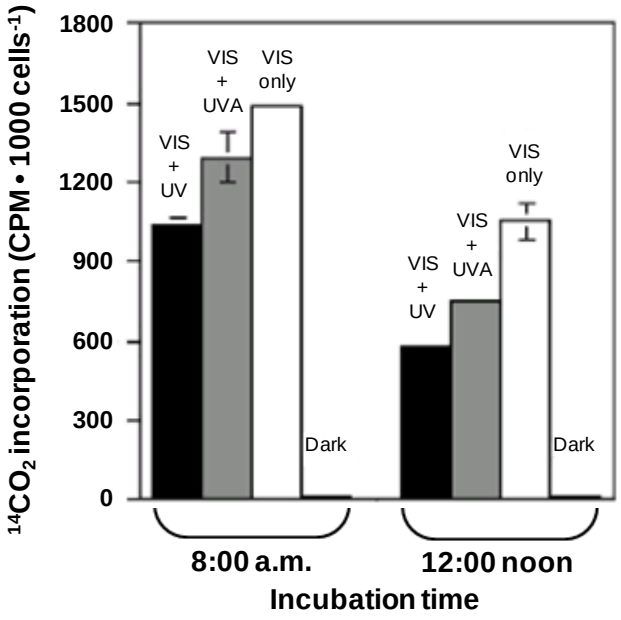


Fig. 2.4 Photosynthetic $^{14}CO_2$ incorporation into Cyanidiales cell suspensions as a function of irradiance exposure and sampling time. Results are the mean and SE from three samples for each treatment. Adapted from Lehr et al., 2007.

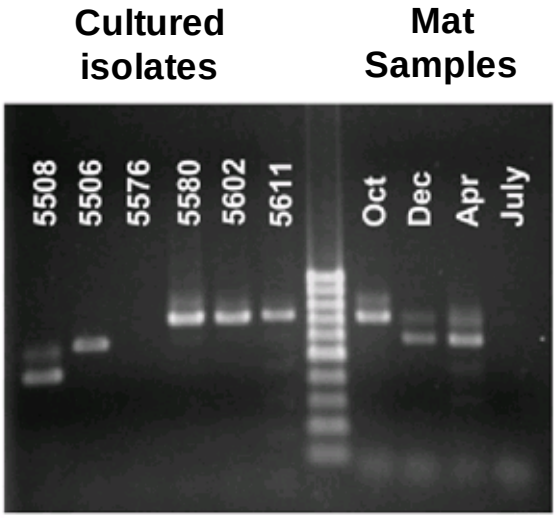


Fig. 2.5 Separation of microsatellite PCR products by electrophoresis. For each microsatellite PCR, 10 µL of PCR product was loaded per lane. Molecular weight standards (100–1000 bp), shown in the middle of the gel, separate pure culture isolate microsatellite amplicons from mat DNA extracted at different seasonal time points. Adapted from Lehr et al., 2007.

field of functional genomics to determine what genes are actively produced within a cell under a given set of environmental conditions. Community mRNA can be used to reflect the physiology and genome regulation of an organism in response to varying environmental conditions. Through advances in sequencing and microarray technologies, non-model species with important ecological relevance can be assessed for contributions to important environmental systems in a cost-effective manner. Microarray technology has been successfully used to study the functional genomics of ecologically relevant species in pure culture settings, however *in situ* studies are still rare, and considered to be in a state of infancy. This technology is constrained by the availability genome sequence data and/or the development of an extensive cDNA library. As well, even with access to a large sequence dataset, probe sequences derived from *in vitro* or mixed cultures may not be representative of microorganisms active in an environmental system. Ultra-high throughput transcriptome sequencing of is an up-and-coming method used to combat this restraint. Without the need for prior sequence information, high throughput sequencing can generate a catalog of unique transcripts that can subsequently be used as a stand-alone representative of expression profiling or applied to downstream microarray development. Custom microarrays derived via this method should be more informative and show widespread representation of dominate microorganisms active in a given environment.

In *in vitro* settings, the fully sequenced and small genome of *C. merolae* has been utilized in microarray technology to better understand the process of organelle inheritance (24). As an important process in eukaryotic cell division, organelle

inheritance is poorly understood due to the fact that most eukaryotic organisms house multiple organelles (vacuole, chloroplast or mitochondria) that are not highly synchronized in their division and/or are difficult to track mechanistically (24). A single set of organelles, and ease of cellular synchronization has made *C. merolae* highly applicable to studying the transcriptome of the cell cycle (24, 25). Microarray technology using the *C. merolae* genome is also helping uncover how plant cells regulate and conduct nitrogen assimilation (33); here, DNA microarrays uncovered a transcription factor critical to the initiation of nitrogen assimilation. As well, the Cyanidiales extremophilic lifestyle is making it a target for research focused on identifying genes involved in regulating temperature, acid, and salt tolerance (32, 60, 61). So far, several candidate genes associate with increased expression following high salt exposure are promising in the field of genetic engineering; researchers believe flowering plants genetically transformed with *C. merolae* tolerance genes might be better able to combat salt stress introduced as a result of global warming (61).

Few studies have looked at *in situ* transcriptomics of microorganisms inhabiting acidic habitats. Recently, Moreno-Paz et al. (2008) examined the physiology of microorganisms associated with biofilm formation and dispersion within the extremely acidic waters (pH<2) of the Rio Tinto river (Spain). Through microarray analysis, results uncovered the up-regulation of metabolic pathways involved in acetate production as well as expression differences between biofilm bacteria in alternative growth phases. Results distinguished the main mechanisms leading to microbial biofilm formation in natural acid habitats (48). Similar studies have not been conducted in acidic geothermal

waters, nor have they looked at the physiology of eukaryotes in these extreme environments. Further research is needed to determine the environmental or physiological parameters that control the ecology of eukaryotes in extreme habitats.

## References


1. Aguilera, A, Souza-Egipsy, V, Gomez, F, and Amils, R. 2007. Development and structure of eukaryotic biofilms in an extreme acidic environment, Río Tinto (SW, Spain). Microb. Ecol. 53:294-305.

2. Albertano, P, Ciniglia, C, Pinto, G, and Pollio, A. 2000. The taxonomic position of *Cyanidium*, *Cyanidioschyzon*, and *Galdieria*: an update. Hydrobiologia. 433:137-143.

3. Allen, MB. 1959. Studies with *Cyanidium caldarium*, an anomalously pigmented chlorophyte. Arch. Microbiol. 32:270-277.

4. Azua-Bustos, A. 2010. Chilean Cave *Cyanidium.* In: Seckbach, J. and Chapman, D.J. (eds), Red algae in the genomic age. Springer Science and Business Media, B.V. pp. 425-438.

5. Barbier, G, Oesterhelt, C, Larson, MD, Halgren. RG, Wilkerson, C, Garavito, RM, Benning, C, and Weber, AP. 2005. Comparative genomics of two closely related unicellular thermo-acidophilic red algae, *Galdieria sulphuraria* and *Cyanidioschyzon merolae*, reveals the molecular basis of the metabolic flexibility of *Galdieria sulphuraria* and significant differences in carbohydrate metabolism of both algae. Plant Physiol. 137:460–474.

6. Bonheyo, G, Frias-Lopez, J, and Fouke, BW. 2005. A test for airborne dispersal of thermophilic bacteria from hot springs, In: Inskeep, WP, and McDermott, TR, (eds), Geothermal Biology and Geochemistry in Yellowstone National Park. Montana State University Publication, Bozeman, pp. 327-340.

7. Brock, T.D.1967. Life at high temperatures. Science. 158:1012–1018.

8. Brock, TD. 1978. The genus *Cyanidium*, In: Brock, TD (eds), Thermophilic microorganisms and life at high temperatures. Springer Verlag, New York, pp. 255-302.

9. Castenholz, RW, and McDermott, TR. 2010. The Cyanidiales: ecology, biodiversity, and biogeography, In: Seckbach, J. and Chapman DJ (eds), Red algae in the genomic age. Springer Science and Business Media, B.V. pp. 357-371.

10. Chapin, SF. 2002. Carbon input to terrestrial ecosystems. In: Matson, PA, and Mooney, HA., (eds) Principles of Terrestrial Ecosystem Ecology. Springer Science and Business Media, Inc, New York.

11. Ciniglia, C, Yoon, HS, Pollio, A, Pinto, G, and Bhattacharya, D. 2004. Hidden biodiversity of the extremophilic Cyanidiales red algae. Mol. Ecol. 13:1827–1838.

12. Cockell, CS, and Rothschild, LJ. 1999. The effects of UV radiation A and B in diurnal variation in photosynthesis in three taxonomically and ecologically diverse microbial mats. Photochem. Photobiol. 69:203–210.

13. Cunningham, FX, Lee, H, and Gantt, E. 2007. Carotenoid biosynthesis in the primitive red alga *Cyanidioschyzon merolae*. Eukaryotic Cell. 6:533-545.

14. Darienko, T, and Hoffman, L. 2010. Subaerial algae and cyanobacteria from the archaeological remains of Carthage (Tunisia), including the record of a species of *Cyanidium* (Rhodophyta). Algological Studies. 135:41-60.

15. De Luca, P, and Moretti, A. 1983. Floridosides in *Cyanidium caldarium*, *Cyanidioschyzon merolae* and *Galdieria sulphuraria* (Rhodophyta, Cyanidiophyceae). J. Phycol. 19:368–369.

16. Diaz, MM, and Maberly, SC. 2009. Carbon-concentrating mechanisms in acidophilic algae. Phycologia. 48:77–85.

17. Doemel, W, and Brock, TD. 1970. The upper temperature limit of *Cyanidium caldarium*. Arch. Microbiol. 72:326-332.

18. Doemel, WN, and Brock, TD. 1971. The physiological ecology of *Cyanidium caldarium*. J. Gen. Microbiol. 67:17–32.

19. Enami, I, Akutsu, H, and Kyogoku, Y. 1986. Intracellular pH regulation in an acidophilic unicellular alga, *Cyanidium caldarium*: $^{31}$P-NMR determination of intracellular pH. Plant Cell Physiol. 27:1351–1359.

20. Enami, I, Hideyuki, A, and Shen, J. 2010. Mechanisms of acido-tolerance and characteristics of photosystems in and acidophilic and thermophilic red alga, *Cyanidium caldarium*. In Seckbach, J, and Chapman, DJ (eds), Red algae in the genomic age. Springer Science and Business Media, B.V. pp. 373-389.

21. Ferris, MJ, Sheehan, KB, Kuhl, M, Cooksey, K, Wiggleworth-Cooksey, B, Harvey, R, and Henson, JM. 2005. Algal species and light microenvironment in a low-pH, geothermal microbial mat community. Appl. Environ. Microbiol. 71:7164-7171.

22. Foissner, W. 2006. Biogeography and dispersal of micro-organisms: a review emphasizing protists. Acta Protozool. 45:111-136.

23. Friedmann, I. 1964. Progress in the biological exploration of caves and subterranean waters in Israel. Int. J. Speleol. 1:29–33.

24. Fujiwara, T., Misumi, O., Tashiro, K., Yoshida, Y., Nishida, K., Yagisawa, F., Imamura, S., Yoshida, M., Mori, T., Tanaka, K., Kuroiwa, H., and Kuroiwa, T. 2009. Periodic gene expression patterns during the highly synchronized cell nucleus and organelle division cycles in the unicellular red alga Cyanidioschyzon merolae. DNA Res. 16:59–72.

25. Fujiwara, T, Kuroiwa, H., Yagisawa, F., Ohnuma, M., Yoshida, Y., Yoshida, M., Nishida, K., Misumi, S., Watanabe, S., Tanaka, K., and Kuroiwa, T. 2010. The coiled-coil protein VIG1 is essential for tethering vacuoles to mitochondria during vacuole inheritance of *Cyanidioschyzon merolae*. The Plant Cell. 22:772-781.

26. Geitler, L. 1933. Diagnoses neuer blaualgen von den Sunda-Insela. Arch. Hydrobiol. Suppl. 12.

27. Gross, W, Küver, J, Tischendorf, G, Bouchaala, N, and Busch, W. 1998. Cryptoendolithic growth of the red alga *Galdieria sulphuraria* in volcanic areas. Eur. J. Phycol. 33:25–31.

28. Gross, W, and Oesterhelt, C. 1999. Ecophysiological studies on the red alga *Galdieria sulphuraria* isolated from southwest Iceland. Plant Biol. 1:694–700.

29. Gross, W. 2000. Ecophysiology of algae living in highly acidic environments. Hydrobiologia. 433:31-37.

30. Gross, W, and Gross, S. 2001. Physiological characterization of the acidophilic red alga *Galdieria sulphuraria* isolated from a mining area. Nova Hedwigia, Beiheft. 123:523–530.

31. Gross, W, Oesterhelt, C, Tischendorf, G, and Lederer, F. 2002. Characterization of a non-thermophilic strain of the red algal genus *Galdieria* isolated from Soos (Czech Republic). Eur. J. Phycol. 37:477–482.

32. Hashimoto, M., Misumi, O., Hujiwara, T., Sakajiri, T., and Kuroiwa, T. 2006. Gene expression profiles by temperature change of the red alga *Cyanidioschyzon merolae*. J. Plant Res. 119 suppl. 74.

33. Imamura, S., Kanesaki, Y., Ohnuma, M., Inouye, T., Sekine, Y., Fujiwara, F., Kuroiwa, T., and Tanaka, K. 2009. R2R3-type MYB transcription factor, CmMYB1, is a central nitrogen assimilation regulator in Cyanidioschyzon merolae. Proc. Natl. Acad. Sci. USA. 106(30):12548-12553.

34. Inskeep, WP, and McDermott, TR. 2005. Geomicrobiology of acid-sulfate-chloride springs in Yellowstone National Park. In: Inskeep, WP and McDermott, TR, (eds) Geothermal biology and geochemistry in Yellowstone National Park. Montana State University, Bozeman, pp. 143-162.

35. Jackson, CR, Langner, HW, Donahoe-Christiansen, J, Inskeep, WP, and McDermott, TR. 2001. Molecular analysis of microbial community structure in an arsenite-oxidizing acidic thermal spring. Environ. Microbiol. 3:532-542.

36. Johnson, BA. 1998. Biodiversity and ecology of acidophilic microorganisms. FEMS Microbiol. Ecol. 27:307-317.

37. Kelly, DJA, Budd, K, Lefebvre, DD. 2007. Biotransformation of mercury in pH-stat cultures of eukaryotic freshwater algae. Arch Microbiol. 187:45-53.

38. Kura-Hotta, M, and Enami, I. 1981. Light-induced H+ efflux from intact cells of *Cyanidium caldarium*. Plant Cell Physiol. 22:1175–1183.

39. Kuroiwa, T, Nishida, K, Yoshida, Y, Fujiwara, T, Mori, T, Kuroiwa, H, and Misumi, O. 2006. Structure, function and evolution of the mitochondrial division apparatus. Biochim. Biophys. Acta. 1763:510–521.

40. Lehr, CR, Frank, SD, Norris, TB, D'Imperio, S, Kalinin, AV, Toplin, JA, Castenholz, RW, and McDermott, TR. 2007. Cyanidia (Cyanidiales) population diversity and dynamics in an acid-sulfate-chloride spring in Yellowstone National Park. J. Phycol. 43:3–14.

41. Macur, RE, Langner, HW, Kocar, BD, and Inskeep, WP. 2004. Linking geochemical processes with microbial community analysis: successional dynamics in an arsenic-rich, acid-sulphate-chloride geothermal spring. Geobiology. 2:163-177.

42. Mathur, J, Bizzco, RW, Ellis, DG, Lipson, DA, Poole, AW, Levine, R, and Kelley, ST. 2007. Effects of abiotic factors on the phylogenetic diversity of bacterial communities in acidic thermal springs. Appl. Environ. Microbiol. 73:2612-2623.

43. Matsuzaki, M, Misumi, O, Shin-i, T, Maruyama, S, Takahara, M, Miyagishima, SY, Mori, T, Nishida, K, Yagisawa, F, Nishida, K, Yoshida, Y, Nishimura, Y, Nakao, S, Kobayashi, T, Momoyama, Y, Higashiyama, T, Minoda, A, Sano, M, Nomoto, H, Oishi, K, Hayashi, H, Ohta, F, Nishizaka, S, Haga, S, Miura, S, Morishita, T, Kabeya, Y, Terasawa, K, Suzuki, Y, Ishii, Y, Asakawa, S, Takano, H, Ohta, N, Kuroiwa, H, Tanaka, K, Shimizu, N, Sugano, S, Sato, N, Nozaki, H,

Ogasawara, N, Kohara, Y, and Kuroiwa, T. 2004. Genome sequence of the ultra small unicellular red alga *Cyanidioschyzon merolae* 10D. Nature. 428:653–657.

44. Meneghini, G. 1839. Nuova specie di alga descritta dal sig. dott. giuseppe meneghini di padova. Nuovo Giorn. Lett. 39:67–68.

45. Merola, A., Castaldo, R., De Luca, P., Gambardella, R., Musacchio, A. and Taddei, R. 1981. Revision of *Cyanidium caldarium*: three species of acidophilic algae. Giorn. Bot. Ital. 115:189–195.

46. Miller, SR. 2007. Diversity of the cosmopolitan thermophile *Mastigocladus laminosus* at global, regional, and local scales. In: J. Seckbach (ed) Algae and cyanobacteria in extreme environments. Springer, Dordrecht. pp. 401–410.

47. Minoda, A., Nagasawa, K., Hanaoka, M., Horiuchi, M., Takahashi, H., and Tanaka, K. 2005. Microarray profiling of plastid gene expression in a unicellular red alga, *Cyanidioschyzon merolae*. Plant Molecular Biology. 59:375-385.

48. Moreno-Paz, M., Gomez, M.J., Arcas, A., and Parro, V. 2010. Environmental transcriptome analysis reveals physiological differences between biofilm and planktonic modes of life of the iron oxidizing bacteria *Leptospirillum spp.* in their natural microbial community. BMC Genomics. 11:404.

49. Nagasaka, S., Nishizawa, NK., Mori, S. and Yoshimura, EY. 2004. Metal metabolism in the red alga *Cyanidium caldarium* and its relationship to metal tolerance. Biometals. 17:177–181.

50. Nordstrom, DK, Ball, JW and McCleskey, RB. 2005. Ground water to surface water: chemistry and thermal outflows in Yellowstone National Park. In: Inskeep WP, and McDermott TR, (eds) Geothermal biology and geochemistry in Yellowstone National Park. Montana State University, Bozeman, pp. 73-94.

51. Nozaki, H, Takano, H, Misumi, O, Terasawa, K, Matsuzaki, M, Maruyama, S, Nishida, K, Yagisawa, F, Yoshida, Y, Fujiwara, T, Takio, S, Tamura, K, Chung, SJ, Nakamura, S, Kuroiwa, H, Tanaka, K, Sato, N, and Kuroiwa, T. 2007. A 100% complete sequence reveals unusually simple genomic features in the hot-spring red alga *Cyanidioschyzon merolae.* BMC Biol. 5:28.

52. Oesterhelt, C, Schmalzlin, E, Schmitt, JM, and Lokstein, H. 2007. Regulation of photosynthesis in the unicellular acidophilic red alga *Galdieria sulphuraria*. The Plant Journal. 50:500-511.

53. Pattanaik, B, Schuman, R, and Karsten, U. 2007. Effects of ultraviolet radiation on cyanobacteria and their protective mechanisms. In: Seckbach, J. (eds) Algae and Cyanobacteria in Extreme Environments. Springer, Dordrecht. pp. 31-45.

54. Phelps, D, and Buseck, PR. 1980. Distribution of soil mercury and the development of soil mercury anomalies in the Yellowstone geothermal area, Wyoming. Econ. Geol. 75:730-741.

55. Pinto, G, Ciniglia, C, Cascone, C, and Pollio, A. 2007. Species composition of Cyanidiales assemblages in Pisciarelli (Campi Flegrei, Italy) and description of *Galdieria Phlegrea* sp. nov. In: J. Seckbach (ed) Algae and cyanobacteria in extreme environments. Springer, Dordrecht. pp. 487–502.

56. Qin, J, Lehr, CR, Yuan, C, Le, X.C, McDermott, TR, and Rosen, BP. 2009. Biotransformation of arsenic by a Yellowstone thermoacidophilic eukaryotic alga. Proc. Natl. Acad. Sci. U.S.A. 106:5213-5217.

57. Raven, JA, Beardall, J, and Johnston AM. 1982. Inorganic carbon transport in relation to H+ transport at the plasmalemma of photosynthetic cells. In: Marme, D, Marre, E, Hertel, R, (eds), Plasmalemma and tonoplast: their functions in the plant cell. Elsevier Biomedical Press BB, Amsterdam, The Netherlands. pp. 41–47.

58. Redman, RS, Litvintseva, A, Sheehan, KB, Henson, JM, and Rodriguez, RJ. 1999. Fungi from geothermal soils in Yellowstone National Park. Appl. Environ. Microbiol. 65:5193-5197.

59. Reeb, V, and Bhattacharya, D. 2010. The thermo-acidophilic Cyanidiophyceae (Cyanidiales), In: Seckbach, J. and Chapman DJ, (eds) Red algae in the genomic age. Springer Science and Business Media, B.V, pp. 410-426.

60. Sakajiri, T., Misumi, O., Mori, T., Fujiwara, T., Hashimoto, M., and Kuroiwa, T. 2006. Monitoring expression profiles of *Cyanidioschyzon merolae* under salt and pH stress using DNA microarrays. J. Plant Res. 119 suppl. 75.

61. Sakajiri, T., Asano, K., Hirooka, S., Tashiro, K., Misumi, O., Fujiwara, T., and Kuroiwa, T. 2008. Microarray analysis reveals S-adenosylmethionine (SAM) synthetase involvement in salt tolerance of *Cyanidioschyzon merolae*. Cytologia. 73:341–368.

62. Seckbach, J, Ikan, R, Nagashima, H, and Fukuda, I. 1993. New phylogenetic status for acid hot spring algae. In: Sato, S, Ishida, H, and Ishikawa, H. (eds), Endocytobiology. V. Tòbingen University Press, Tobingen, pp. 241–254.

63. Seckbach, J. 1994. The natural history of *Cyanidium* (Geitler 1933): past and present perspectives, In: Seckbach, J. (eds), Evolutionary pathways and enigmatic algae: *Cyanidium caldarium* (Rhodophyta) and related cells. Kluwer, Dordrecht, The Netherlands, pp. 99–112.

64. Skowronski, T, Szubinska, S, Pawlik, B, and Jakubowski, M. 1991. The influence of pH on cadmium toxicity to the green alga *Stichococcus bacillaris* and on cadmium forms present in the culture medium. Environmental Pollution. 74:89-100.

65. Skuja, H. 1970. Alghe cavernicole nelle zone illuminate delle Grotte di Castellana (Murge di Bari). Le Grotte d'Italia 4:193–202.

66. Smith, DW, and Brock, TD. 1973. The water relations of the alga *Cyanidium caldarium* in soil. J. Gen. Microbiol. 79:219-231.

67. Stauffer, RE, Jenne, EA, and Ball, JW. 1980. Chemical studies of selected trace elements in hot-spring drainages of Yellowstone National Park. US Geological Survey Professional Paper 1044-F.

68. Takacs-Vesbach, MK, Jackson-Weaver, O, and Reysenbach AL. 2008. Volcanic calderas delineate biogeographic provinces among Yellowstone thermophiles. Environ. Microbiol. 10:1681-1689.

69. Toplin, JA, Norris, TB, Lehr, CR, McDermott, TR, and Castenholz, RW. 2008. Biogeographic and phylogenetic diversity of thermoacidophilic Cyanidiales in Yellowstone National Park, Japan, and New Zealand. Appl. Environ. Microbiol. 74:2822-2833.

70. Tortell, P. 2000. Evolutionary and ecological perspectives on carbon acquisition in phytoplankton. Limnol. and Oceanogr. 45:744-750.

71. Walker, JJ, Spear, JR, and Pace, NR. 2005. Geobiology of a microbial endolithic community in the Yellowstone geothermal environment. Nature. 434:1011-1013.

72. Weber, AP, Oesterhelt, C, Gross, W, Brautigam, A, Imboden, LA, Krassovskaya, I, Linka, N, Truchina, J, Schneidereit, J, Voll, H, Voll, LM, Zimmermann, M, Jamai, A, Riekhof, WR, Yu, B, Garavito, RM, and Benning, C. 2004. EST-analysis of the thermo-acidophilic red microalga *Galdieria sulphuraria* reveals potential for lipid A biosynthesis and unveils the pathway of carbon export from rhodoplasts. Plant Mol. Biol. 55:17–32.

73. Whitaker, RJ, Grogan, DW, and Taylor, JW. 2003. Geographic barriers isolate endemic populations of hyperthermophilic archaea. Science 301:976–978.

74. Wiegert, RG, and Mitchell, R. 1973. Ecology of Yellowstone thermal effluent systems: intersects of blue-green algae, grazing flies (*Paracoenia*, Ephydridae) and water mites (*Partnuniella*, Hydrachnellae). Hydrobiologia. 41:251-271.

75. Yoon HS, Hackett JD, Pinto G, Bhattacharya D. 2002. The single, ancient origin of chromist plastids. Proc. Natl. Acad. Sci. U.S.A. 99:15507–15512.

76. Yoon, HS, Hackkett, JD, Ciniglia, C, Pinto, G, and Bhattacharya, D. 2004. A molecular timeline for the origin of photosynthetic eukaryotes. Mol. Biol. Evol. 21:809-818.

77. Yoon, HS, Ciniglia, C, Wu, M, Comeron, JM, Pinto, G, Pollio, A, and Bhattacharya, D. 2006. Establishment of endolithic populations of extremophilic Cyanidiales (Rhodophyta). BMC Evol. Biol. 6:1-12.

78. Yoon, HS., Muller, KM, Sheath, RG., Ott, FD., and Bhattacharya, D. 2006. Defining the major lineages of red algae (Rhodophyta). J. Phycol. 42:482-492.

79. Yoshimura, E, Nagasaka, S, Satake, K, and Mori, S. 2000. Mechanism of aluminum tolerance in *Cyanidium caldarium.* Hydrobiologia. 433:57–60.

80. Zenvirth, D, Volokita, M, and Kaplan, A. 1985. Photosynthesis and inorganic carbon accumulation in the acidophilic alga *Cyanidioschyzon merolae*. Plant Physiology 77:237–239.

# CHAPTER 3

# CYANIDIALES TRANSCRIPTOMICS IN ACIDIC GEOTHERMAL ENVIRONMENTS

## Abstract

Ultra-violet (UV) and high-intensity visible (VIS) radiation are environmental stressors known to harm photosynthetic organisms through the generation of reactive intermediates that damage photosynthetic machinery. This study shows the potential of using a thermoacidophilic red alga of the order Cyanidiales to model *in situ* algal gene expression dynamics as a function of UV exposure and seasonal shifts in UV-VIS intensity. These algae exhibit a dynamic seasonal biomass fluctuation referred to as 'mat decline' where viability drastically decreases as seasonal UV-VIS irradiance intensity increases. In Yellowstone National Park (YNP), temporal experiments coupling UV irradiance manipulations (filtering) with whole-community transcription profiling revealed significant cyanidial gene expression changes occurring as a result of exposure to UV, and that patterns of response adjust across low and high irradiance time periods. Separate analyses examined genes responding to either increasing seasonal UV or VIS intensity, or by the combined effects of both irradiance wavelengths (UV and VIS). Results not only corroborated known physiological changes to solar irradiance, but also suggested the strategies employed to deal with excess VIS and UV intensity may be highly integrated. Finally, a suite of comparative analyzes determined the relative utility of environmental transcriptomics technologies in studying ecologically-relevant

expression patterns. Results suggest *in situ* expression profiles will improve understanding of how photosynthetic organisms are responding to environmental stressors as they are observed in nature.

## Introduction

The Cyanidiales are an order of unicellular eukaryotic red algae that thrive in acidic (pH 0.2-4.0) and high temperature (38-56$^{o}$C) geothermal environments. No other photosynthetic microorganisms are known to inhabit this combination of niche conditions, resulting in the Cyanidiales being a visually dominant component of the microbial community in high-temperature acidic environments. Due to their simple morphology, only three genera (*Cyanidium*, *Cyanidioschyzon*, and *Galdieria*) are formally recognized, based primarily on variations in cellular morphology and mode of replication (1). Members of the Cyanidiales differ from most plants in many aspects of their biology, such as their ability to influence the biogeocycling of arsenic and mercury (9, 26), their tolerance towards extreme environmental stresses (acid, temperature, salt) (23, 24), as well as their relatively simple cellular structure and minimally redundant genome (16). Many aspects of cyanidial ecology, however, remain poorly explored, presenting a high potential for new discoveries.

Abiotic stressors such as light, temperature, pH, and osmolarity are important factors known to influence algal growth and photosynthesis capacity (14). Recently, Lehr et al. (2007) described an annual cyclic decline in aqueous cyanidial mat communities in Yellowstone National Park (YNP), reporting that viable algal numbers declined significantly during the mat decline period, and that shifts in population size

corresponded to changes in 18s rDNA clone library composition and chloroplast short sequence repeat (cpSSR) profiles. In assessing possible causes, results from the study suggested temperature, UV irradiance, and/or visible (VIS) light irradiance were all important environmental factors contributing to the mat decline event. Although known as an important algal abiotic stressor, few mechanisms underlying cyanidial response to irradiance have been investigated so far, and only using pure cultures in laboratory settings (17, 18). Developing and applying molecular tools is a critical next step in helping to elucidate the mechanisms behind cyanidial response to irradiance *in situ*. Several genomic resources are available for these algae, including the completed *Cyanidioschyzon merolae* strain 10D genome and an expressed sequence tag (EST) dataset for *Galdieria sulphuraria* (16, 37). Though the *C. merolae* and *G. sulphuraria* genomes are used as models for studying the origin, evolution, and basic regulatory mechanisms of photosynthetic eukaryotes (6), molecular studies have not yet been undertaken to study how these algae deal with high levels of abiotic stress as a part of their natural environment.

With the genomic tools and genome sequences available, it is now possible to link *in situ* Cyanidiales gene expression with measurable environmental change at the level of the whole transcriptome, potentially identifying specific factors that may provide for a more fundamental understanding of mat decline. For this, the current study employed a custom oligonucleotide-microarray derived from ecologically-relevant Cyanidiales genomic and cDNA sequences, and used this array in field studies to examine cyanidial response to UV and VIS irradiance *in situ*. Gene expression patterns were monitored over

a time period of increasing UV-VIS intensity and as a function of irradiance that was manipulated through the use of filters which included or excluded UV radiation effects. Microarray analysis was also compared with pyrosequencing data from *in situ* cDNA libraries to identify genes that showed a statistically significant association with UV-VIS irradiance. By identifying both broad genome scale expression patterns, as well as individual responsive genes, work here aimed to build a basic framework of the Cyanidiales cellular response to UV-VIS irradiance. Finally, algal mats were also analyzed to assess the potential connectivity of the transcriptome to the proteome, helping to further understand algal responses to environmental change in nature.

## Materials and Methods

### Site Description, Sample Collection, and Aqueous Chemistry Analysis

All field experiments were conducted in YNP at the acidic geothermal features Dragon Spring in the Norris Geyser Basin (44°43’54.8” N, 110°42’39.9” W, spring number NHSP106 in the YNP thermal inventory) (Fig. 2.1A) and Lemonade Creek (44°48’3.3” N, 110°43’43.5” W, spring number APTNN033 in the YNP thermal inventory) (Fig. 3.2) located in Amphitheater Springs. Both sites have been described previously (12, 15). All sampling was from the visibly green algal mats. Samples were taken at varying time points at both locations from July 2007 to July 2010, and were collected axenically by transferring algal covered rocks into sterile Falcon$^{TM}$ tubes along with site water, vigorously shaking the tube by hand to dislodge the algal cells from the rock, transferring the algae to a sterile 15 mL Falcon tube, and flash freezing the material

in a dry-ice/ethanol bath. The entire process of dislodging mat material and transferring the falcon tube to the dry-ice/ethanol bath was completed within ~30-45 sec. The frozen samples were then transported back to the lab and stored at -80$^0$C until used for RNA or protein extraction.

To monitor air and water temperature, a data-logger (Datalogger Spectrum 1000, Veriteq Instruments, Richmond, BC, Canada) equipped with a thermistor sensor cable was used to record hourly water temperature measurements from Lemonade Creek, as well as the surrounding ambient air. If this data was unavailable due to wildlife tampering or equipment failure, single point measurements were recorded at the time of sampling using a hand held thermometer or an Accumet AP110 portable temperature/pH meter (Fisher Scientific, Pittsburgh, PA), which was also used for pH determination. UV irradiance was measured at sampling times using an ILT-1700 radiometer (International Light, Newburyport, MA, USA) as previously described (12). Long-term hourly measurements of visible (VIS) radiation were recorded using a LI-1400 datalogger equipped with a pyranometer (LI-COR, Lincoln, NE) that was installed at the Lemonade Creek field site.

Aqueous geochemistry of Lemonade Creek was monitored throughout the entire sampling period. Briefly, site water was filter-sterilized (0.22 μm) directly into sterile 50 mL Falcon tubes, and transported to the lab for analysis. A subset of filtered samples was acidified (1.0 mL of 16M $HNO_3$) in the field prior to transport and analyzed for total dissolved elements. Major anions and cations were measured using anion exchange

chromatography and inductively couple plasma optical emission spectrometry (ICP-OES) as described previously (12).

Two types of irradiance screening materials were used in the field experiments, both transparent to photosynthetically active radiation (PAR) and installed in triplicate (Fig. 3.2). The OP-4 filter material (Cyro Industries, Woodcliff Lake, NJ) transmitted about 90.0% of VIS, UVA, and UVB radiation, whereas the UF-5 filter material (Autohaas N. Am, Philadelphia, PA) transmitted about 90.0% VIS but blocked 99.9% of all UV below 390 nm (21).

Microarray Design

DNA sequences from a total of 6,837 genes were used as an input for microarray probe design, and were derived from three different sources:

i) Genomic DNA from two ecologically relevant isolates originating from Dragon Spring (CCMEE isolate 5508) and Lemonade Creek (CCMEE isolate 5578) was sequenced on a Solexa Genome Analyser (Illumina, San Diego, CA). Sequences were assembled to the *C. merolae* strain 10D reference genome using the Burrows-Wheeler Transformation (BWT) alignment tool (13). Base calls made to generate consensus sequences were determined by ≥ 90% agreement from a minimum of 20 aligned sequences. Genome coding regions from both isolates were identified with a BLASTn search against the *Cyanidioschyzon merolae* strain 10D genome database. Since the primary research site for the microarray study was Lemonade Creek, coding regions identified in isolate 5578 were considered most important for inclusion in the probe design. Therefore all full-length coding regions (5,978 sequences) identified in isolate 5578 were used in the

custom microarray construction. Coding regions unique to isolate 5508 (isolated from Dragon Spring), as well as isolate 5508 DNA sequence that complemented gapped or non full-length 5578 isolate sequences (425 sequences), were also included. Intergenic regions unique to either isolate 5578 (45 sequences) or 5508 (10 sequences) were also incorporated into the array design, and were identified based on their absence from the *Cyanidioschyzon merolae* strain 10D genome, and thus were considered potential strain-specific coding regions.

ii) ORF sequences (201 sequences) either unique to the *Cyanidioschyzon merolae* strain 10D genome or sequences from the 10D genome that filled gaps in the 5508 and 5578 isolate genomes were incorporated into the design.

iii) Coding regions identified in the cDNA pyrosequencing libraries (see below) but that were absent in the 5508, 5578, and the *C. merolae* strain 10D genomes were included and represented another potential source of strain-specific Cyanidiales transcribed regions. Coding regions were identified by BLASTx searches against the NCBI non-redundant protein database; 88 reads were identified from the January library while 90 reads were included from the July pyrosequencing data set.

From the 6,837 input sequences, 138 sequences were considered ‘exemplars’, meaning quality probes could not be adequately designed. The final array design included 6,699 unique gene sequences incorporated into a 12-plex Nimblegen Systems (Madison, WI) custom expression array. Five unique 60-mer oligonucleotide probes were designed per gene using a multi-step approach to select for probes with optimal hybridization. All

sequence probes with optimal hybridization properties were represented three times on each chip (i.e. three technical replicates per chip) for a total of 45 probes per gene.

RNA Preparation

RNA was extracted using a modified version of the FastRNA Pro Soil-Direct Kit (MP Biomedicals, Solon, OH). Frozen samples (~10 mL) were first thawed to the point of convenient removal from the Falcon tubes, transferred to pre-chilled 30 mL sterile Oakridge tubes, and vortexed in 2 minute intervals to break up the ice chunks (with rests on ice in between vortexing). Once completely thawed, the mat material was centrifuged (8000 x g, 5 min, 4$^{o}$C), and washed twice with 10 mL of ice-cold 0.85% NaCl (wt/vol). The algal pellet was quickly weighed to ensure a maximum of ~100 mg starting material was used, and subsequently suspended in 1 mL of RNA Pro Soil Lysis Buffer. Beginning with the FastRNA Pro Soil-Direct Kit, the suspension was transferred to a Lysing Matrix E tube and the RNA extraction continued with the kit according to the manufacturer's instructions.

Following extraction, RNA was DNase treated using a modified version of the TURBO DNA-free kit (Ambion, Austin, TX). Briefly, RNA was treated with 2 µL TURBO DNase per 50 µL sample volume and incubated at 37$^{o}$C for 1 hr. This process was repeated using 1 µL TURBO DNase per 50 µL sample volume and again incubated at 37$^{o}$C for 1 hr. RNA was purified by one round of extraction with phenol-chloroform-isoamyl alcohol (25:24:1). RNA was precipitated by the addition of 2.5 volumes of 100% ice-cold ethanol (EtOH) and 0.1 volume of 7.5M ammonium acetate and incubated overnight at -20$^{o}$C followed by centrifugation. The RNA pellet was washed once with

70% ethanol and suspended in nuclease-free water. DNA was confirmed to be absent by PCR tests containing 100 ng of RNA preparation, GoTaq Flexi DNA polymerase (Promega, Madison, WI), and the *rbc*L primers rbcL090F and rCR described previously (4, 40).

Microarray Sample Preparation,
Labeling, Hybridization, and Processing

The RNA from three replicate samples collected for each irradiance treatment was converted to double-stranded cDNA using a modified version of the SuperScript Double-Stranded cDNA Synthesis Kit (Invitrogen, San Diego, CA). Briefly, 10 µL of total RNA and 1 µM oligo dT (22) primer (Integrated DNA technologies, Coralville, IA) were mixed, incubated at 70$^{o}$C for 10 min., and placed in an ice water slurry for 5 min. *Dithiothreitol (*DTT) was added to a final concentration of 10 mM, each dNTP was added to a final concentration of 0.5 mM, and one volume of first-strand reaction buffer was added. This mixture was incubated at 42$^{o}$C for 2 min. 400 U of SuperScript II reverse transcriptase was then added and the mixture was incubated at 42$^{o}$C for 1 hr. The balance of the cDNA synthesis continued according to the manufacturer's instructions.

After synthesis, the double-stranded cDNA was purified by one round of extraction with phenol-chloroform-isoamyl alcohol (25:24:1). Double-stranded cDNA was precipitated by the addition of 3M NaOAc (pH 5.2), 2.5 volumes 100% ice-cold EtOH, and 24 µg of glycogen, and incubated at -80$^{o}$C overnight followed by centrifugation. The cDNA pellet was washed with 70% ice-cold EtOH, air dried at RT for 10 min, and suspended in nuclease-free water. The cDNA was cleaned using the Qiaquick PCR Purification Kit (Qiagen, Valencia, CA) and assessed for purity using the

RNA 6000 NanoChip assay on an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA).

Labeling was performed by mixing 1µg of cDNA sample with 1 U/µL Klenow fragment and 1 U/µL Cy3 fluorophore. This mixture was incubated at 37$^{o}$C for 2h. For each array, 4 µg Cy3-labeled cDNA was hybridized using the NimbleGen Hybridization Kit (NimbleGen Systems, Madison, WI) according to the manufacturer's instructions. Arrays were scanned using a GenePix 4000B scanner (Molecular Devices, Sunnyvale, CA) at a photomultiplier tube (PMT) detector voltage of 550. Scanned images were aligned, background corrected, and normalized using a robust multichip average (RMA) and NimbleScan software (Nimblegen Systems, Madison, WI). The probe detection limit was defined per array as being greater than four median absolute deviations (MAD) from the signal intensity of several negative control probes included in the array design. Low intensity probes which fell below this baseline expression threshold were excluded from analysis.

Microarray Data Analysis

Array data was analyzed using the FlexArray (version 1.6.1) software package available from Génome Québec *(http://genomequebec.mcgill.ca/FlexArray)*. A paired t-test was performed to identify statistically significant differences among the two UV irradiance treatments at each time point. Gene lists were trimmed to identify genes with fold-change differences of at least ±2 and p-values < 0.05. Hierarchical and k-means clustering analysis was performed with the Genesis software package (33) on data from each time point, using a Euclidean squared similarity metric and data centering.

Additional *k*-means analysis was performed using the EPCLUST software from the European Bioinformatics Institute (*http://www.bioinf.ebc.ee/EP/EP/EPCLUST/*). Genes were selected for clustering analysis if mRNA expression levels in the +UV (UV-exposed) Cyanidiales mats deviated from that in the –UV (UV-protected) mats by a factor of 2.0 (p-value <0.05) in at least one time point. This same selection criterion was also applied to separate analyses examining seasonal UV-VIS intensity effects. In this analysis mRNA expression levels in the July 2010 +UV or –UV algal mats needed to deviate from that in January 2010 +UV or –UV algal mats by a factor of 2.0 (p-value <0.05) to be included in the data set.

Finally, regression analysis was also employed as a means of assessing potential incremental impacts of seasonal shifts in VIS and UV irradiance on cyanidial gene expression patterns. Gene expression levels were analyzed against VIS and UV irradiance measurements ($W \cdot m^{-2}$) recorded at four time points over an 11 month sampling period using a one-way variable model. Calculations of r and p-value ($p < 0.05$) assessed the statistical significance. Microarray data have been deposited in NCBI's Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) and are accessible through the GEO Series accession number GSE37673.

cDNA Library Construction

RNA was collected and prepared from frozen January (low UV) and July (high UV) non-filtered cyanidial mats sampled at Dragon Spring as described above. RNA was converted to double-stranded cDNA using the Clontech SMART PCR cDNA synthesis kit (Clontech, Mountain View, USA), following the manufacturer's instructions except

that the CDS-3M provided in the Trimmer Direct kit (Envrogen, Moscow, Russia) was used instead of the 3' SMART CDS Primer II-A (22). cDNA was amplified using the Advantage 2 PCR kit (Clontech, Mountain View, USA) under the following condition: 18 cycles of 95$^{o}$C for 7s, 66$^{o}$C for 30s, and 72$^{o}$C for 6 min. Double-stranded cDNA was purified using the Qiaquick PCR Purification Kit (Qiagen, Valencia, CA) and adapters incorporated during the first strand synthesis were partially removed by *Sfi*I digestion (22). Following adapter digestion, the cDNA was purified using the Qiaquick kit.

454 Pyrosequencing and Assembly

Amplified cDNA from Dragon Spring was barcoded, pooled in equimolar amounts for equivalent sequencing depth, and sequenced at EnGenCore (University of South Carolina) on a Roche GS-FLX 454 automated pyrosequencer. Pyrosequencing reads were quality trimmed according to Kunin and colleagues (2010), followed by sequence comparisons and annotation based on a set of sequential BLAST searches. Beginning with standalone BLASTx and custom Perl scripts, sequences were first searched against the fully-sequenced *Cyanidioschyzon merolae* strain 10D genome (16). Remaining reads were annotated using BLASTx searches against the NCBI non-redundant protein database. The e-value cutoff was set at 1 x $10^{-5}$ for all searches. Sequences were checked for artificial replicate sequences using the open-source program CD-HIT (http://microbiomes.msu.edu/replicates) and protocols previously described (32). Statistical analysis of OTU richness via rarefaction and Chao1 estimates were performed in mothur.

Protein Preparation, SDS-PAGE Analysis,

In-gel digestion, and Nanospray LC-MS/MS

Microarray and 454 pyrosequencing cDNA libraries were also compared to a limited proteomic analysis. For the latter, protein extractions were performed on the same UV-exposed and UV-protected Lemonade Creek algal mats used in the microarray analysis. Briefly, mat samples were thawed and collected by centrifugation, washed in ice-cold 50mM phosphate buffer (pH7.0), and mixed with grinding buffer [50 mM Tricine-NaOH (pH7.8), 300 mM sorbitol, 10 mM NaCl, 5 mM $MgCl_2$, and protease inhibitor cocktail] and 1.0 g of 0.1-mm acid-washed glass beads (Sigma, St. Louis, MO) (36, 38). Cell lysis was achieved by bead beating the samples 3x for 1 min at 6.5 m/s in a FastPrep bead beater (MP Biomedicals, Solon, OH) with 5 min rests on ice in between each bead beating cycle. Proteins were precipitated by the addition of 5 volumes of ice-cold acetone and incubated at -80$^{o}$C overnight followed by centrifugation. The protein pellet was washed twice with 80% acetone, dried at RT for 10 minutes, and suspended in 8M Urea. Protein concentration was measured using the RC/DC Protein Assay (Bio-Rad, Hercules, CA). Following quantification, samples were kept frozen at -80$^{o}$C until use.

Purified proteins were denatured and separated on a 4 to 20% SDS-PAGE polyacrylamide gel (Bio-rad, Hercules, CA), following by staining with Coomassie brilliant blue R250 (Bio-rad, Hercules, CA). The most abundant protein bands were excised, destained with 50% acetonitrile in 50 mM ammonium bicarbonate (pH 7.9), and vacuum dried. Gel slices were rehydrated with 1.5 mg/mL DTT in 25 mM ammonium bicarbonate (pH 8.5) at 56$^{o}$C for 1 hr in a water bath. Gel slices were alkylated with 10 mg/mL iodoacetamide (IAA) in 25 mM ammonium bicarbonate (pH 8.5), and incubated

at RT in the dark for 45 min. Gel slices were washed once with 100 mM ammonium acetate (pH 8.5) for 10 min, washed twice with 50% acetonitrile in 50 mM ammonium bicarbonate (pH 8.5) for 10 min, vacuum dried, and rehydrated with 3 µL of 100 µg/mL Trypsin Gold (Promega, Madison, WI) in 25 mM ammonium bicarbonate (pH 8.5). Slices were covered in a solution of 10 mM ammonium bicarbonate with 10% acetonitrile (pH 8.5), and digested overnight at 37$^{o}$C followed by centrifugation. Peptides were analyzed by liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS). The search engine MASCOT (Matrix science, London, UK) was used to compare the masses of identified peptides to masses of sequences in the NCBInr database. Acceptable protein identifications required the detection of two significant peptides based on MASCOT MOWSE scores of greater than 32 (p-value < 0.05).

## Results

### Environmental Context

Previous temporally intensive work conducted at Dragon Spring (Norris Geyser Basin, YNP) suggested UV irradiance, VIS irradiance, and water temperature were all important physical factors influencing seasonal changes in cyanidial population dynamics and cell viability (12). Consequently, for the present study at Lemonade Creek (Amphitheater Springs, YNP), measurements for UV, VIS, water temperature, and aqueous chemistry were either continuously recorded with field instrumentation or measured during each field sampling, to document environmental conditions that might influence gene expression patterns.

Lemonade Creek was selected for this study because of the robust mats formed by the Cyanidiales algae, and because this creek is chemostat-like with respect to stable chemistry. Lemonade Creek is classified as an acid-sulfate-type geothermal feature due to a chemical signature that includes significant acidity and high concentrations of sulfate

Table 3.1 Aqueous geochemistry of Lemonade Creek water

| Constituent | Concentration (mM unless otherwise noted) | |
|---|---|---|
| | All Monthly Sampling (n=20) | Microarray Sampling Months (n=4) |
| *pH* | 2.76 ± .11 | 2.74 ± 0.29 |
| *Element* | | |
| Na | 2.1 ± 0.1 | 2.15 ± 0.12 |
| Si | 4.57 ± 0.26 | 4.53 ± 0.23 |
| K | 0.67 ± 0.04 | 0.67 ± 0.03 |
| Ca | 0.218 ± 0.040 | 0.201 ± 0.002 |
| Mg | 85.0 ± 6.0 µM | 80.0 ± 2.5 µM |
| Al | 0.434 ± 0.104 | 0.35 ± 0.06 |
| Fe | 55.3 ± 3.1 µM | 54.8 ± 2.2 µM |
| As | B/D | B/D |
| P | B/D | B/D |
| Mn | 2.38 ± 0.27 µM | 2.52 ± 0.57 µM |
| Zn | 2.23 ± 0.47 µM | 1.95 ± 0.52 µM |
| *Ion* | | |
| $Cl^-$ | 0.63 ± 0.18 µM | 0.64 ± 0.18 |
| $SO_4^{2-}$ | 7.38 ± 0.68 | 7.52 ± 0.47 |
| $F^-$ | 20.8 ± 6.0 µM | 19.9 ± 5.2 µM |

Water samples were taken directly above the algal mats. Concentrations are reported as the mean ± 1 SD of 20 samples taken approximately monthly over the 2-year study. Also depicted is the average concentration measured during the period examined by microarray analysis. Months where aqueous chemistry was impacted by spring snowmelt were excluded from measurement (i.e. April-June, depending on the timing of spring snowmelt). B/D, below detection.

(15). These chemical features, including pH, were measured from water samples taken directly above the cyanidial mats. Most chemical constituents showed relatively low

levels of variation when examined monthly over a two-year sampling period (Table 3.1). Concentrations and variations for the four specific time points included in the microarray analysis were quite similar to the long-term trends (Table 3.1). Na, Si, Ca, Mg, Fe, and$SO_4$ were all quite stable, displaying coefficients of variation of 0.1-6.3% (Table 3.1), whereas Al, Mn, $Cl^-$, and $F^-$ were more variable (coefficients of variation = 17.1-28.1%). For the latter, there were no obvious trends to the variation, but rather seemingly random oscillations between sampling dates. Al, Zn, and $F^-$ levels increased at the July 2010 microarray sampling date (coinciding with mat decline) (Table A3.1), but based on previous pure culture work documented in the literature, the concentrations of each constituent never approached levels that could be considered lethal or potentially growth inhibiting (20, 39). Chemistry variations were also within the range of previous observations noted for Lemonade Creek's geothermal source waters (15). The data is consistent with the interpretation that the underground flows feeding Lemonade Creek's four different geothermal tributaries may derive from difference source waters, with one or more contributing to the more variable chemical constituents. An important exception to the temporal consistency of the spring water occurred during spring snowmelt (e.g. April-June, depending on the start of spring runoff), when the physical and chemical properties of the creek water changed dramatically. Consequently, high snowmelt months were excluded from the analysis, and not part of the sampling regime.

In contrast to aqueous geochemistry, and as anticipated, the measurement of environmental physical factors previously linked to algal population dynamics (12) showed significant changes across the Julian calendar year (Fig. 3.1). Sporadic data loss

was encountered with VIS measurements due to either equipment malfunction during extreme weather conditions or wildlife tampering, however, expected seasonal trends were still obvious. Low water temperatures observed in the spring were due to Lemonade Creek serving as the primary drainage route for the canyon in which it is located. Following sustained air temperatures levels above 16°C (~Julian day 104), significant

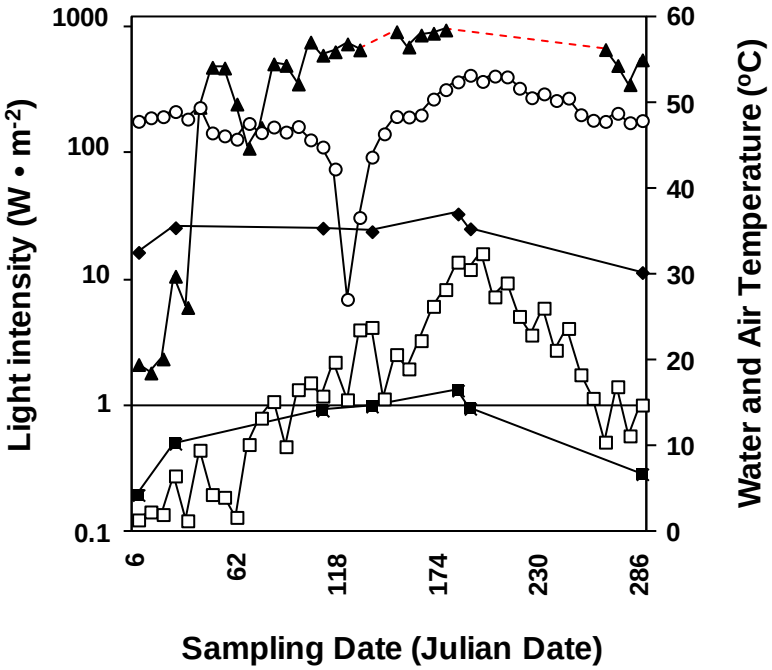


Fig 3.1 Temporally intensive measurements of environmental physical factors at Lemonade Creek. Water temperature (○), air temperature (□), UVA (315-400 nm,◆), UVB (280-315 nm,■), and VIS light (400-700 nm,▲) intensity as a function of time (shown in Julian days). Visible light, water, and air temperature intensities are the average of seven maximum daily point reading recorded over the course of a week. To account for differences in seasonal photoperiod length, only visible light measurements recorded between the hours of 8 AM and 4 PM were used to calculate the daily maximum. UV irradiance was measured during field sampling trips. Large swings in water temperature are reflective of weather conditions; spring snowmelt at Lemonade Creek typically starts in mid-April (~Julian day 118) and lasts until mid to late May (~Julian day 146). Discontinuities in visible light recordings, as a result of either equipment malfunction or wildlife tampering, are displayed as dashed red lines between data points.

decreases in water temperature (~27$^{o}$C) were observed (Fig. 3.1), corresponding to frigid snowmelt which dominated the creek flow. Spring runoff also resulted in high water volume, which altered the creek bed, and swept away the algal mats. Once spring snowmelt subsided, water temperatures increased rapidly (Fig. 3.1) due to the creek flow being exclusively made up of water from the thermal tributaries. This re-established optimum growth temperatures for the Cyanidiales, which subsequently formed lush mats (e.g. Fig. 3.2). Following re-colonization, water temperatures continued to rise, peaking at 54$^{o}$C after the summer solstice (Julian day 172), which also coincided with maximal summer air temperatures (Fig. 3.1).

Solar irradiance levels also changed in an expected temporal fashion (Fig. 3.1). Highest intensities were measured from late-June through August, with UVA readings peaking at 33.4 W·m$^2$, UVB at 1.3 W·m$^2$, and maximal VIS levels of 869.6 W·m$^2$. Lowest irradiance values were recorded near the winter solstice, decreasing to 16.6 W·m$^2$ for UVA, 0.2 W·m$^2$ for UVB, and 10.8 W·m$^2$ for VIS. Beside the Earth's axial tilt that underlies the seasonal change in irradiance, accurate winter VIS readings were at times obscured due to periods of extensive snow cover that blocked VIS light sensors deployed at the field site (i.e. Julian days 6-21, Fig. 3.1). For logistical reasons, sampling during the winter months was less frequent and therefore maintenance of field equipment was not always optimum with respect to data acquisition. However, point measurements obtained with the portable light meter illustrated that VIS levels during winter were in the range of 80 W·m$^2$ (Julian day 28), more than ten-fold lower than maximal summer levels (Fig. 3.1), thus displaying significant seasonal variation.

UV Irradiance-induced Changes in Gene Expression

Temporal algal biomass sampling and RNA extractions were used to examine gene expression patterns as a function of relative UV and VIS exposure during low UV (winter), increasing UV (spring), highest UV (summer), and decreasing UV (fall) periods of the year (Fig. 3.1). To identify differentially regulated genes, a pair-wise comparison of gene expression levels in the +UV mats (UV-exposed algae) was compared to the –UV mats (UV-protected algae) and are discussed herein in terms of fold changes in the +UV:-UV expression ratio.

A general evaluation of the Cyanidiales transcriptional response to changing seasonal UV irradiance levels is depicted in Fig 3.3. Apparent UV-influenced expression patterns were most evident in January 2010, where the +UV:-UV transcript ratio of 182 genes was statistically significant (p-value $< 0.05$; fold change $\geq \pm 2$). The majority (92%) of the UV-sensitive genes at this time point were down-regulated (Fig 3.3), with most annotated as repetitive DNA elements (Table A3.2). Though the total number of genes significantly up-regulated in January 2010 was proportionally small (8% of the total), important metabolic pathways were activated in response to UV exposure, including the regulation of genes involved in translation (*rps10*, *rpl14*, *rpl6*) and electron transport (*nad4L* and*nad5*) (Table A3.2).

Somewhat surprisingly, UV-linked expression patterns were far less evident during summer, the time of year when UV levels are highest (Fig. 3.1; Fig. 3.3) and mat decline effects are visually evident. Only three genes exhibited significant +UV:-UV expression ratios in July 2010, all being up-regulated between 2 to 3-fold (Fig. 3.3). One

up-regulated transcript (CMR353C) encoded for a protein involved in respiratory burst response, a well known plant defense mechanism that is activated during pathogen infection events (27). The three other significantly up-regulated genes all encoded for hypothetical proteins (Table A3.3)

Figure 3.2 Lemonade Creek mat with filter setups (12 February). Green mat color is typical of Cyanidiales mats in the 37-55$^{o}$C range. Pairs of OP-4 (+UV) and UF-5 (-UV) filters for each of three biological replicates are shown.

With respect to UV-influenced expression patterns observed during transitional UV intensity periods, (April 2010, UV levels increasing; November 2009, UV levels decreasing), the total number of genes with UV-responsive changes decreased as seasonal irradiance intensified (Fig. 3.1; Fig. 3.3). In April 2010, 17 genes displayed significant +UV:-UV expression ratios, with the majority (59%) down-regulated (Fig. 3.3). A

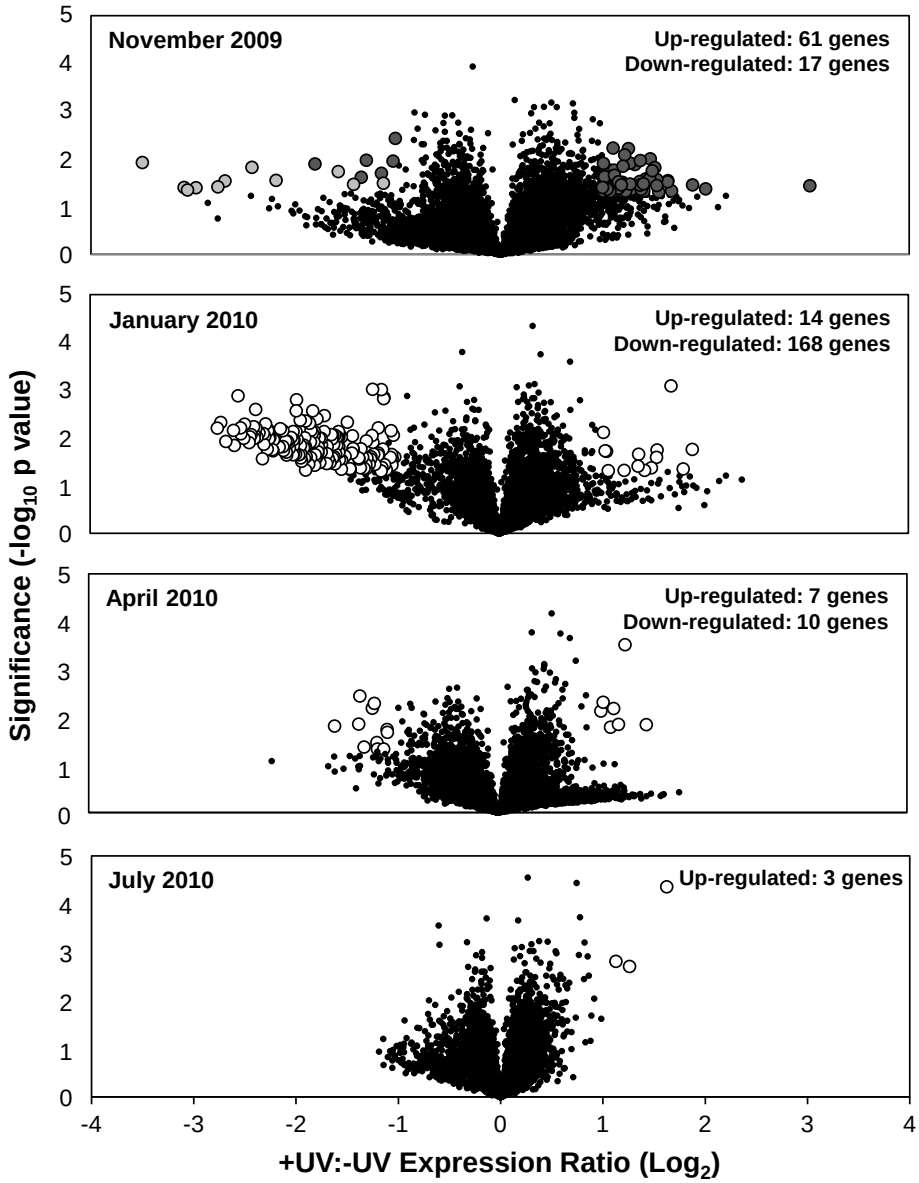


Fig. 3.3 Volcano plot displaying genes with significant UV exposure response at four time points. Genes with statistically significant +UV:-UV expression ratios (≥2 fold-change; p-value < 0.05) in January 2010, April 2010, and July 2010 are illustrated as open circles, while genes with no significant UV transcriptional response are shown as black circles. In November 2009, nuclear-encoded genes with significant +UV:-UV expression ratios are illustrated as dark-grey circles, plastid or mitochondrial-derived genes with significant UV-response are depicted as light-grey circles, and genes with no significant response are shown as black circles. The total number of genes with significant +UV:-UV expressions ratios at the four time points are displayed within each panel.

functional assessment in April 2010 revealed the overall repression of the translational machinery, with many genes encoding for ribosomal proteins (*rps4*,*rps12*, *rpl5*, *rpl32*) (Table A3.4). As well, the transcript level of *infB*, encoding protein synthesis initiation factor 2, was down-regulated roughly 2-fold (Table A3.4). Components of the translational machinery also showed statistically significant +UV:- UV expression ratios in November 2009; however, unlike April 2010, these genes were not always repressed in response to UV-exposure. Specifically, nuclear-encoded genes involved in translation, including several large ribosomal proteins (*rpl23A*, *rpl34*, *rpl35A*), a ribosome biogenesis enzyme (*rio1*), and a eukaryotic translation factor (*eIF-1*) were all activated between 2 to 8-fold in response to UV-exposure, while a chloroplast- derived small ribosomal protein (*rps18*) was down- regulated roughly 6-fold (Table A3.5). In general, November 2009 UV-linked expression patterns revealed that all up- regulated genes (n=61) were nuclear-encoded, while of the 17 genes down-regulated in response to UV-exposure, 65% were either plastid or mitochondrial (Fig. 3.3; Table A3.5). For example, UV-exposure resulted in the systematic repression of electron transport complexes, including genes coding for the cytochrome b6-f (*petG*, *petM*), NADH dehydrogenase (*nad1*,), succinate dehydrogenase (*sdhB*), cytochrome c oxidase (*cox1*, *cox2*), and cytochrome c reductase (*cytB*) (Table A3.5). Additional comparisons examining genes with significant +UV:-UV expression ratios that were shared among the two transitional time points (i.e. November 2009 and April 2010) and either low (January 2010) or high (July 2010) UV periods of the year, uncovered little overlap in encoded function (Table A3.4; Table A3.5). As well,

no overlap in encoded function was noted amongst the transitional time points themselves.

## Temperature and Irradiance Specific Patterns of Gene Expression

A *k*-means clustering analysis was performed to group genes according to similarities in their expression profiles across seasonal periods that corresponded to the cyanidial mat decline event (Fig. 3.3). As presented here, the analysis began with samples taken in November 2009 and progressed to July 2010, during which time UVA levels increased from 13.1 to 33.4 $W \bullet m^2$, UVB from 0.28 to 1.34 $W \bullet m^2$, and VIS from 80.02 to 869.66 $W \bullet m^2$ (Fig. 3.4). During this same time frame, water temperature changes were more modest, increasing from 46 to 54$^{o}$C (Fig. 3.4). For a gene to be included in the clustering analysis, a statistically significant change (defined as ≥2-fold) in the +UV:-UV expression ratio needed to occur in at least one of the four time points. With this as the selection criteria, expression analysis identified and visualized responsiveness of 258 genes in three clusters. Within each cluster, genes displayed similar patterns of expression changes (Tables A3.6-A3.8).

In cluster A, 32% and 99% of the +UV:-UV expression ratios were down-regulated ≥2-fold in November 2009 and January 2010, respectively (Fig. 3.5). As solar irradiance increased in April 2010 (Fig. 3.4), the +UV:-UV expression ratios for cluster A increased, with 28% of the genes up-regulated ≥2-fold (Fig. 3.4; Fig. 3.5). When irradiance levels peaked in July 2010 (Fig. 3.4), however, the transcription response for 98% of cluster A genes appeared neutral to UV effects, indicating no gene response at the highest seasonal UV intensities (Fig. 3.4; Fig. 3.5). The majority (84%) of genes in

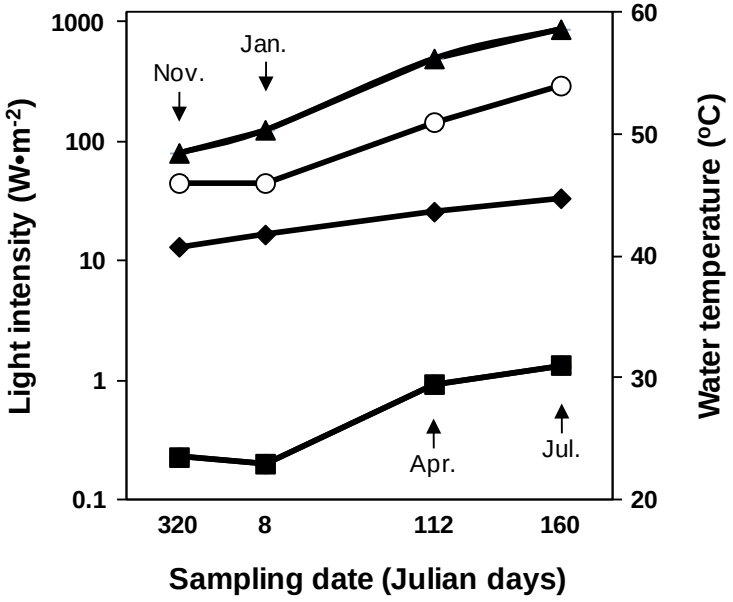


Fig. 3.4 Changes in water temperature and UV-VIS intensity at Lemonade Creek occurring at four time points. Illustrated are water temperature (○), UVA (315-400 nm, ◆ ), UVB (280-315 nm, ■), and VIS light (400-700 nm, ▲) measurements recorded at the time of Cyanidiales microarray sample collection.

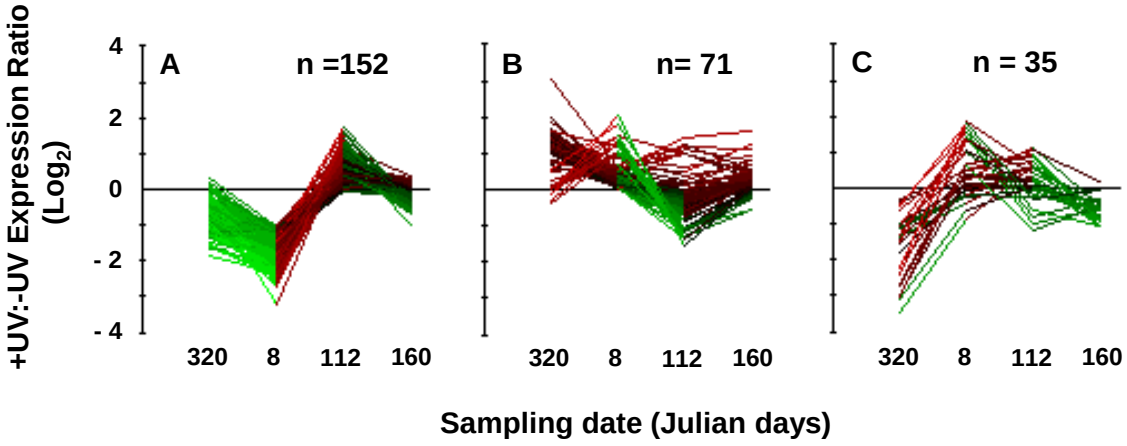


Figure 3.5 Gene expression patterns as a function of water temperature and UV-VIS intensity at Lemonade Creek. *K*-means clustering analysis (258 genes in total) of the mRNA profiles in response to seasonal UV-irradiance shifts. The data are plotted as the $Log_2$ fold change of the +UV:-UV expression ratio across the four time points shown in Fig. 3.4. The number of genes in each cluster is indicated. *Red* and *green* colors depict expression change, either an increase or decrease in the +UV:-UV expression ratios under the UV-exposed cyanidial mats.

clusters A coded for DNA repetitive elements (Table A3.6). Genes with identifiable functions included three class I elements (retrotransposons) (CMF037X, CMK165C, CMP121C), and two class II elements (transposons) (, CMO148C) (Table A3.6).

All cluster B genes displayed +UV:-UV expression ratios that were up-regulated in November 2009 (Fig. 3.5), with 87% up-regulated ≥2-fold. This included the activation of translation-related genes coding for functions involved in translational machinery (*rplB, rpl5, rpl23A, rpl34, rpl35A*), ribosomebiogenesis (*rio1*), and RNA transport (*eIF1, eIF2, eIF3*) (Table A3.7). Across the remainder of the sampling times, <22% of cluster B genes exhibited any significant UV response (Fig. 3.5).

Genes grouping in cluster C displayed patterns of UV-sensitivity that varied considerably over the four sampling time points. In November 2009, 60% of cluster C genes displayed +UV:-UV expression ratios that were down-regulated (Fig. 3.5), whereas in January 2010, 49% of all cluster C genes appeared up-regulated by UV (Fig. 3.5). Tracking the seasonal progression into spring (April 2010), 29% of all cluster C genes were now repressed ≥2-fold (Fig. 3.5). Irradiance response patterns were altered again by July 2010, when maximal irradiance intensities were recorded (Fig. 3.4); here all but two genes were neutral with respect to UV-exposure response (Fig. 3.5). Metabolic functions associated with cluster C uncovered a large number of genes (23%) involved in electron transport, including genes associated with the electron carrier complexes NADH dehydrogenase (*nad1, nad4L, nad5*), succinate dehydrogenase *(sdhB, sdhC*), and cytochrome c oxidase (*cox1, cox2, cytB*) (Table A3.8). Genes encoding proteins involved in ribosome biosynthesis (*rpl6, rpl14, rpl20, rpl32, rps4, rps10, rps12, rps18*), and

phylloquinone and menaquinone biosynthesis (*menB, menC*)were also part of cluster C (Table A3.8).

UV and VIS Irradiance Effects

The *k*-means clustering analysis suggested other environmental factors were muting UV effects; i.e. based on the criteria used in this study and statistical analysis of the +UV:-UV expression ratios at the July 2010 sampling, it appeared that 99% of the genes represented on the microarray were not influenced by UV during the season with the highest UV irradiance intensity (Fig. 3.4; Fig. 3.5). As VIS irradiance was documented to differ by roughly an order of magnitude between summer and winter (Fig. 3.4), separate analyses undertook a direct comparison of gene expression levels at the lowest (January 2010) and highest (July 2010) UV-VIS irradiance periods of the year (Fig 3.4), identifying cyanidial genes which were regulated by UV or VIS separately, or by UV-VIS combined. For a gene to be included in the analyses, a statistically significant change ($\geq \pm$2-fold) in the July:January expression ratio needed to occur in either the OP-4 (UV + VIS exposed) or UF-5 (VIS-only exposed) filtered algal mats. Gene responses were then compared to identify those that were unique to the algal mats shielded from UV (UF-5 filter, VIS only), those with similar transcription levels under both filter types (influenced by UV + VIS), and those that exhibited significant expression differences under the OP-4 filtered algal mats (exposed to all wavelengths), but that were not influenced by VIS wavelengths alone; the latter were concluded to be specifically influenced by UV alone.

Apparent VIS-specific gene expression changes were prominent; 189 genes

displayed statistically significant July:January expression ratios in the absence of UV irradiance. Genes repressed in response to increased VIS intensity primarily encoded for repetitive DNA elements (Table 3.2; Table A3.9). A significant number of genes (n= 62) were up-regulated in response to increasing VIS intensity, notably genes associated with the translational machinery (*rpl6*, *rpl14*, *rpl32*,), ATP synthesis (*atp6*), electron transport (*cox1*, *cytB*, *nad1*, , *nad4L*, *nad5*), and chaperone/heat shock response (*groEL*, *tcp1*, *hsp105*) (Table A3.9).

Given the apparent correspondence between VIS intensity and gene expression changes, additional analyses sought to assess whether such VIS responses were incremental or potentially triggered by a specific VIS irradiance threshold intensity. Separate linear regression analysis associated (p-value < 0.05) the expression of 42 genes with incremental VIS increases (Table A3.10). Genes exhibiting expression levels negatively correlated to VIS intensity were involved with DNA replication (*mcm4*, *mcm8*) and cell division (*ftsZ*). A number of genes involved in heat shock response and protein folding/refolding showed a positive association with shifts in VIS intensity, including the heat shock regulon member *dnaJ* and several chaperone genes (e.g. *tcp1*, *pcmT*, *clpS*) (Table A3.10).

Fewer genes were identified when making the same July:January expression comparison for algal mats exposed to both UV and VIS (Table A3.12), which would indicate possible UV influences. Comparing statistically significant July:January expression ratios in both the UV + VIS exposed and VIS-only exposed algal mats identified 162 genes that shared the same expression tendencies; i.e. genes up- or down-

regulated were in agreement and illustrated evidence of internal consistency within the study (Table A3.11). There were interesting patterns evident that provided evidence of UV influences, specifically showing that UV irradiance appeared to attenuate VIS repression affects, i.e. the level of repression in the UV + VIS exposed algal mats was less than algal mats exposed to VIS irradiance alone. There were 42 genes where the VIS repression was reduced >50%, with all but 4 coding for repetitive DNA elements (Table A3.11). These results correlate well with the *k*-means clustering analysis which found DNA repetitive elements were down-regulated in response to UV-VIS exposure (Table A3.6). Genes coding for proteins associated with cell division (cyclin A2, *ftsZ, mnd1, mad2*) and DNA replication (*mcm2*, *mcm4*, *mcm5*, *mcm6*, *mcm7*) (Table 3.2; Table A3.11) were significantly down-regulated to the same degree under both the UV+VIS and VIS exposure conditions, suggesting these genes were actually being affected by VIS wavelengths. Heat-shock and chaperone genes were up-regulated in the July:January expression ratios under both the UV+VIS and VIS-specific genes expression changes (Table A3.9; Table A3.11), likewise suggesting VIS-specific regulation.

Gene expression changes occurring specifically as a result of seasonal increases in UV intensity were determined by identifying those genes which were differentially regulated when exposed to both UV+VIS (i.e. algal mats under the OP-4 filters), but that were not represented in the list of genes found to be differentially regulated by VIS alone. In this category a total of 91 genes were significant in their July:January expression ratios with most (62%) down-regulated (Table 3.2). Notably, a photolyase enzyme (CmPHR5), well known to be involved in the repair of UV-induced damage (2), was up-

Table 3.2 Summary of gene expression sensitivity occurring as a result of both UV and VIS intensity (UV + VIS), only UV intensity (UV only), or only VIS intensity (VIS only). Shown are the functional role categories for genes significantly up- (+) or down-regulated (-) in response to increasing seasonal irradiance. Note that only repetitive DNA elements were affected by both UV and VIS exposure (Table A3.11), and in each case UV appeared to attenuate VIS repression affects.

| | UV + VIS | UV Only | VIS Only |
|---|---|---|---|
| Functional Role Category | Number of Genes | Number of Genes | Number of Genes |
| ABC Transporters | | | +1 |
| Arachidonic acid metabolism | | | +1 |
| Arginine and Proline metabolism | | -1 | +1 |
| Base Excision Repair | -1 | | |
| Carotenoid biosynthesis | | +1 | |
| Cell Cycle | -4 | | +2/-1 |
| Citrate Cycle (TCA cycle) | | -1 | |
| DNA Repetitive Elements | -55 | -2 | +1/-66 |
| DNA Replication | -5 | -2 | |
| Fatty Acid Biosynthesis | | -1 | |
| Folate biosynthesis | | | +2 |
| Glutathione metabolism | | +1 | |
| Glycerolipid Metabolism | -1 | | |
| Glyoxylate and dicarboxylate metabolism | | +1 | |
| Lipoic Acid Metabolism | | -1 | |
| N-glygan biosynthesis | | | +1 |
| Nitrogen metabolism | | +2 | |
| Other types of O-glycan biosynthesis | | | |
| Oxidative Phosphorylation | -1 | -1 | +6 |
| Pentose and glucuronate conversions | | +1 | |
| Peroxisome | +1 | | |
| Phagosome | -1 | | |
| Phe, Tyr, and Trp biosynthesis | -1 | -1 | |
| Photosynthesis | -1 | -2 | +2 |
| Photosynthesis-Anteanna Proteins | | -2 | |
| Porphyrin and Chlorophyll Metabolism | -5 | -2 | |
| Protein Export | | | |
| Protein processing | +1 | | +1 |
| Pyrimidine Metabolism | -1 | | |
| Ribosome | +1/-1 | -5 | +3/-1 |
| Ribosome biogenesis in eukaryotes | | | +1 |
| RNA Degradation | | | +1 |
| RNA Polymerase | | +1 | |
| SNARE Interactions | | | +2 |
| Sulfur Relay System | | -1 | +1 |
| Translation Factors | | -3 | |
| Ubiquinone biosynthesis | | | +3 |
| Ubiquitin mediated proteolysis | | | |
| All Others | +30/-52 | +29/-30 | +59/-33 |

regulated ~2.5-fold exclusively as a result of increased UV intensity (Table A3.12). Repression of photosynthetic activity was also found to be a UV-specific response, and included genes associated with photosystem II (*psbT*, *psbW*), the phycobilisome complexes (*apcA*, *cpcG*), and chlorophyll metabolism (CMA002Z, CMA141Z) (Table 3.2; Table A3.12).

Additional regression analysis identified eight genes exhibiting expression that could be statistically linked with either UV-A, UV-B or both (Table A3.13). Correlation with both UV-A and UV-B, might be expected because these wavelengths co- varied across the study (Fig. 3.4), however, three genes were found to be influenced by UV-B alone, including a DNA repetitive element ($r = -0.951$), a reactive oxygen intermediate ($r = -0.966$), and a putative glutathione S-transferase ($r = 0.954$). The latter is potentially interesting with respect to mat decline in that this gene codes for a radical scavenging enzyme known to be involved in stress responses in red algae (25).

Finally, regression analysis was also used to evaluate whether water temperature was potentially impacting gene expression patterns. Seasonal changes in water temperature were somewhat modest, increasing by $8^{o}C$ from winter to summer (Fig. 3.4), and failed to correlate (negatively or positively) with the expression of any of the 6,699 genes represented in the microarray.

Transcriptomic Method Comparison

In addition to enhancing the microarray design, the pyrosequencing of cDNA libraries constructed from non-filtered cyanidial mats sampled at Dragon Spring during both low (January) and high (July) UV irradiance periods was used for comparing

transcriptomes characterized via different methods; i.e. pyrosequencing versus microarrays. A total of 35,062 and 72,838 quality-filtered pyroreads were acquired from the January and July cDNA libraries, respectively. Reads were annotated using a BLASTx search against the *C. merolae* strain 10D genome database. The remaining reads that did not match the genome from *C. merolae* strain 10D, were identified through a secondary BLASTx search against the NCBI-nr protein database. Only BLASTx hits with *e*-values less than $1 \times 10^{-5}$ were used for annotation designations. Analysis of the pyrosequencing reads revealed 343 (January, 17,907 reads) and 259 (July, 47,091 reads) different *Cyanidioschyzon* genes (Fig. 3.6A).

Once annotated, a comparison was made to the microarray analysis conducted at Dragon Spring, looking at the degree of overlap in transcript diversity and overall method sensitivity. To do this, microarray data sets that were also derived from non-filtered algal mats sampled in July and January at Dragon Spring were analyzed for positive probe signal. Genes with signal intensities above that of negative control probes (+2 SD) included in the array design were used for comparison. Statistically significant expression of 6,117 genes was detected for the July microarray while 6,026 genes were detected on the January microarray (Fig. 3.6A). Transcripts identified in the pyrosequencing cDNA libraries showed a high degree of overlap with the microarrays. Approximately of 87% all transcripts detected in the cDNA libraries were likewise found to be expressed in the corresponding microarray (Fig. 3.6A).

Comparative analyses also looked at whether abundant genes detected in either the July or January pyrosequencing library displayed enhanced probe signal intensity in

the microarrays. For an identified pyrosequencing gene to be included in this analysis, transcript abundance was somewhat arbitrarily set at ≥0.1% frequency of the total

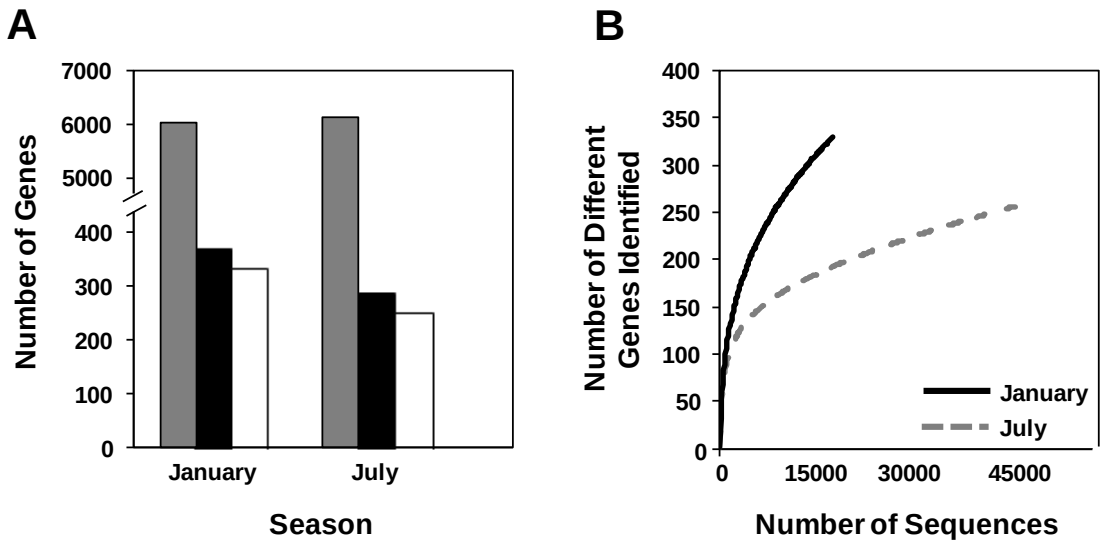


Figure 3.6 Transcript diversity detected by pyrosequencing libraries and microarrays across two seasonal periods at Dragon Spring. **(A)** Grey bars, total number of genes for which statistically significant expression was detected in the microarrays; Black bars, number of genes identified in the cDNA pyrosequencing libraries; White bars, number of genes detected in both the microarray and cDNA library **(B)** Collector's curves for the January (black line) or July (grey line) cDNA pyrosequencing libraries at Dragon Spring. The number of unique reference genes identified via BLASTx (e-values < 1 x $10^{-5}$) is shown as a function of the number of sequences analyzed.

number of reads in either the January (≥ 18 reads • $gene^{-1}$) or June (≥ 38 reads • $gene^{-1}$) cDNA library, thus ensuring comparisons were being made between significantly expressed genes. Results uncovered a large number of abundant July cDNA pyrosequencing transcripts that displayed significant microarray probe signal intensity; 52% of all the abundant July transcripts showed probe signal intensities that ranked in the top 30% among all 6,117 statistically significant array genes. Likewise, 54% of all

abundant January pyrosequencing transcripts ranked in the top 30% in probe signal intensity on their corresponding microarray.

Collector's curve analysis predicted a large level of transcript diversity remained undetected in the pyrosequencing libraries (Fig. 3.6B). Results suggest significantly greater sequencing depth would be necessary in order to capture the full profile of gene expression activity via pyrosequencing. Apparent transcriptome coverage in the summer cDNA library was clearly more extensive than that obtained for the winter sample (Fig. 3.6B), and no doubt was due, at least in part, to the fact that the winter pyrosequencing library was roughly half the size of that obtained for the summer library. Interestingly though, gene richness was found to be substantially higher in the winter library, suggesting a higher proportion of the genome is transcriptionally active during the time of the year with minimal UV intensities.

Though different in their diversity and coverage levels, both pyrosequencing libraries detected significant apparent allelic richness. Examples are illustrated by a random sampling of ten genes with varying degrees of transcript abundance from each cDNA library, using a Chao1 richness metric (set at 97% gene sequence identity) to estimate total allelic diversity. There were numerous sequence variants in both cDNA libraries (Table A3.14; Table A3.15). In general, and as would be expected, predicted sequence richness increased with read numbers, however, there were exceptions showing that allelic richness varied as a function of the specific gene being examined, and that transcript diversity was not always associated with relative read abundance.

Despite low coverage depth, both pyrosequencing libraries detected putative genes which were absent from the *C. merolae* strain 10D genome. Secondary BLASTx searches against the NCBI-nr protein database resulted in the identification of an additional 29 and 31 genes from the January and July pyrosequencing libraries, respectively. In the July library, 13 eukaryotic genes were identified (e-score range: $5 \times 10^{-9}$ to $5 \times 10^{-21}$), with the majority (65%) having a best hit to hypothetical proteins derived from either a fungus or plant species (Table A3.16). Similar trends were observed in the January pyrosequencing library, where 13 eukaryotic genes were identified (e-score range: $9 \times 10^{-5}$ to $4 \times 10^{-29}$) with most (69%) annotated as a fungus or plant hypothetical proteins (Table A3.17). The remaining genes identified from both libraries were annotated as being of prokaryote origin (Table A3.16; Table A3.17). Though cDNA library construction targeted poly(A) tailed mRNA, to specifically select for eukaryotic mat community transcripts, both pyrosequencing datasets suggest bacterial transcripts from the mat community were also amplified from the poly-T primer used in library construction. This is not surprising given recent reports which estimate a large proportion of bacterial mRNA may be poly(A) tailed (19, 28).

Comparing Microarray Analyses with Protein Expression

Limited proteomics-based experiments were performed to identify major proteins potentially responding to UV irradiance, and to compare protein occurrence with relative expression levels in the above microarrays. One dimensional proteomic analysis of -UV and +UV algal mats collected in January 2010 at Lemonade Creek revealed identical protein banding patterns (data not shown), indicating no major regulatory changes at the

translational level in response to UV irradiance. Proteins with visually dominant expression in the 1-D gel were targeted for identification using tandem mass spectrometry (LC-MS/MS). Due to strong similarities in the +UV and -UV protein banding patterns, for downstream LC-MS/MS work only the +UV (OP-4 filter) mat samples were used in peptide identification. Twelve protein bands were excised (Fig. 3.7A) and identified based on peptide mass and fragmentation matches in the NCBInr database.

Several proteins exhibited multiple isoforms. Of the twelve bands excised and identified, four bore the same annotation as another band. As well, several excised bands identified more than one protein. Collectively, LC-MS/MS analysis led to the identification of eight abundant proteins (Table 3.3). Proteins with strong band presence were found to be involved in light-harvesting (*apcB*, *apcE*, *cpcA*, *cpcB*, *cpcG*), ATP synthesis (*atpB*), and carbon fixation (*rbcL*) related-metabolisms. The heat shock protein 70 (Hsp70) was also enriched in the proteome. Matching microarray analysis conducted at Lemonade Creek in January 2010 found no significant transcriptional response to UV exposure for the eight LC-MS/MS identified proteins. Fig. 3.7B illustrates this association by plotting the +UV:-UV expression ratio for the eight protein-encoding genes; none were up- or down-regulated ≥2-fold ($p$-value $< 0.05$), and thus in agreement with the lack of observed differences between the +UV and –UV protein banding patterns.

Table 3.3 Proteins Identified by LC-MS/MS

| Protein | Band # | Mass (Da) | Gene Name | KEGG Functional Group |
|---|---|---|---|---|
| Allophycocyanin beta chain | 1,8 | 17,340 | *apcB* | Photosynthesis-antenna proteins |
| Phycobilisome linker polypeptide | 7 | 96,440 | *apcE* | Photosynthesis-antenna proteins |
| Phycocyanin alpha chain | 1,2,5,8 | 17,270 | *cpcA* | Photosynthesis-antenna proteins |
| Phycocyanin beta chain | 1,3,9 | 18,250 | *cpcB* | Photosynthesis-antenna proteins |
| Phycobilisome rod-core linker polypeptide | 6 | 27,650 | *cpcG* | Photosynthesis-antenna proteins |
| ATP synthase CF1 beta chain | 10 | 50,410 | *atpB* | Photosynthesis |
| Ribulose bisphosphate carboxylase large chain | 4,10,11 | 54,130 | *rbcL* | Carbon Fixation |
| Heat shock protein Hsp70 | 12 | 71,530 | *Hsp70* | Endocytosis, Trasport and Catabolism |

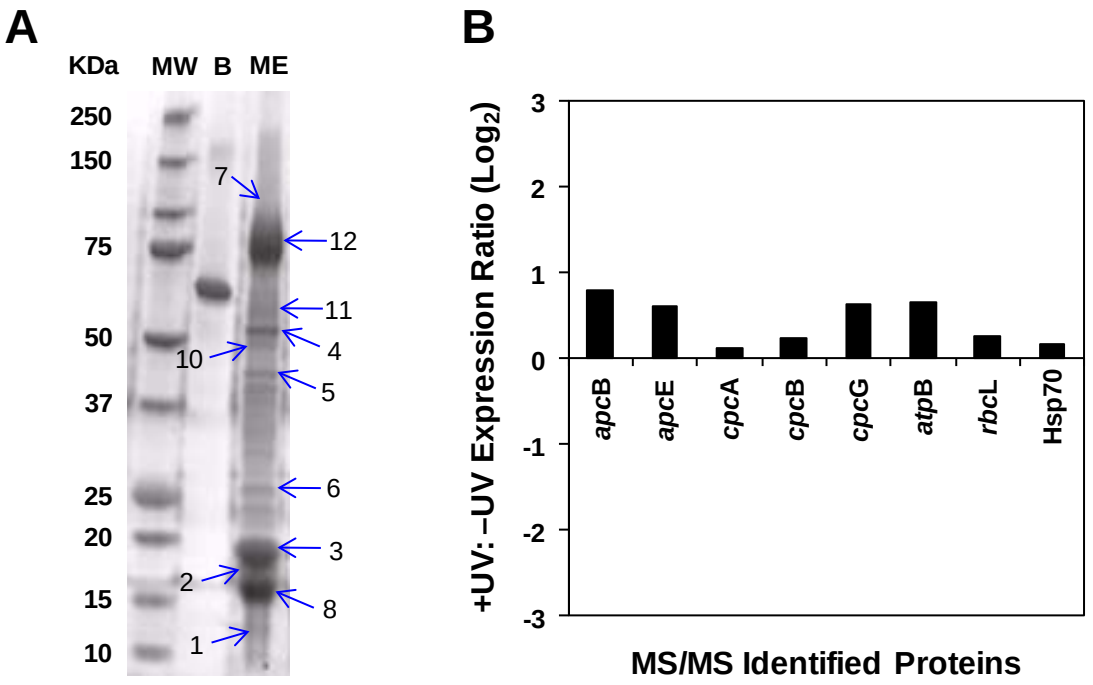


Figure 3.7 Limited comparison of Lemonade Creek proteome and transcriptome work conducted at Lemonade Creek in January 2010. **(A)** 1D SDS-PAGE of protein extracts from +UV cyanidial mats. MW, molecular weight standards; B, BSA standard; ME, cyanidial protein extracts. Protein bands targeted for LC-MS/MS work are indicated with arrows. **(B)** Relative January 2010 +UV:-UV transcriptional expression ratio ($Log_2$) for abundant cyanidial proteins identified via LC-MS/MS.

A separate assessment examined whether the dominantly expressed proteins in January 2010 were also abundantly transcribed, irrespective of UV-exposure. To do this, the January 2010 +UV and –UV microarrays were analyzed by assessing probe signal intensity for the eight abundant protein-encoding genes. For the –UV samples, signal intensities for the eight protein-encoding genes ranked in the top 40% among all genes with significant expression on the array. Intensity was also high for the +UV samples; here the eight genes ranked in the top 20% among all genes examined on the array. Results suggest that for both the +UV and -UV cyanidial mats the abundant LC-MS/MS identified proteins are also highly transcribed.

## Discussion

This study represents one of the first research efforts aimed at conducting genomics-enabled microbial ecology in an environment that is chemo-stat like in nature, yet that can be manipulated so as to reliably link *in situ* gene expression patterns with measurable environmental variables. The focus was to build on a previous investigation that suggested UV and VIS irradiance are keystone environmental factors controlling the health and activity of the Cyanidiales in their natural habitat (12). To date, the effects of *in situ* UV-VIS irradiance on algal gene expression have not been documented. In this study, genome-wide transcriptomics revealed a pronounced seasonality in terms of cyanidial response to both UV exposure as well as to UV-VIS intensity that appears to adjust across low and high irradiance time periods.

A systematic quantification of important physical factors at Lemonade Creek found significant changes in UV-VIS irradiance and water temperature (Fig. 3.1), and

that the intensity of these environmental physical factors increased during months of the year when cyanidial mat decline effects are strongest. Temperature is an important environmental selector for these algae, and could potentially contribute to mat decline (12). In this study, peak temperatures in July and August correspond to this annual event (Fig 3.1), and are near the upper temperature limit for the Cyanidiales (3). As such, temperature could represent a stress for these algae, however, direct regression analysis failed to identify any statistically significant relationship between temperature and the expression (positive or negative) of any of the 6,699 genes included in the microarray design. Likewise, a temporally intensive examination of spring aqueous chemistry revealed no changes that could be correlated with algal population dynamics (Table 3.1).

Solar irradiance levels also changed in an expected temporal fashion (Fig. 3.1), with UV and VIS irradiance intensities both increasing in step with the lengthening daily photoperiod. Experiments coupling UV irradiance filtering with whole-community transcription profiling provided evidence that seasonal UV exposure affects *in situ* cyanidial gene expression patterns. Initial analyses examined UV sensitivity by assessing the +UV:-UV expression ratios at four time points that differed with respect to UV irradiance intensity (Fig. 3.1). From this perspective, greatest UV sensitivity was observed in November 2009 and January 2010 (Fig. 3.3), periods of the year featuring the lowest UV levels and shortest photoperiods (Fig. 3.4). At the November 2009 time point, 78 genes exhibited statistically significant +UV:-UV expression ratios, with most being up-regulated (Fig. 3.3). In contrast, over 90% of the differential gene expression changes observed in January 2010 involved down-regulation in response to UV exposure (Fig.

3.3). Importantly, as the study progressed into summer and higher UV-VIS irradiance intensities, the frequency of statistically significant +UV:-UV expression ratios declined dramatically (Fig. 3.3), with only three instances of apparent UV-linked differential expression by July 2010 (Table A3.3). This suggested that either UV effects *per se* were no longer occurring, or that other environmental influences were now rendering the UV affects less obvious in this type of expression ratio analysis. The *k*-means clustering analysis was also based on +UV:-UV expression ratios, and also illustrated the most prominent UV effects in November 2009 and January 2010 (Fig. 3.5). Of the three distinct genes clusters formed, the majority (>95%) of genes in each cluster exhibited significant +UV:-UV expression ratios during either November 2009 or January 2010 (Table A3.6-A3.8). But by July 2010, UV irradiance-specific expression changes again appeared to be essentially absent (Fig. 3.5).

One of the most notable examples of genes being apparently repressed in response to UV exposure involved genes encoding for DNA repetitive elements (Table A3.6). Repetitive DNA elements are known to influence gene expression either indirectly by controlling chromatin structure or directly by acting as transcriptional activators or repressors (29, 31). Until the current study, however, extensive *in situ* transcriptional activity of repetitive elements has not been formally documented, nor were repetitive elements associated with UV-irradiance response. At this juncture, the potential regulatory linkage between UV and repetitive DNA elements is somewhat speculative, however the near wholesale response of these repetitive elements to UV is almost certainly not coincidental. Potentially, if their function in the Cyanidiales is that of a

repressor, then by repetitive elements being deactivated, genes involved in UV stress response could correspondingly be up-regulated. The latter is in agreement with earlier studies demonstrating that the down-regulation of repressor repetitive DNA elements following a thermal stress event resulted in the activation of several heat shock and chaperone-encoding genes important in conferring thermotolerance (29).

The above +UV:-UV expression ratio analysis was valuable for identifying significant UV effects (positive or negative) at the four specific time points, but the limitations of this kind of approach became evident in the spring and summer seasons, time points coinciding with a 10-fold increase in VIS irradiance. This likely began masking or counteracting UV-based gene regulation and was most apparent in July 2010, where the *k*-means analysis showed that the +UV:-UV expression ratios were neutral with respect to UV-response (Fig. 3.5). Consequently, separate analyses compared gene expression during the lowest (January 2010) and highest (July 2010) UV-VIS irradiance periods of the year, identifying genes that were being regulated by UV or VIS separately, or by the collective effects of UV+VIS.

Apparent combined UV+VIS response was evident, where 168 genes exhibited significant July:January expression ratios under both the UV+VIS exposed and VIS-only exposed algal mats (Table A3.2). Nearly 80% of the differential gene expression changes involved down-regulation, with most genes repressed to the same degree under both the UV+VIS and VIS exposure conditions (Table A3.11), suggesting the genes were actually being affected by the VIS wavelengths. Apparent combined UV-VIS response was only evident for genes annotated as repetitive DNA elements, where UV was found to

attenuate VIS repression affects (Table A3.11). Genes affected by VIS irradiance were involved in replication, repair, and growth-related metabolisms (Table 3.2; Table A3.11). This suggests high VIS dosages may be negatively effecting algal replication and growth, and thus could be a contributing factor to cyanidial mat decline. This interpretation is in agreement with Cyanidiales population dynamics and irradiance filtering experiments reported by Lehr et al. (2007), who noted an appreciable decline in cyanidial viability and photosynthetic activity associated with UV-VIS irradiance.

Gene regulation inferred to be exclusively in response to UV intensity saw a coordinated decrease in photosynthetic activity as seasonal UV intensified, including the repression of genes encoding for structural subunits of the phycobilisome light-harvesting antenna and photosystem II, and enzymes associated with chlorophyll metabolism (Table A3.12). These results expand on previous pure culture studies, which note the primary targets of UV-induced photodamage are the photosystems and accessory pigments (8, 9, 25). Importantly, these expression changes are occurring under *in situ* conditions, suggesting the expression profiles are likely ecologically relevant, and that environmental microarrays may be useful in identifying the mechanisms by which phototrophs respond to UV in nature.

Additional analysis sought to separate the VIS affects from UV affects. The UF-5 (i.e. –UV) filters blocked 99.9% of all UV wavelengths and thus allowed for a direct examination of VIS-only effects. A large number of genes associated with oxidative phosphorylation, including cytochromes B and C and theNADH ubiquinone oxidoreductase were specifically up-regulated in response to increased VIS intensity

(Table 3.2; Table 3.9). Studies of pure culture phototrophs have observed the activation of alternative energy production processes in the wake of UV-VIS related stress (8, 9, 25), noting that inhibition of photosynthesis typically is associated with the activation of cellular respiration. Though decreased photosynthetic activity was found to be a UV-specific response, the VIS-only activation of cellular respiration suggests the strategies employed by these algae to deal with both excess VIS and UV intensity are potentially integrated.

Further VIS-only work also utilized direct regression analysis. Expression of genes coding for key functions such as DNA replication, and for organelle and cell cycle division displayed statistically significant negative correlations ($r = -0.94$ to $-0.97$) with VIS intensity (Table A3.10). There was also a notable positive statistical relationship ($r = 0.87$ to $0.90$) between VIS intensity and the expression of several heat shock and chaperone genes, specifically: *dnaJ*, *tcp1*, *clpS*, and *pcmT* (Table A3.10), suggesting the algal cells were experiencing photodamage and activating protective and repair systems. Strong up-regulation of heat shock/chaperone genes in response to visible light-induced damage is in agreement with earlier studies demonstrating the induction of heat shock response following high intensity VIS exposure and damage to PSII (8, 30, 39).VIS dosage during the summer period may result in the activation repair mechanisms that function to protect PSII from VIS-induced photoinhibition.

Given the mRNA expression changes encountered under *in situ* conditions, environmental microarray data sets were assessed for their efficacy at capturing gene expression patterns occurring in natural microbial communities. This was examined in

two separate ways. First, the relative information yield of two independent transcriptome data sets was examined. Comparative analyses showed that environmental microarrays were far more effective at detecting transcript diversity and gene expression changes than transcriptome data sets developed via pyrosequencing (Fig. 3.6A). Minimal pyrosequencing coverage depth likely reduced the number of different gene transcripts detected (Fig. 3.6B), as well as the ability to resolve gene expression changes. This is consistent with previous metatranscriptomic analyses of marine bacterioplankton communities, which found pyrosequencing libraries indentified only a small subset of the total mRNA pool present, and failed to capture small expression differences (7, 32).

Second, algal mats were analyzed to evaluate the potential connectivity of the transcriptome to the proteome. Specifically, one dimensional proteomic analysis (Fig. 3.7A) in combination with limited LC-MS/MS analysis identified several dominant cyanidial proteins (Table 3.3) that displayed high probe signal intensity on matching environmental microarrays, suggesting there was good correspondence between highly abundance proteins and gene transcriptional activity for these algae. This interpretation is in agreement with transcriptome and proteome experiments reported by Trauger et al. (2008), who found a high degree of similarity in the mRNA and protein levels for genes altered in a hyperthermophilic bacterium following a thermal stress event. It must be noted, however, that proteomic work in this study was limited, and specifically focused on abundant proteins; recent work has suggested that low-abundance proteins and mRNAs are completely uncorrelated in their expression levels (34). Issues aside, the collective data suggests that complimentary information obtained by both proteomic and

transcriptomic techniques can be used to tease out ecologically-relevant molecular changes occurring in *in situ* microbial populations.

References

1. Albertano, P, Ciniglia, C, Pinto, G, and Pollio, A. 2000. The taxonomic position of *Cyanidium*, *Cyanidioschyzon*, and *Galdieria*: an update. Hydrobiologia. 433:137–43.

2. Asimgil, H, Kavakli, IH. 2004. Purification and characterization of five members of photolyase/cryptochrome family from *Cyanidioschyzon merolae*. Plant Science. *In Press*.

3. Brock, TD, and Darland, ML. 1970. Limits of microbial existence: temperature and pH. Science. 169:1316-1318.

4. Ciniglia, C, Yoon, HS, Pollio, A, Pinto, G, and Bhattacharya, D. 2004. Hidden biodiversity of the extremophilic Cyanidiales red algae. Molecular Ecology. 13:1827-1838.

5. Collen, J, Roeder, V, Rousvoal, S, Collin, O, and Kloareg, CB. 2006. An expressed sequence tag analysis of thallus and regenerating protoplasts of *Chondrus crispus* (gigartinales, rhodophyceae). Journal of Phycology. 42:104-112.

6. Fujiwara, T., Misumi, O., Tashiro, K., Yoshida, Y., Nishida, K., Yagisawa, F., Imamura, S., Yoshida, M., Mori, T., Tanaka, K., Kuroiwa, H., and Kuroiwa, T. 2009. Periodic gene expression patterns during the highly synchronized cell nucleus and organelle division cycles in the unicellular red alga Cyanidioschyzon merolae. DNA Res. 16:59–72.

7. Gifford, SM, Sharma, S, Rinta-Kanto, JM, and Moran, MA. 2011. Quantitative analysis of a deeply sequenced marine microbial metatranscriptome. The ISME Journal. 5:461-472.

8. Huang, L, McCluskey, MP, Ni, H, and LaRossa, R. 2002. 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(24): 6845-6858.

9. Kelly, DJA., Budd, K, Lefebvre, DD. 2007. Biotransformation of mercury in pH-stat cultures of eukaryotic freshwater algae. Arch Microbiol. 187:45-53.

10. Kim, DK, Hong, SJ, Bae, JH, Yim, N, Jin, ES, and Lee, CG. 2011. Transcriptomic analysis of *Haematococcus lacustris* during astaxanthin accumulation under high irradiance and nutrient starvation. Biotechnol.Bioprocess Eng. 16:398-705.

11. Kunin, V, Engelbrektson, A, Ochman, H, and Hugenholtz, P. 2010. Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. Environ. Microbiol. 12:118–123.

12. Lehr, CR, Frank, SD, Norris, TB, D'Imperio, S, Kalinin, AV, Toplin, JA, Castenholz, RW, and McDermott, TR. 2007. Cyanidia (Cyanidiales) population diversity and dynamics in an acid-sulfate-chloride spring in Yellowstone National Park. J. Phycol. 43:3–14.

13. Li, H, and Durbin R. 2009. Fast and accurate short read alignment with burrows-wheeler transform. Bioinformatics. 25:1754-60.

14. Li, Z, Wakao, S, Fischer, BB, and Niyogi, KK. 2009. Sensing and responding to excess light. Annu. Rev. Plant Biol. 60:239-260.

15. Mathur, J., Bizzoco, RW, Dean, EG, Lipson, DA., Poole, AW, Levine, R, and Kelly, ST. 2007. Effects of abiotic factors on the phylogenetic diversity of bacterial communities in acidic thermal springs. Appl. Environ. Microbiol. 73(8):2612-2623.

16. Matsuzaki, M, Misumi, O, Shin-I, T, Maruyama, S, Takahara, M, Miyagishima, SY, Mori, T, Nishida, K, Yagisawa, F, Nishida, K, Yoshida, Y, Nishimura, Y, Nakao, S, Kobayashi, T, Momoyama, Y, Higashiyama, T, Minoda, A, Sano, M, Nomoto, H, Oishi, K, Hayashi, H, Ohta, F, Nishizaka, S, Haga, S, Miura, S, Morishita, T, Kabeya, Y, Terasawa, K, Suzuki, Y, Ishii, Y, Asakawa, S, Takano, H, Ohta, N, Kuroiwa, H, Tanaka, K, Shimizu, N, Sugano, S, Sato, N, Nozaki, H, Ogasawara, N, Kohara, Y, and Kuroiwa, T. 2004. Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. Nature. 428:653–657.

17. Minoda, A, Nagasawa, K, Hanaoka, M, Horiuchi, M, Takahashi, H, and Tanaka, K. 2005. Microarray profiling of plastid gene expression in a unicellular red alga, *Cyanidioschyzon merolae*. Plant Molecular Biology. 59:375-385.

18. Minoda, A, Weber, APM, Tanaka, K, and Miyagishima, S. 2010. Nucleus-independent control of the rubisco operon by the plastid-encoded transcription factor Ycf30 in the red alga *Cyanidioschyzon merolae*. Plant. Physiol. 154:1532-1540.

19. Mohanty, BK, and Kushner, SR. 2010. Bacterial/archaeal/organellar polyadenylation. Wiley Interdisciplinary Reviews: RNA. 2:256-276.

20. Nagasaka, S, Nishizawa, NK, Satoshi, M, and Yoshimura, E. 2004. Metal metabolism in the red alga Cyanidium caldarium and its relationship to metal tolerance. BioMetals. 17:177-181.

21. Norris, TV, McDermott, TR, and Castenholz, RW. 2002. The long-term effects of UV exclusion on the microbial composition and photosynthetic competence of bacteria in hot-spring microbial mats. FEMS Microbiol. Rev. 39:193-209.

22. Novaes, E, Drost, DR, Farmerie, WG, Pappas Jr, GJ, Grattapaglia, D, Sederoff, R, and Kirst, M. 2008. High-throughput gene and SNP discovery in *Eucalyptus grandis*, an uncharacterized genome. BMC Genomics. 9:312.

23. Ohta N, Sato N, Kuroiwa T. 1998. Structure and organization of the mitochondrial genome of the unicellular red alga *Cyanidioschyzon merolae* deduced from the complete nucleotide sequence. Nucleic Acids Res. 26:5190–5298.

24. Ohta N, Matsuzaki, M, Misumi, O, Miyagishima, S, Nozaki, H, Tanaka, K, Shin-I, T, Kohara, Y, and Kuroiwa, T. 2003. Complete sequence and analysis of the plastid genome of the unicellular red alga *Cyanidioschyzon merolae*. DNA Res. 10:67–77.

25. Park, S, Jung, G, Hwang, Y, and Jin, E. 2010. Dynamic response of the transcriptome of a psychrophilic diatom, Chaetoceros neogracile, to high irradiance. Planta. 231:349-360.

26. Qin, J, Lehr, CR, Yuan, C, Le, X.C, McDermott, TR, and Rosen, BP. 2009. Biotransformation of arsenic by a Yellowstone thermoacidophilic eukaryotic alga. Proc. Natl. Acad. Sci. U.S.A. 106:5213-5217.

27. Sagi, M, Davydov, O, Orazova, S, Yesbergenova, Z, Ophir, R, Stratmann, JW, and Fluhr, R. 2004. Plant respiratory burst oxidase homologs impinge on wound responsiveness and development in *Lycopericon esculentum*. The Plant Cell. 16:616-628.

28. Saravanamuthu SS, Von Gotz F, Salunkhe P, Chozhavendan R, Geffers R, Buer J, Tummler B, and Steinmetz EJ. 2004. Evidence for polyadenylated mRNA in *Pseudomonas aeruginosa*. J Bacteriol. 186:7015–7018.

29. Schmidt, AL, and Anderson, LM. 2006. Repetitive DNA elements as mediators of genomic change in response to environmental cues. Biol. Rev. 81:531-543.

30. Schroda, M, Kropat, J, Rudiger, OW, Vallon, O, Wollman, FA, and Beck CF. 2001. Possible role for molecular chaperones in assembly and repair of photosystem II. Biochemical Society Transactions. 29:413-418.

31. Shapiro, JA, and Von Sternberg, R. 2005. Why repetitive DNA is essential to genome function. Biol. Rev. 80:227-250.

32. Stewart, FJ, Ottesen, EA, and DeLong, ED. 2010. Development and quantitative analyses of a universal rRNA-subtraction protocol for microbial metatranscritomics. The ISME Journal. 4:896-907.

33. Sturn, A, Quackenbush, J, and Trajanoski, Z. 2002. Genesis: cluster analysis of microarray data. Bioinformatics. 18:207-208.

34. Taniguchi, Y, Choi, PJ, Li, GW, Chen, H, Babu, M, Hearn, J, Emili, A, and Xie, XS. 2011. Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. Science. 329:533-537.

35. Trauger, SA, Kalisak, E, Kalisak, J, Morita, H, Weinberg, MV, Menon, AL, Poole, FL, Adams, MWW, and Siuzdak, G. 2008. Correlating the transcriptome, proteome and metabolome in the environmental adaptation of a hyperthermophile. Journal of Proteome Research. 7:1027-1035.

36. Turner, L, Houghton, JD, and Brown, SB. 1997. Purification and identification of apophycocyanin α and β subunits from soluble protein extracts of the red alga *Cyanidium caldarium*. Light exposure is not a prerequisite for biosynthesis of the protein moiety of this photosynthetic accessory pigment. Planta. 201:78-83.

37. Weber, AP, Oesterhelt, C, Gross, W, Brautigam, A, Imboden, LA, Krassovskaya, I, Linka, N, Truchina, J, Schneidereit, J, Voll, H, Voll, LM, Zimmermann, M, Jamai, A, Riekhof, WR, Yu, B, Garavito, RM, and Benning, C. 2004. EST-analysis of the thermo-acidophilic red microalga *Galdieria sulphuraria* reveals potential for lipid A biosynthesis and unveils the pathway of carbon export from rhodoplasts. Plant Mol. Biol. 55:17–32.

38. Wilson, KE, and Huner, PA. 2000. The role of growth rate, redox-state of the plastoquinone pool and the trans-thylakoid ΔpH in photoacclimation of *Chlorella vulgaris* to growth irradiance and temperature. Planta. 212:93-102.

39. Yokthongwattana, K, Chrost, B, Behrman, S, Casper-Lindley, C, and Melis, A. 2001. Photosystem II damage and repair cycle in the green alga, Dunaliella salina: involvement of a chloroplast-localized Hsp70. Plant Cell Physiol. 42:1389-1397.

40. Yoon, HS, Hackett, JD, and Bhattacharya, D. 2002. A single origin of the peridinin- and fucoxanthin- containing plastids in the dinoflagellates through tertiary endosymbiosis. Proc. Natl. Acad. Sci. U.S.A. 99(18):11724-11729.

41. Yoshimura, E, Nagasaka, S, Satake, K, and Mori, S. 2000. Mechanism of aluminum tolerance in *Cyanidium caldarium*. Hydrobiologia. 433:57-60.

# CHAPTER 4

# CYANIDIALES DIVERSITY IN YELLOWSTONE NATIONAL PARK

## Abstract

The Cyanidiales are an order of unicellular red algae that are unique among phototrophs in that they thrive in acidic and high temperature habitats associated with geothermal activity. Their taxonomy was originally based on morphological features, resulting in the classification of three formally recognized genera: *Cyanidioschyzon*, *Cyanidium*, and *Galdieria*. Much remains to be learned about the distribution and diversity of these algae in geothermal environments. In this study, culture-dependent work was combined with molecular surveys to examine Cyanidiales diversity and distribution in aqueous, soil and endolithic environments within the Yellowstone National Park (YNP) geothermal complex. Phylogenetic reconstruction using the *rbcL* gene cloned from throughout YNP supports the existence of at least two YNP lineages: *Cyanidioschyzon* and *Galdieria*, and shows the separation of taxa based on ecophysiological conditions. *Galdieria* populations were found exclusively in soil and endolithic habitats, while *Cyanidioschyzon* was the sole phylotype identified in aqueous environments, but was also detected in all soil and endolithic habitats investigated. The obligate photoautotroph *Cyanidioschyzon* is evidently able to thrive in non-aqueous habitats that are enhanced with respect to moisture and/or humidity. Interestingly, no *Cyanidium* phylotypes were detected in the surveys, suggesting that perhaps YNP's geothermal landscape is void of this genus. Finally, culture-dependent work examined

species distribution and abundance in non-aqueous environments in more detail, providing evidence that gravimetric soil moisture content is an important physical factor controlling cyanidial species composition.

## Introduction

The thermo-acidophilic, unicellular, and asexual red algae belonging to the order Cyanidiales are unique in that their environment is defined by acidic pH (0.5-3.5) and elevated temperature (38-56°C). These red algae are a dominant component of acidic geothermal environments and can be found in springs, soils, and endolithic (rock sub-surface) habitats throughout the world (4, 22, 23). Taxonomy was originally based on morphological differences, with the Cyanidiales simplistic structure leading to the identification of three different genera: *Cyanidioschyzon*, *Cyanidium*, and *Galdieria* (5, 12, 13). It has been suggested that due to limited morphological variation their genetic diversity is grossly underestimated, and that more novel species lineages remain to be discovered (5). This, along with many other aspects of cyanidial ecology remain poorly explored, presenting a high potential for new discoveries with regards to their distribution and diversity in varying geothermal locations.

One such Cyanidiales habitat is the Yellowstone National Park (YNP, Wyoming, USA) geothermal complex, a ~ 9,000 $km^2$ region with over 11,000 hydrothermal features that vary widely in pH, temperature, and chemistry (18). A variety of physical environmental factors could exert selection pressures on these algae, potentially creating distinct distribution patterns within YNP. Doemel and Brock (1971) suggested moisture content may act to control Cyanidiales species distribution, though currently there is

limited information to support this hypothesis. In general, pure culture isolates have been tested for their response to loss of water, covering a range of desiccation levels that exceed those encountered in the geothermal environment (20). Recently, Ciniglia et al. (2004) looked at biodiversity of these algae in soil and endolithic environments through the use of environmental polymerase chain reaction (PCR) at a hydrothermal system near Pisciarelli (Naples, Italy). Using the ribulose 6-phosphate carboxylase oxygenase (*rbcL*) gene as a marker for phylogenetic diversity, they were able to identify a new lineage of *Galdieria*, as well as detect genera not yet found in to the area they studied. Members of the *Galdieria* genus are known to inhabit a diverse array of microhabitats; their unique ability to grow heterotrophically on at least 50 different carbon sources affords them metabolic flexibility to thrive in environments with limiting light (e.g. 1% of daylight) and nutrients (11).

Previous YNP studies documented Cyanidiales diversity using a cultivation-dependent approach (26), however, it is widely acknowledged that cultivation bias could prevent the full algal population from being present in an isolate collection. In the study summarized herein, a combined molecular survey and cultivation-based approach was used to link cyanidial diversity, abundance, and distribution with influential environmental factors. Noting that much of the previous work in YNP has focused on cyanidial biodiversity in aqueous environments, work presented herein was less biased and included samples from soil, endolithic, and aqueous habitats. As well, using varying growth conditions known to select for specific Cyanidiales, the proportional occurrence

of specific genera was estimated from soil and endolithic samples, with subsequent *rbcL* gene sequencing determining the actual species composition.

## Materials and Methods

### Molecular Surveys: Site Description and Sample Collection

YNP Cyanidiales diversity was surveyed in endolithic, soil, and aqueous sites sampled from various thermo-acidic areas within YNP (Fig 4.1; Fig. 4.2). For the twenty-five sampling sites, global positioning coordinates, general collection location, type of environment, temperature, and pH are recorded in Table 4.1. Soil material was collected using autoclaved stainless steel spatulas; while large autoclaved bore pipet tips were used for acquiring aqueous mat samples (aqueous samples are defined here as those samples that are submerged under water). For endolithic samples, an autoclaved hammer was used to break off the upper rock layer in order to expose the underlying endolithic algae inside the rock matrix, which was then scraped and collected using a sterile spatula. Samples were transferred to sterile 50 mL Falcon$^{TM}$ tubes, flash frozen in a dry-ice/ethanol bath, and stored at -80$^{0}$C until used for DNA extraction.

### DNA Extraction, *rbcL* Cloning, Sequencing, and Phylogenetic Analysis

Total DNA was extracted from all samples as described previously (3). The nearly full-length ribulose 6-phosphate carboxylase oxygenase (*rbcL*) gene was amplified using protocols and the Cyanidiales-specific primers rbcL-90F and rCR (Table A4.1) previously described (5, 28). PCR products were purified using the QIAquick PCR

purification kit (Qiagen, Valencia, CA) and cloned into the pCR2.1–TOPO vector (Invitrogen, Carlsbad, CA). From each of the 25 samples collected, 10-12 clones (282 total) were sequence-screened to identify clones of interest for full-length sequencing, which was carried out using commercial vendors (High-Throughput Sequencing Solutions, Seattle, WA; Nevada Genomics Center, Reno, NV; or Molecular Research Core Facility, Pocatello, ID). Sequences were edited using Sequencher (Gene Codes Corporation, Ann Arbor, MI), and aligned using ClustalX (www-igbmc.u-strasbg.fr/BioInfo/ClustalX/Top.html). Phylogenetic analysis on aligned full-length sequences used the maximum-likelihood [ML] method in the PAUP software package (http://paup.csit.fsu.edu/about.html). Bootstrap values were generated using 100 pseudoreplicates in a heuristic search (starting with a random tree), and tree bisection-reconnection (TBR) branch-swapping algorithm was used to find the best ML tree. Accession numbers for the *rbcL* genes of previously sequenced rhodophytes used here for phylogenetic comparisons are listed in Table A4.2. Clone sequences described in this study are available as GenBank accession numbers JQ269605-JQ269638.

*Cyanidium*-specific PCR Primers and Amplification

In order to maximize the possibility of discovering potential YNP *Cyanidium* populations, *Cyanidium rbcL* gene sequences were specifically targeted for amplification. *Cyanidium* gene sequences available at the NCBI website fell into three main *rbcL* groups, with each amplifiable using one of the three reverse primers (MR-rbcLR, DBV182- rbcLR, or Sybil-rbcLR) in combination with the universal Cyanidiales forward primer rbcL-90F (Table A4.1). *Cyanidium*-specific PCRs were performed as described

above on DNA obtained from each molecular survey site. The annealing temperature was 51$^{o}$C for MR-rbcLR, 55$^{o}$C for DBV182- rbcL, and 45$^{o}$C for Sybil-rbcLR. A positive control from each of the three *Cyanidium* groups was included in each PCR to monitor PCR fidelity.

Cultivation Experiments

Soil and endolithic sites directly adjacent to Dragon Spring in Norris Geyser Basin (YNP, WY) were sampled and used to estimate the proportional occurrence of the different Cyanidiales genera in soils that varied naturally in moisture content. Environmental material was collected as described above, placed in sterile Falcon$^{TM}$ tubes, transferred to the lab, and prepared for total viable counts on media that would select for Cyanidiales capable of autotrophic (i.e. total viable algal counts) or heterotrophic growth. Samples were serially diluted (10-fold steps), and plated in triplicate onto solid Allen's media (1) amended with 0.5 mL of trace element solution at pH 3.5 and solidified with 20g/L Gelrite$^{TM}$ (Research Products International Corp., Mt. Prospect, IL). All autotrophic spread plates were incubated at 37$^{o}$C (to reduce moisture loss from the solid medium) under constant illumination from cool-white fluorescent lamps (80 $\mu E \cdot m^{-2} \cdot s^{-1}$). For heterotrophic growth, the Allen's agar was amended with 10 mM mannose (serving as the carbon source and electron donor) (11, 19) and the plates were incubated in continuous darkness at 37$^{o}$C.

In all cultivation work (and other, see below) viable cells counts were calculated for each site and media type using statistically significant colony counts (30-300 colonies/plate). Total direct counts were also performed using a Zeiss Axioskop 2

microscope (Zeiss, Jena, Germany) and a Petroff-Hauser counting chamber (Hausser Scientific, Horsham, PA). Colony and direct microscopic cell counts were normalized based on dry weight of the soil or rock material, which was determined by drying overnight at 65$^{o}$C, and weighing the following day to determine gravimetric moisture content of the original samples.

Species composition of the soil and endolithic samples was determined using both culture-dependent and culture-independent molecular methods. For the latter, environmental DNA extraction, *rbcL* amplification, and cloning methods were as described above. Thirty clones were selected and sequenced to estimate cyanidial species composition. In addition, species composition was approximated for cultivation work by performing colony PCRs on colonies arising on the heterotrophic Allen's media plates. Briefly, for each soil sample, thirty algal colonies (~1 mm diameter) were picked with autoclaved toothpicks and directly transferred to PCR tubes for use as a DNA template. The *rbcL* primer set and thermocycler conditions were identical to methods described above, with the addition of an initial 10 min. incubation at 94$^{o}$C to lyse the cells. Sequencing and phylogenetic analysis followed protocols described above.

Terbinafine Sensitivity Assay

During initial cultivation experiments, an acidophilic fungus caused occasional overgrowth problems, sometimes preventing accurate colony counts. Therefore, in another experimental thrust, the use of antifungal agents was investigated. In these experiments, dilution samples were plated onto autotrophic and heterotrophic Allen's media (described above) supplemented with terbinafine (5 µg/mL) or nystatin (10 U/mL).

The efficacy of using terbinafine as a selective agent was also examined among 21 pure culture isolates representing *Cyanidioschyzon merolae* (true morphotype, YNP Type IB), isolates that genotype as *Cyanidioschyzon* sp. but morphologically appear as *Galdieria* (YNP Type IA) (26), and YNP Type II and New Zealand Type IV *Galdieria* isolates (26). Isolates were obtained from the CCMEE culture collection at the University of Oregon (http://cultures.uoregon.edu), and tested for sensitivity by streaking in triplicate onto four different types of solid Allen's media: with or without terbinafine and in combination with or without mannose. Streak plates were then grown under autotrophic or heterotrophic culture conditions as described above. Some isolates displayed low/no growth rates on the solid media used, and so consequently were tested for sensitivity by cultivating in liquid Allen's media of the same composition, but lacking a solidifying agent. Incubation temperatures and the light regimen were the same, with terbinafine resistance scored as positive if optical density ($OD_{595}$) increased to > 0.1.

## Results

### Culture-independent Analysis of *rbcL* Phylogeny and Cyanidiales Species Distribution

Cyanidiales phylogenetic diversity was assessed at twenty-five sampling locations that varied with respect to temperature, pH, moisture content, geographic location, and/or habitat type (Fig 4.1; Fig. 4.2; Table 4.1). Sequence diversity at each site was quite low, with no evidence of temperature niche preference within a single sampling location (see samples DSB-DSH, Table 4.1). Consequently, a single representative *rbcL* clone from each identified cyanidial species was selected for near full-length (1215 nt) sequencing

and inference of phylogenetic relationships. From a total of 34 chimera-checked, full-length clones, Cyanidiales *rbcL* diversity throughout the Yellowstone geothermal complex resolved into two major lineages, *Galdieria* and *Cyanidioschyzon*, with strong bootstrap support (Fig. 4.3). Members of the largest group of clones, indentified as *Cyanidioschyzon*, were 99.3% identical to the *rbcL* of the fully-sequenced *C. merolae* strain 10D genome, and were the only Cyanidiales phylotype detected in aqueous habitats. *Cyanidioschyzon* sequences were also found in endolithic and soil environments (Table 4.1), though for the latter, moisture appeared to be important for colonization in that all of the soils sampled appeared either saturated or very wet (actual moisture content not determined). For the endolithic samples, rock materials were taken from locations immediately adjacent and above (< 2 cm) a hot spring outflow channel (e.g sample LCATERR, Table 4.1) or at the time of sampling were constantly bathed in steam (e.g. SFFL and SFFR, Table 4.1).

The other major lineage observed was *Galdieria*-A, which was found exclusively in soil and endolithic habitats, and never in any aqueous samples (Table 4.1). These YNP clones grouped together with *Galdieria sulphuraria* pure cultures isolates from YNP and Sonoma, California, but were distinct from other *Galdieria*-A clones previously derived from Italy, Russia, and Mexico (Fig. 4.3). Together, and as expected, these *Galdieria*-A clades branched separately from a *Galdieria*-B clade, which was comprised of *rbcL* alleles PCR-cloned from geothermal environments in Italy (Fig. 4.3). As with the *Cyanidioschyzon* representatives, there was little variation among the YNP *Galdieria*-A *rbcL* alleles. Finally, among a total of 282 partial and full-length clones, *Cyanidium* was

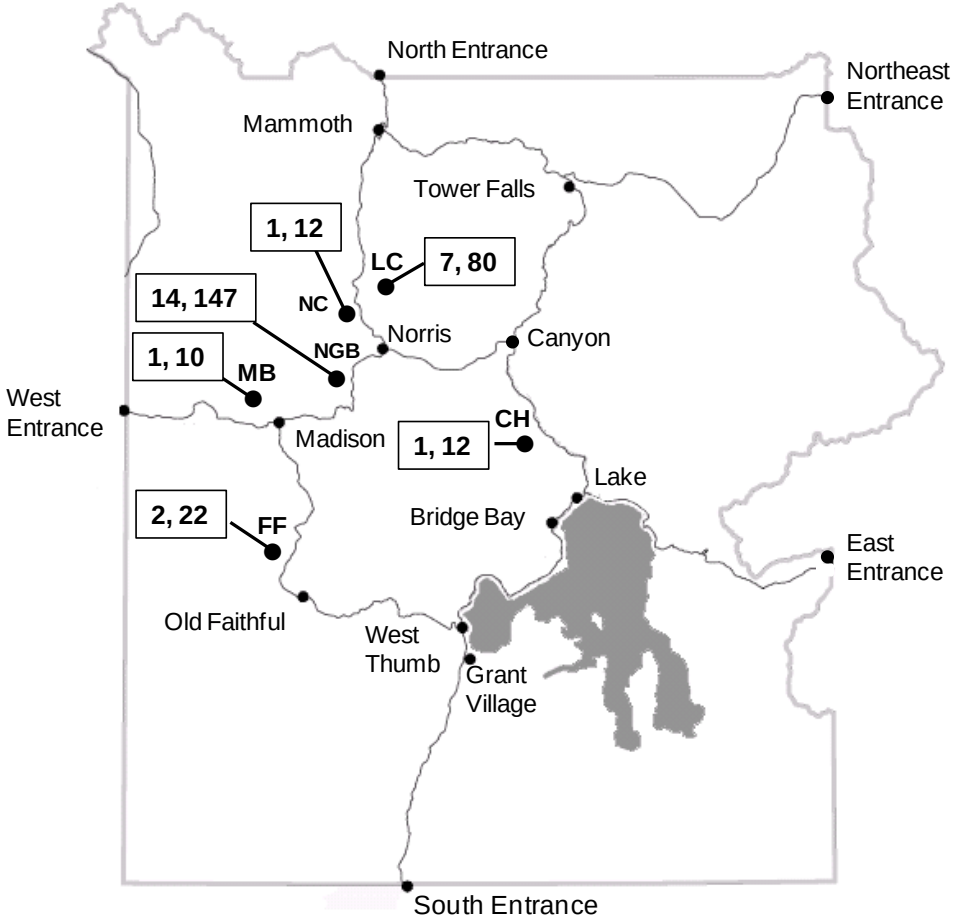


Figure 4.1 Map of YNP, with large black dots indicating major acidic geothermal areas sampled for cyanidial biodiversity: Fairy Falls (FF), Monument Basin (MB) Norris Geyser Basin (NGB), Nymph Creek (NC), Lemonade Creek (LC), and Crater Hills (CH). Specific geothermal area information is described in detail in Table 4.1. Boxes describe the number of sites sampled at each major geothermal area, and total number of clones obtained.

Table 4.1 Description of the sampling sites included in the molecular *rbcL* survey of YNP. In this table G-A is *Galdieria*-A, Cz is *Cyanidioschyzon*.

| | | | | | | | Cyanidiales clone identification | |
|---|---|---|---|---|---|---|---|---|
| Location | Location Code | Environment | Global Positioning | Temp (°C) | pH | Accession no.[a] | G-A | Cz |
| Crater Hills, YNP | CHJ | Acid Crust | 44°39'17"N, 110°28'54"W | 50 | - | JQ269634, JQ269635 | 7 | 5 |
| Dragon Spring, Norris Geyser Basin | DSB | Acid stream | 44°43'54"N, 110°42'39"W | 44 | 3.3 | JQ269617 | | 12 |
| Dragon Spring, Norris Geyser Basin | DSC | Acid stream | 44°43'54"N, 110°42'39"W | 49 | 3.3 | JQ269618 | | 11 |
| Dragon Spring, Norris Geyser Basin | DSD | Acid stream | 44°43'54"N, 110°42'39"W | 54 | 3.3 | JQ269605 | | 11 |
| Dragon Spring, Norris Geyser Basin | DSE | Acid stream | 44°43'54"N, 110°42'39"W | 48 | 3.3 | JQ269619 | | 12 |
| Dragon Spring, Norris Geyser Basin | DSF | Acid stream | 44°43'54"N, 110°42'39"W | 42 | 3.3 | JQ269620 | | 12 |
| Dragon Spring, Norris Geyser Basin | DSH | Acid stream | 44°43'54"N, 110°42'39"W | 34 | 3.3 | JQ269606 | | 12 |
| Dragon Spring, Norris Geyser Basin | DS1 | Acid stream edge | 44°43'54"N, 110°42'39"W | # | # | JQ269623, JQ269631 | 8 | 3 |
| Dragon Spring, Norris Geyser Basin | DS2 | Acid stream edge | 44°43'54"N, 110°42'39"W | # | # | JQ269629, JQ269638 | 6 | 4 |
| Dragon Spring, Norris Geyser Basin | DS3 | Acid stream edge | 44°43'54"N, 110°42'39"W | # | # | JQ269624, JQ269633 | 7 | 3 |
| Dragon Spring, Norris Geyser Basin | DS5 | Acid stream edge | 44°43'54"N, 110°42'39"W | # | # | JQ269625, JQ269632 | 9 | 1 |
| Fairy Falls Trailhead, YNP | SFFL | Endolithic | 44°30'5"N, 110°49'54"W | - | - | JQ269615, JQ269630 | 5 | 7 |
| Fairy Falls Trailhead, YNP | SFFR | Endolithic | 44°30'5"N, 110°49'54"W | - | - | JQ269616 | | 10 |
| Lemonade Creek, YNP | LCASUB | Acid stream | 44°48'4"N, 110°43'44"W | 53 | 3.1 | JQ269627 | | 12 |
| Lemonade Creek, YNP | LCATERR | Acid stream edge | 44°48'4"N, 110°43'44"W | 31 | 3.1 | JQ269607, JQ269608 | 4 | 7 |
| Lemonade Creek, YNP | LCBCEL | Acid stream edge | 44°48'4"N, 110°43'44"W | 39 | 3.1 | JQ269609, JQ269610 | 1 | 11 |
| Lemonade Creek, YNP | LCBSUB | Acid stream | 44°48'4"N, 110°43'44"W | 56 | 3.1 | JQ269628 | | 10 |

# Located immediately above (<2 cm) the Dragon Spring source water outlet

[a] Genbank accession number of representative clones

**Table 4.1**-*Continued*

| Location | Location Code | Environment | Global Positioning | Temp (°C) | pH | Accession no.[a] | Cyanidiales clone identification G-A | Cyanidiales clone identification Cz |
|---|---|---|---|---|---|---|---|---|
| Lemonade Creek, YNP | LCBTERR | Acid stream edge | 44°48'4"N, 110°43'44"W | 30 | 3.1 | JQ269611, JQ269612 | 2 | 10 |
| Lemonade Creek, YNP | LCCBLGR | Acid stream | 44°48'4"N, 110°43'44"W | 38 | 3.1 | JQ269613 | | 11 |
| Lemonade Creek, YNP | LCCYEGR | Acid stream | 44°48'4"N, 110°43'44"W | 47 | 3.1 | JQ269614 | | 12 |
| Nymph Creek, YNP | NCB | Acid stream edge | 44°45'11"N, 110°43'27"W | 42 | 3.2 | JQ269621 | | 12 |
| River Group, Monument Basin, YNP | RIVER1B | Acid Crust | 44°41'3"N, 110°45'14"W | 45 | - | JQ269636 | | 10 |
| Succession Spring, Norris Geyser Basin | SSI | Acid stream | 44°43'46"N, 110°42'47"W | 50 | - | JQ269622 | | 12 |
| Succession Spring, Norris Geyser Basin | SSII | Acid stream | 44°43'46"N, 110°42'47"W | 50 | - | JQ269626 | | 11 |
| Twin Springs, Norris Geyser Basin | TS | Acid stream | 44°43'36"N, 110°42'33"W | 42 | - | JQ269637 | | 12 |

# Located immediately above (<2 cm) the Dragon Spring source water outlet

[a] Genbank accession number of representative clones

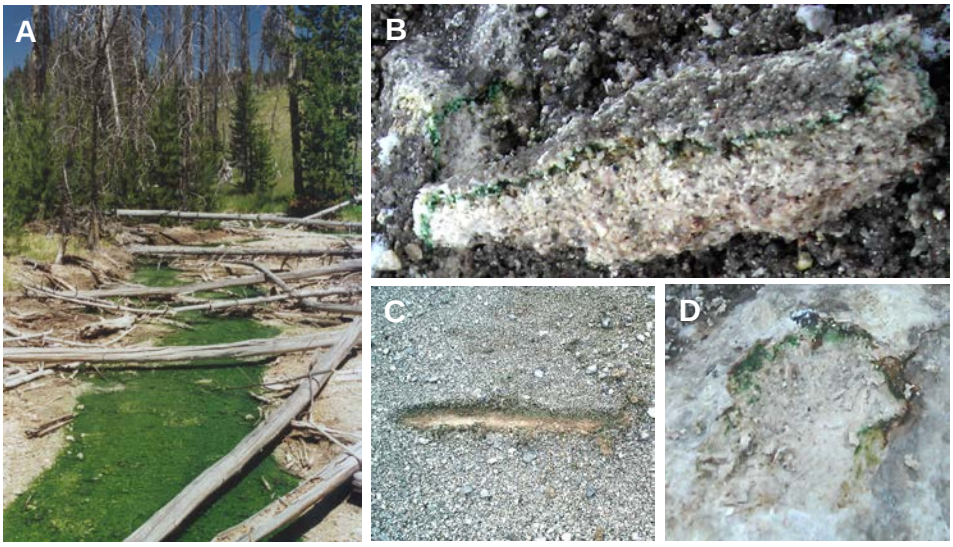


Figure 4.2 Photographs of cyanidial habitats within YNP. The diversity of sampled habitats included: (A) aqueous mats, (B) endolithic algal biomats growing inside loose, crumbly rock fabric, (C) thermally-influenced sandy soils, and (D) endolithic algal cells growing directly beneath the rock surface.

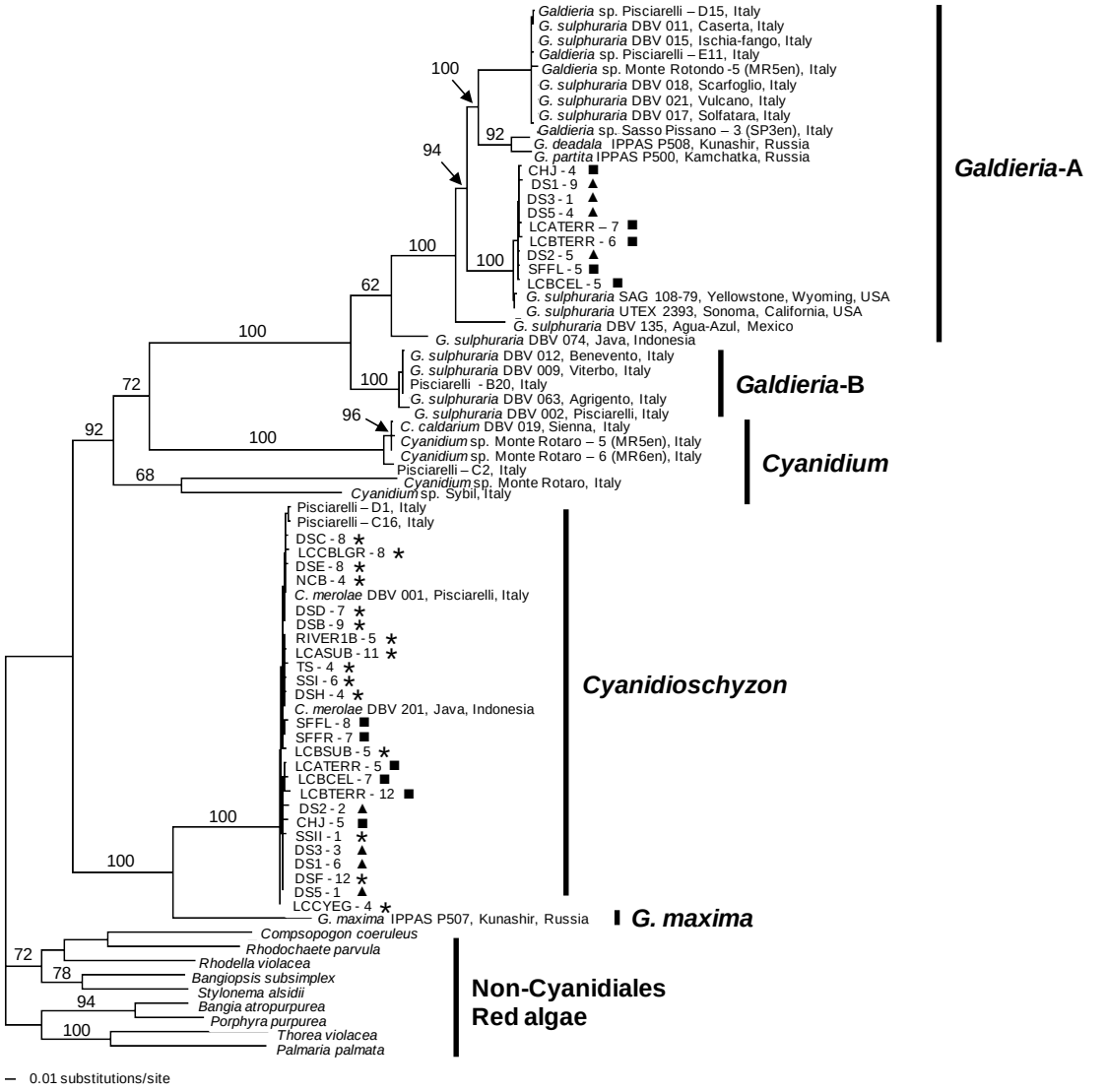


Figure 4.3 Phylogeny of the YNP Cyanidiales inferred from a maximum likelihood (ML) analysis based on *rbcL* gene sequences using PAUP. Only bootstrap values > 60% are shown. Sequences from the environmental survey are marked to denote habitat type, specifically: aqueous (asterisks), endolithic (■), and soil (▲) sample sites.

absent. To investigate this issue further, additional PCRs were conducted using *Cyanidium*-specific *rbcL* forward primers with the same DNA extracts from all 25 sampling locations; all PCRs were negative for *Cyanidium rbcL* sequences.

Cultivation Experiments

The above culture-independent molecular survey indicated that while *Galdieria* phylotypes were never found in aqueous environments, both *Cyanidioschyzon* and *Galdieria* were encountered in soil and endolithic habitats. Consequently, cultivation-dependent experiments were conducted to better understand Cyanidiales diversity and proportional occurrence in these non-aqueous environments. Previous studies have suggested that moisture content is an important environmental parameter determining cyanidial niche separation (5, 20, 28), and thus sampling efforts in this aspect of the study focused on Dragon Spring (Fig. 4.4) located in the Norris Geyser Basin. The site was selected because it represents all known Cyanidiales habitats in a very centralized location. Dragon Spring itself is an acidic hot spring feature with source water pH of 3.1 and temperatures ranging from 68-72$^{o}$C, though the Cyanidiales in Dragon Spring are located in the lower temperature reaches of the outflow channel (Table 4.1). Dragon Spring is surrounded by heated acidic soils (Fig. 4.2C) and endolithic habitats (Fig. 4.2B, Fig. 4.2D) that are colonized by the Cyanidiales. The aqueous, soil and endolithic environments are located within 6.4 m of each other, and thus there should be no long distance geographical barriers constraining the colonization of *Galdieria* and *Cyanidioschyzon*. Rather, distribution patterns should provide inferential information reflecting niche preferences. Furthermore, the site provided a naturally occurring transect with respect to soil moisture i.e. with increasing distance (uphill) from the edge of the spring outflow channel, soil moisture content decreased. This provided a reasonable sampling transect for directly examining soil moisture effects with respect to

*Cyanidioschyzon* and *Galdieria* habitat preference. All soils were compositionally comprised of coarse sand with small amounts of unidentified organic debris. Soils with greatest moisture content were found directly adjacent to the edge of the Dragon Spring outflow channel and were visually darker than the drier, lighter colored soils found uphill from the spring (Fig. 4.4). Moisture content of the endolithic materials could not be ascertained or predicted *a priori*, but this entire area is quite acidic, with pH ranging from 2.1-2.7 and gravimetric moisture content ranging from 8-31% (Table 4.2).

An examination of cyanidial cell counts in the sampled soil habitats determined that viable autotrophic counts ranged from $1.87 \times 10^3$ to $2.02 \times 10^5$ cells per gram of dry soil, and are interpreted here as total Cyanidiales viable counts (Fig. 4.5A). Heterotrophic colony counts ranged from $3.83 \times 10^3$ to $2.09 \times 10^5$ cells per gram of dry soil and were used to estimate the proportional occurrence of viable *Galdieria* (capable of growth on organic carbon substrates) at each soil sample site (Fig. 4.5A). Both autotrophic and heterotrophic viable counts appeared related to soil moisture (Fig. 4.5A). Non-linear regression analysis (a one-variable model, R software version 2.12.1, http://www.r-project.org), established this relationship was statistically significant, finding that gravimetric soil moisture content accounted for 97.7% of the variance in cyanidial viable cell density (p-value = $3.71 \times 10^{-6}$) (Fig. 4.5A). Similar analysis of direct microscopic counts were in agreement, also showing a strong relationship ($r^2 = 0.965$, p-value = 0.02) between gravimetric moisture content and the direct counts measured for all four soil sites (Fig. 4.5B). For all soil samples there was, however, a significant discrepancy between total direct counts and viable plate counts, percent viability decreased by over an

order of magnitude from the high moisture soils (20-23% viability, sample sites A and B) to the low moisture soils (1-2% viability, sample sites E and F) (compare Fig. 4.5A and Fig. 4.5B), implying that overall cell viability was low in the moisture stressed environments.

For most sampled soil sites, autotrophic and heterotrophic colony counts were not statistically significantly different, implying the soil populations were somewhat homogeneous and likely dominated by *Galdieria*. PCR cloning of Cyanidiales *rbcL* from DNA extracted from all soils supported this assessment; 100% of the 120 clones (30 from each soil) were found to be *Galdieria* (Table 4.2). Soil sample F (8% gravimetric water content) differed somewhat however. Here, heterotrophic counts ($3.83 \times 10^3$ cells per gram of dry soil) were roughly 2-fold that of autotrophic counts ($1.87 \times 10^3$ cells per gram of dry soil), indicating the possibility that the algal composition was more heterogeneous with respect to culturability under autotrophic conditions, or that this soil sample contained *Galdieria* that could not grow autotrophically. The latter possibility was pursued by sub-culturing > 100 colonies from different dilutions to autotrophic media and incubating under light. All sub-cultures grew autotrophically.

In contrast to the soil samples, which exhibited a two order of magnitude range in viable counts across the span of soil moistures, the relationship linking gravimetric moisture content with algal viability did not hold for endolithic environmental samples (Fig. 4.5A, samples C and D). Though moisture content of the material scraped from the rock interior varied from 14% to 24% (gravimetric moisture content), autotrophic and heterotrophic counts were essentially identical (differed by less than 6%) ranging from

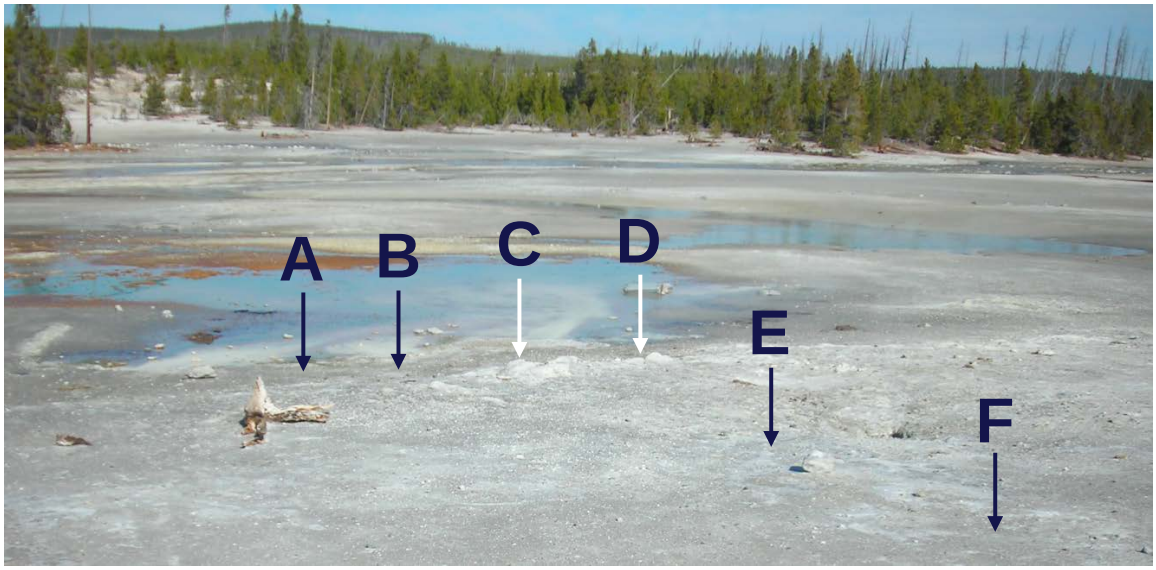


Fig 4.4 Image of Dragon Spring and surrounding thermal areas. Vertical arrows indicate sampling sites used for viable cell counts. Samples were collected during the 2009 field season from soil (black arrows) or endolithic (white arrow) habitats.

Table 4.2 Environmental features and Cyanidiales species composition of soil and endolithic sites investigated for the cultivation experiments.

| | | | | | Proportional clone composition (%)* | |
|---|---|---|---|---|---|---|
| Site | Description | Distance from Edge of Dragon Spring (m) | pH | Soil Moisture Content | *Cyanidioschyzon* | *Galdieria* |
| A | Soil | 0.8 | 2.7 | 30% | | 100 |
| B | Soil | 0.8 | 2.6 | 31% | | 100 |
| C | Endolithic | 1.5 | 2.1 | 24% | 90 | 10 |
| D | Endolithic | 1.5 | 2.7 | 14% | 13 | 87 |
| E | Soil | 4.9 | 2.1 | 18% | | 100 |
| F | Soil | 6.4 | 2.7 | 8% | | 100 |

* A total of 30 PCR amplified *rbcL* clones were sequenced for each soil or endolithic site

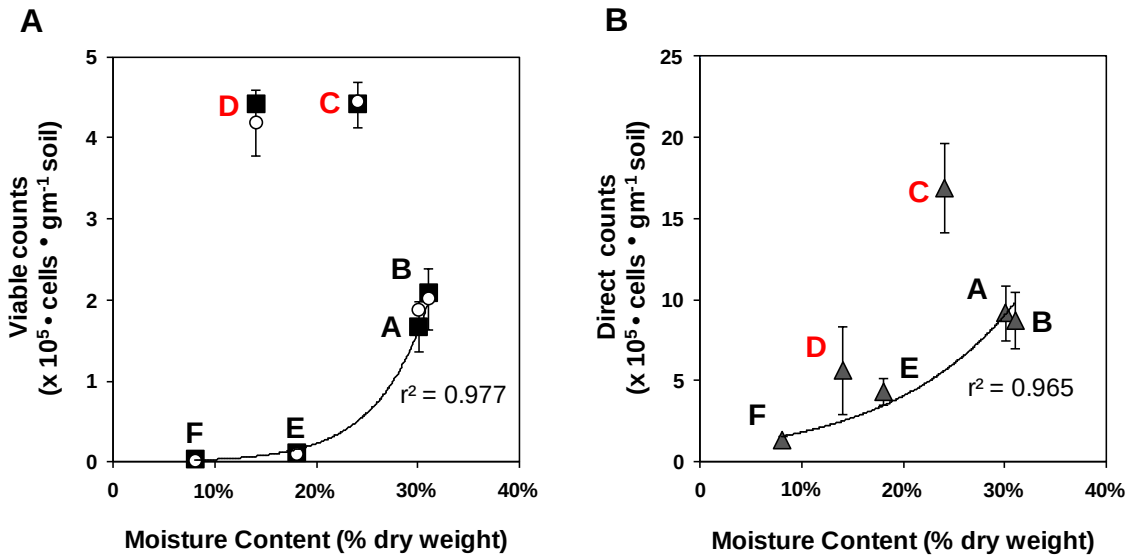


Figure 4.5 Direct, viable autotrophic, and viable heterotrophic cell counts of cyanidial populations as a function of moisture content in non-aqueous environments. (A) Viable cell counts in soil (black letters) and endolithic (red letters) habitats. Colony counts were determined after 14 days of incubation at 37°C under autotrophic (circles) and heterotrophic (squares) culture conditions from samples collected in non-aqueous habitats surrounding Dragon Spring. (B) Direct cell counts (triangles) in soil (black letters) and endolithic (red letters) habitats. For both panels direct, total viable, and heterotrophic cell counts are expressed as cells per gram dry weight of soil or endolithic rock material. Data points are labeled according to sample site and represent the mean of three replicates, with error bars (where visible) representing 1 standard deviation. Solid lines indicate best-fit exponential function of measured data; $r^2$ value for the line is given. Note the difference in the y-axis scale for panels A and B.

4.19 x $10^5$ to 4.45 x $10^5$ cells per gram of dry endolithic material (Fig. 4.5A). Note also that the endolithic viable algal counts were roughly two-fold that of soil counts (statistically greater, p-value < 0.05).

PCR *rbcL* clone sampling found *Cyanidioschyzon* in both endolithic samples; 90% in site C and 13% in site D (30 clones each) (Table 4.2). However, as discussed above, the autotrophic and heterotrophic counts for endolithic materials were essentially the same for both samples (Fig. 4.5A), indicating that there were *Cyanidioschyzon*

present in these samples that would not grow on the Allen's media used in this work; this was particularly apparent for site C. As determined by microscopic counts, site C contained roughly 2-fold greater algal cells than any of the soils and 3-fold greater than the other endolithic sample, site D (Fig. 4.5A). However, percent viability for site C was only 25% of the total counts as compared to approximately 100% for site D (compare Fig. 4.5A and Fig. 4.5B). This was in spite of the fact that moisture levels in the site D sample were half that for site C.

Terbinafine Sensitivity

While not always a problem, some cultivation-based work encountered occasional difficulties with a fast-growing fungus that overgrew the solid media plates, making colony counts at times impossible. To control fungal growth, two antifungal agents were tested, nystatin and terbinafine. Terbinafine proved to be effective in controlling the fungal contaminant, although initial work suggested this ergosterol inhibitor also inhibited the growth of *Cyanidioschyzon* sp., and selected exclusively for growth of *Galdieria* sp. (results not shown). The utility and selectiveness of this antifungal agent was further examined by determining the effect of terbinafine on twenty-one *Cyanidioschyzon* or *Galdieria* pure cultures isolates that originated from geothermal areas in either YNP or New Zealand (NZ). The *Galdieria* isolates were the NZ Type IV and the YNP Type II (26), and the *Cyanidioschyzon* isolates included the Type IA and Type IB as defined by Toplin et al. (2008). Type IA isolates exhibit >99% identity to the formally recognized *Cyanidioschyzon merolae* strain 10D *rbcL* and 18s rDNA genes, yet display morphological and cell division traits of *Galdieria* (26). By contrast, YNP

Table 4.3 Antifungal activity of autotrophically and heterotrophically grown *Galdieria* and *Cyanidioschyzon* pure-culture isolates derived from Yellowstone National Park and New Zealand. Isolates were tested for resistance against the antifungal agent terbinafine. + Growth, - No Growth.

| Strain ID | Growth Conditions | | | |
|---|---|---|---|---|
| | Autotrophic | | Heterotrophic | |
| | Autotrophic | Autotrophic + Terbinafine | Mannose | Mannose + Terbinafine |
| **YNP Type IB (*Cyanidioschyzon merolae*)** | | | | |
| 5610 | + | - | - | - |
| **YNP Type IA (*Cyanidioschyzon* sp.)** | | | | |
| 5506 | + | - | - | - |
| 5507 | + | - | - | - |
| 5508 | + | - | - | - |
| 5576 | + | - | - | - |
| 5584 | + | - | - | - |
| 5585 | + | - | - | - |
| 5609 | + | - | - | - |
| 5625 | + | - | - | - |
| 5631 | + | - | - | - |
| 5639 | + | - | - | - |
| 5640 | + | - | - | - |
| **YNP Type II (*Galdieria* sp.)** | | | | |
| 5511 | + | + | + | + |
| 5572 | + | + | + | + |
| 5573 | + | + | + | + |
| **NZ Type IV (*Galdieria* sp.)** | | | | |
| 5706 | + | + | + | + |
| 5707 | + | + | + | + |
| 5708 | + | + | + | + |
| 5710 | + | + | + | + |
| 5711 | + | + | + | + |
| 5717 | + | + | + | + |

*Cyanidioschyzon* Type IB shares genetic identity, morphology, and cell division features as for *C. merolae* strain 10D (26). When the various *Cyanidioschyzon* and *Galdieria* isolates were examined for their ability to grow autotrophically in the presence of

terbinafine, all *Galdieria* isolates were found to be resistant to the antifungal agent (Table 4.3), though none of the Type IA and IB *Cyanidioschyzon* isolates were resistant. Due to their enigmatic cellular features, Type IA isolates were also screened to assess metabolic capabilities. None grew heterotrophically on mannose (Table 4.3) or glucose (results not shown), which agrees with their classification as *Cyanidioschyzon*, which is currently viewed to be an obligate photoautotroph. The same results were obtained with the Type IB *Cyanidioschyzon* YNP isolate (Table 4.3). This screen confirmed that both types of *Cyanidioschyzon* isolates used in this study were incapable of heterotrophic growth on the tested carbon sources. Other preliminary screens of three YNP (Type II) and six NZ (Type IV) *Galdieria* isolates confirmed their ability to grow heterotrophically on mannose (Table 4.3) and glucose (results not shown).

Given the above demonstration that identified *Galdieria*, but not *Cyanidioschyzon*, as being resistant to this anti-fungal agent, a terbinafine screen was also incorporated into the above cultivation work. All tested colonies (100 colonies from each soil and endolithic sample) arising on the heterotrophic Allen's media plates were also resistant to terbinafine (results not shown). Further, all tested colonies arising on the autotrophic plates containing terbinafine were identified as *Galdieria* (120 colonies total).

## Discussion

This study represents an in-depth effort to improve the understanding of Cyanidiales biodiversity and distribution in YNP, the world's largest and most diverse geothermal complex. The focus was to systematically examine spatially disparate sites

and variations in environmental physical factors and their potential influence on cyanidial distribution and population structure.

The YNP phylogenetic framework established in this study expands on previous culture-dependent work conducted by Toplin et al. (2008), who used pure culture Cyanidiales isolates derived from acidic habitats to determine species composition in YNP. The current study also complements similar studies centered on the geothermal systems near Tuscany and Naples, Italy, which documented cyanidial species separation based on geographic localization and habitat preference (5, 28). Until the current study, the relationship between gravimetric soil moisture content and *in situ* Cyanidiales assemblages had not been formally documented. Based on data collected, results in the current study indicate that in acidic soil habitats, moisture availability is an important factor determining cyanidial abundance and distribution. Other factor(s) appear to influence populations in endolithic habitats.

Phylogenetic analysis of the YNP *rbcL* gene sequence identified two well-supported YNP lineages: *Galdieria*-A and *Cyanidioschyzon* (Fig. 4.3). YNP *Galdieria* environmental clones were closely related to *G. sulphuraria* and were localized exclusively to non-aqueous habitats (Table 4.1). Previous phylogenetic analysis found that *Galdieria* sequences isolated from the continental United States grouped within *Galdieria*-A, a globally dispersed Cyanidiales subclade (5, 28). When analyzed with the YNP environmental clones, pure cultures *G. sulphuraria* (UTEX 2393) isolated from Sonoma, California and *G. sulphuraria* (SAG 107.79) isolated from YNP, formed a distinct United States subclade, which was separate from the other *Galdieria*-A

phylotypes, the latter of which includes the *G. sulphuraria* DBV 135, an Agua-Azul culture isolated from Mexico (ML-bootstrap = 94%)(Fig. 4.3). These findings have ramifications with regards to the global distribution of *Galdieria*-A, suggesting dispersal events are extremely infrequent, and that given enough time isolated *Galdieria* populations could potentially genetically differentiate to form geographically-distinct clades (17).

An interesting feature of Cyanidiales assemblages in YNP endolithic and soil habitats was the large proportional occurrence of *Cyanidioschyzon*. All endolithic and soil sites sampled for the molecular survey were comprised of two species, *Galdieria*-A and *Cyanidioschyzon*; the latter accounted for 42-100% of the total clone abundance in endolithic environments and 10-58% of the abundance in sampled soil habitats (Fig. 4.3; Table 4.1). The localization of this genus to non-aqueous environments serves to validate earlier findings of endolithic *Cyanidioschyzon* populations (26, 27, 28). More importantly, the widespread co-occurrence, and in some cases dominance, of *Cyanidioschyzon* populations in non-aqueous environments demonstrates that the previous findings (26, 27, 28) were not coincidental. Instead, *Cyanidioschyzon* distribution in geothermal environments is more widespread than previously thought.

In contrast to soil and endolithic habitats, *Cyanidioschyzon* was the only phylotype detected from more than 170 *rbcL* clones sequenced from seven geographically distinct aqueous sampling sites (Fig. 4.1; Table 4.1). *Cyanidioschyzon* clones shared 99.3% sequence identity to the *rbcL* gene sequence from *C. merolae* strain 10D. The recent Toplin et al. (2008) study reported the widespread occurrence of a YNP

Cyanidiales isolate named Type IA, which constituted ~95% of a large isolate collection (>120 pure cultures). The enigmatic cellular features of Type IA isolates, where morphologically they resemble *Galdieria*, yet genetically show high sequence identity to *C. merolae*, make them intriguing from an evolutionary standpoint. Though algal cell morphology was not a focus of this study, based on the dominance of YNP Type IA isolates in the previous work, it is reasonable to assume many-to-most of the *Cyanidioschyzon* populations detected here were are most likely Type IA morphotypes.

The geographic and habitat sampling scheme implemented here failed to detect any YNP *Cyanidium* populations, including PCR amplification attempts using *Cyanidium*-specific *rbcL* primers. Studies of hydrothermal systems in Italy have reported the widespread occurrence and facile detection of *Cyanidium* phylotypes in endolithic derived clone libraries (28). Thus far, studies have failed to demonstrate distinct global distribution patterns of *Cyanidium* phylogenetic clades, suggesting dispersal may not be limiting for this taxa (28). Despite these findings, analysis in the current study failed to detect any *Cyanidium* phylotypes in YNP. This consistent with previous 18s rDNA and *rbcL* YNP molecular survey work, as well as phylogenetic analysis from an extensive YNP culture collection (9, 15, 26). The collective data from the current work and these previous studies suggests *Cyanidium* is either absent or present at very low levels in the YNP geothermal complex.

Given the co-occurrence of *Cyanidioschyzon* and *Galdieria* observed in non-aqueous habitats, culture-dependent analysis was used to assess major physical factors shaping Cyanidiales assemblages in these environments. Moisture availability is a

stressor known to affect the survival of terrestrial photosynthetic organisms (2, 25), and has been suggested to influence cyanidial viability (7). In this study, samples collected and cultivated from along a defined moisture transect near Dragon Spring assessed Cyanidiales proportional occurrence in non-aqueous environments. Both autotrophic and heterotrophic viable cell counts as well as total microscopic counts all showed a strong correlation to soil moisture level (Fig. 4.5A; Fig. 4.5B), demonstrating increased colonization and viability in relation to increased gravimetric soil moisture content. As well, for most soil samples, autotrophic and heterotrophic viable counts were nearly identical; implying the Cyanidiales were homogeneous with respect to their culturability, and likely dominated by *Galdieria* populations. An exception to this trend was sample site F, a soil habitat with the lowest sampled gravimetric moisture content (8%); here, heterotrophic plate counts were roughly 2-fold that of the autotrophic counts. This particular observation is difficult to explain, but perhaps suggests a large proportion of the *Galdieria* at Site F were particularly sensitive to high light conditions, displaying reduced growth rates when cultured autotrophically, and consequently not included in the autotrophic viable counts. Similar observations have been noted in *Galdieria* pure culture work, where the growth rate of several *G. sulphuraria* strains was significantly reduced when cultured at light intensities of 200 $\mu E \cdot m^{-2} \cdot s^{-2}$ (19, 24). In contrast to the soil samples, viability rates in the endolithic habitats sampled did not co-vary with gravimetric moisture content (Fig 4.5A). A recent study has suggested endolithic cyanidial populations are instead influenced humidity (5). This seems reasonable given

the otherwise strong relationship between moisture and viable counts in soils (Fig. 4.5A; Fig. 4.5B).

Moisture availability in soil habitats was further linked to *in situ* cyanidial abundance through measurements of total direct counts. As with the viable counts, total cell counts increased in relation to increased gravimetric soil moisture content (Fig. 4.5B). Results did, however, reveal a significant discrepancy between total and viable cell numbers in sampled soil sites, with low moisture soil habitats displaying only 1-2% viability (Fig. 4.5A; Fig. 4.5B). This finding suggests that cell viability in moisture stressed environments is low, and thus total direct counts grossly overestimate the physiologically active algal populations in such soils. In the endolithic environment, total algal direct counts recorded and plotted separately from those obtained for the soil samples, though the data suggest the drier site D sample was more similar to soils (Fig. 4.5B). Collectively, the data infers that moisture *per se* is not a limiting factor for algae inhabiting an endolithic environment.

Experiments in this study also used the gene sequence variation of *rbcL* as a convenient tool for looking at Cyanidiales species composition in the soil and endolithic culturing samples. PCR cloning of the *rbcL* gene sequence from these particular soils implied that *Galdieria* was clearly the dominant Cyanidiales species, and was in agreement with the autotrophic and heterotrophic plate counts, which did not differ and which in turn implied that all isolates were *Galdieria* (Fig. 4.5A). The same phylogenetic composition analysis of the endolithic habitats, however, found that both endolithic environments contained *Cyanidioschyzon*; site C in particular (Table 4.2). This contrasts

the composition predicted by the viable counts, which found the autotrophic and heterotrophic counts to be essentially identical, predicting both endolithic habitats would be composed almost entirely of *Galdieria* (Fig. 4.5A). The relatively low *Cyanidioschyzon* representation in the site D endolithic sample (Table 4.2) may potentially be too low to readily detect in the autotrophic and heterotrophic viable count comparisons, however, the dominant presence of *Cyanidioschyzon* in the site C endolithic sample (90%)(Table 4.2) should have been readily detectable in the viable count comparisons. The large disconnect observed between direct microscopic counts and viable plate counts in the site C sample (Fig. 4.5A; Fig. 4.5B) may not be a viability issue, but rather the discrepancy is due to the inability or poor capacity of some *Cyanidioschyzon* populations to grow on the solid medium used in the study. This was encountered with the *Galdieria* terbinafine testing, where all three YNP Type II *Galdieria* isolates grew exceptionally poor on solid media and required liquid cultivation.

With regards to the terbinafine studies, pure cultures of *Cyanidioschyzon* and *Galdieria* showed a clear pattern of sensitivity to this anti-fungal agent, wherein only *Galdieria* cultures were resistant (Table 4.3). Furthermore, all Cyanidiales forming colonies under heterotrophic growth conditions were also found to be resistant to terbinafine, implying this anti-fungal agent may be a fairly reliable tool for discriminating between these algae, and for selectively culturing *Galdieria* from environmental samples. Terbinafine is reported to function as an ergosterol which inhibits squalene monooxygenase (SQP), a key enzyme involved in sterol biosynthesis (21). The nuclear genome sequence of *C. merolae* strain 10D documents the SQP gene (CMH256C), but a

SQP-based BLASTn search of the *G. sulphuraria* (EST) library failed to identify an SQP homolog. Terbinafine resistance in *Galdieria* indicates the presence of either an alternative SQP enzyme or sterol biosynthesis pathway, neither of which is inhibited by terbinafine. Terbinafine sensitivity among the *Cyanidioschyzon* Type IA and Type IB isolates implies they have similarities beyond the 18S rDNA and *rbcL* genotypes, and that even though the Type IA isolates are morphologically similar to *Galdieria*, they differ in membrane properties.

References

1. Allen, MB. 1959. Studies with *Cyanidium caldarium*, an anomalously pigmented chlorophyte. Arch. Microbiol. 32:270–7.

2. Bhatnagar, A, Makandar, MB, Garg, MK, and Bhatnagar, M. 2008. Community structure and diversity of cyanobacteria and green algae in the soils of Thar Desert (India). Journal of Arid Environments. 72:73-83.

3. Botero, LM, Brown, KB, Brumefield, S, Burr, M, Castenholz, RW, Young, M, and McDermott, TR. 2004. *Thermobaculum terrenum* gen. nov., sp. nov.: a non-phototrophic gram-positive thermophile representing an environmental clone group related to the Chloroflexi (green non-sulfur bacteria) and Thermomicrobia. Arch. Microbiol. 181:269–277.

4. Brock, TD. 1978. Thermophilic microorganisms and life at high temperatures. Springer-Verlag, New York, NY.

5. Ciniglia C, Yoon HS, Pollio A, Pinto G, Bhattacharya D. 2004. Hidden biodiversity of the extremophilic Cyanidiales red algae. Mol. Ecol. 13:1827-1838.

6. Ciniglia, C, Valentino, GM, Cennamo, P, DeStefano, M, Stanzione, D, Pinto, G, and Pollio, A. 2005. Influences of geochemical and mineralogical constraints on algal distribution in acidic hydrothermal environments: Pisciarelli (Naples, Italy) as a model site. Arch. Hydrobiol. 162:121–42.

7. Doemel, WN, and Brock, TD. 1971. The physiological ecology of *Cyanidium caldarium*. J. Gen. Microbiol. 67:17–32.

8. Fan, KW, Aki, T, and Chen, F. 2010. Enhanced production of squalene in the thraustochytrid *Aurantiochytrium mangrovei* by medium optimization and treatment with terbinafine. World J. Microbiol. and Biotechnol. 26:1303-1309.

9. Ferris, MJ, Sheehan, KB, Kuhl, M, Cooksey, K, Wiggleworth-Cooksey, B, Harvey, R, and Henson, JM. 2005. Algal species and light microenvironment in a low-pH, geothermal microbial mat community. Appl. Environ. Microbiol. 71:7164-7171.

10. Gross, W, and Schnarrenberger, C. 1995. Heterotrophic growth of two strains of the acido-thermophilic red alga Galdieria sulphuraria. Plant Cell Physiology. 36(4):633-638.

11. Gross, W, Kuver, J, Tischendorf, G, Bouchaala, N, and Busch, W. 1998. Cryptoendolithic growth of the red alga *Galdieria sulphuraria* in volcanic

areas. Eur. J. Phycol. 33:25-31.

12. Gross, W, Heilmann, I, Lenze, D, and Schnarrenberger, C. 2001. Biogeography of the Cyanidiaceae (Rhodophyta) based on 18S ribosomal RNA sequence data. Eur. J. Phycol. 36:275–280.

13. Heilmann, I, and Gross, W. 2001. Genetic diversity of thermo-acidophilic red algae according to random amplified polymorphic DNA (RAPD) analysis. Nova Hedwigia 123:531–539.

14. Hofmann, BA, Farmer, JD. 2000. Filamentous fabrics in low-temperature mineral assemblages: are they fossil biomarkers? Implications for the search for a subsurface fossil record on the early Earth and Mars. Planet. Space. Sci. 48:1077-1086.

15. Lehr, CR, Frank, SD, Norris, TB, D'Imperio, S, Kalinin, AV, Toplin, JA, Castenholz, RW, and McDermott, TR. 2007. Cyanidia (Cyanidiales) population diversity and dynamics in an acid-sulfate-chloride spring in Yellowstone National Park. J. Phycol. 43:3–14.

16. Mason, CA, Hamer, G, and Bryers, JD. 1986. The death and lysis of microorganisms in environmental processes. FEMS Microbiol. Rev. 39:373-401.

17. Miller, SR, Castenholz, RW, and Pedersen, D. 2007. Phylogeography of the thermophilic cyanobacterium *Mastigocladus laminosus*. Applied and Environ. Microbiol. 73(15):4751-4759.

18. Nordstrom, DK, Ball, JW, and McCleskey, RB. 2005. Ground water to surface water: chemistry of thermal outflows in Yellowstone National Park, In: Inskeep, WP, McDermott, TR, (eds) Geothermal biology and geochemistry in Yellowstone National Park. Montana State University, Bozeman. pp. 73–94.

19. Oesterhelt, C, Schmalzlin, E, Schmitt, J, Lokstein, H. 2007. Regulation of photosynthesis in the unicellular acidophilic red alga *Galdieria sulphuraria*. Plant J. 51(3):500-11.

20. Pinto, G, Ciniglia, C, Cascone, C, and Pollio, A. 2007. Species composition of Cyanidiales assemblages in Pisciarelli (Campi Flegrei, Italy) and description of *Galdieria Phlegrea* sp. nov. In: Seckbach, J, (eds) Algae and cyanobacteria in extreme environments. Springer, Dordrecht. pp. 487–502.

21. Ryder, NA. 1990. Inhibition of squalene epoxidase and sterol side-chain methylation by allyamines. Biochem. Soc. Trans. 18:45-46.

22. Seckbach, J. 1994. Evolutionary pathways and enigmatic algae: *Cyanidium caldarium* (Rhodophyta). Kluwer Academic Publishers, Dordrecht, The Netherlands. pp. 349-350.

23. Seckbach, J. 1999. The Cyanidiophyceae: hot spring acidophilic algae. In: Seckbach, J, (eds) Enigmatic microorganisms and life in extreme environments. Kluwer Academic Publishers, Dordrecht, The Netherlands. pp. 425–435.

24. Sloth, J, Wiebe, M, and Eriksen, N. 2006. Accumulation of phycocyanin in heterotrophic and mixotrophic cultures of the acidophilic red alga *Galdieria sulphuraria*. Enzyme Microb. Technol. 38:168-175.

25. Souffreau, S, Vanormelingen, P, Verleyen, E, Sabbe, K, and Vyverman. 2010. Tolerance of benthic diatoms from temperate aquatic and terrestrial habitats to experimental desiccation and temperature stress. Phycologia. 49(4):309-324.

26. Toplin, JA, Norris, TB, Lehr, CR, McDermott, TR, and Castenholz, RW. 2008. Biogeographic and phylogenetic diversity of thermoacidophilic Cyanidiales in Yellowstone National Park, Japan, and New Zealand. Appl. Environ. Microbiol. 74:2822-2833.

27. Walker, JJ, Spear, JR, and Pace, NR. 2005. Geobiology of a microbial endolithic community in the Yellowstone geothermal environment. Nature. 434:1011-1013.

28. Yoon HS, Ciniglia C, Wu M, Comeron JM, Pinto G, Pollio A, Bhattacharya D. 2006. Establishment of endolithic populations of extremophilic Cyanidiales (Rhodophyta). BMC Evol. Biol. 6:78.

# CHAPTER 5

# CONCLUSIONS AND FUTURE STUDIES

## Conclusions

The results of the research presented in this dissertation suggest that a number of environmental factors influence the population dynamics of the Cyanidiales, an order of unicellular eukaryotic red algae that flourish in thermoacidic environments (pH 0.2-4.0, temperature 40-56$^{o}$C). The in-depth temporal analysis of *in situ* UV-VIS irradiance on cyanidial gene expression patterns indicated both UV exposure and seasonal UV-VIS intensity influenced cyanidial populations. Greatest cyanidial sensitivity to UV exposure occurred during fall and winter, periods of the year featuring the lowest UV levels and shortest photoperiods. Both these time points saw the statistically significant repression of a large number genes coding for repetitive DNA elements, which prior to the present body of work, were not known to be associated with UV-irradiance response nor display such extensive *in situ* transcriptional activity. These results have raised intriguing new questions into the role repetitive elements play in helping the Cyanidiales respond to environmental forces, suggesting they may help facilitate cyanidial adaptation to UV stress conditions. Future research should be directed at characterizing the activity of repetitive DNA elements in response to UV-VIS irradiance, with the goal of uncovering the molecular mechanism(s) by which repetitive elements alter UV-VIS regulated expression.

Careful characterization of irradiance-specific expression changes occurring during summer, a period of the year with increased UV-VIS irradiance intensity and a lengthened photoperiod, suggested increased VIS irradiance intensity was masking UV-based gene regulation affects. By comparing cyanidial gene expression during the lowest and highest UV-VIS irradiance periods of year, separate analyses identified genes regulated by UV or VIS separately, or by the collective effects of UV and VIS. VIS-specific effects regulated genes involved in replication, energy, and growth-related metabolisms, suggesting VIS-related stress may be negatively effecting algal health and growth, potentially resulting in decreased viability. Furthermore, direct regression analysis uncovered a statistically significant positive relationship between VIS intensity and the heat shock response, indicating VIS-stress also activated cellular protective and repair systems. Expression changes specific to UV intensity were also observed, and involved the significant repression of genes associated with the photosystems and accessory pigments (the primary targets of UV-induced photodamage). These results expand on previous pure culture studies, and highlight the value of using environmental microarrays to identify ecologically-relevant gene expression patterns. Interestingly, apparent combined UV-VIS response was only evident for genes annotated as repetitive DNA elements, where UV was found to attenuate VIS repression affects. Results again address the importance of identifying the mechanism(s) but which repetitive DNA elements affect cyanidial irradiance stress response.

The results presented in this dissertation further demonstrate the ability of environmental microarrays to approximate ecologically-relevant gene expression patterns

by assessing the relative information yield of transcriptomes characterized via different methods; i.e. pyrosequencing versus microarrays. Multiple comparative approaches were taken to assess the degree of method sensitivity; results indicated environmental microarrays were far more effective at detecting transcript diversity and gene expression changes than transcriptomes developed via pyrosequencing. Collector's curve analysis suggested minimal sequencing depth reduced the fraction of gene expression activity and diversity captured in the pyrosequencing libraries. To increase metatranscriptome resolving power as it relates to gene expression activity, future sequencing efforts should include internal mRNA standards, which allow for the absolute quantification of transcript copy number regardless of sequencing depth (4).

Limited characterization of the algal proteome was also included in the in-depth temporal analysis to assess the potential connectivity of the algal transcriptome to the proteome. Results identified several dominant cyanidial proteins that displayed high probe signal intensity on matching environmental microarrays, which indicated there is good correspondence between highly abundant cyanidial proteins and gene transcriptional activity. Collectively, the data suggested that complimentary information obtained by both proteomic and transcriptomic techniques can be used to tease out ecologically-relevant molecular changes occurring in *in situ* microbial populations. Future *in situ* studies should increase the incorporation of proteomic analysis, in order further the understanding of cyanidial response to environmental change.

Much remains to be learned about the distribution and diversity of the Cyanidiales in geothermal environments. The results presented in this dissertation also shed light on

the influence specific physical factors exert on cyanidial species composition as well as the proportional occurrence of specific genera. Cyanidiales diversity was assessed in aqueous, soil and endolithic environments within the Yellowstone National Park (YNP) geothermal complex. Results from phylogenetic reconstruction using the *rbcL* gene identified two well-supported YNP lineages: *Galdieria*-A and *Cyanidioschyzon*. YNP *Galdieria*-A environmental clones were localized exclusively to non-aqueous habitats, while *Cyanidioschyzon* clones were detected in all aqueous, soil, and endolithic habitats sampled. Culture-dependent analysis assessed major physical factors shaping Cyanidiales assemblages in non-aqueous environments and the results suggested gravimetric soil moisture content was an important physical factor controlling cyanidial viability and abundance in soil habitats. Interestingly, viability rates in the endolithic environments sampled did not co-vary with gravimetric moisture content. These results have raised intriguing new questions into what environmental parameter(s) influence endolithic cyanidial populations. One such environmental factor may be habitat humidity, where water vapors trapped within the rock pores create a high relative humidity that can support the growth of microbial communities (14). Future research should be directed at characterizing the endolithic microclimate, with the goal of uncovering whether or not relative humidity is exerting selection pressures on endolithic cyanidial populations.

## Recommendations for Future Studies

### Diurnal UV-VIS Effects on *in situ* Cyanidial Gene Expression Patterns

The environmental microarray work presented in chapter 3 focused on examining genome-level response of the Cyanidiales to UV-VIS irradiance exposure as it related to cyanidial mat decline, an annual seasonal algal viability reduction caused, at least in part, by seasonal shifts in water temperature, UV irradiance, and/or visible (VIS) light irradiance. Gene expression patterns were monitored over a time period of increasing UV-VIS intensity which correlated well with the mat decline event, but which was very broad, and represented only a single snap-shot of the genes responding to seasonal UV-VIS intensity. In the future, *in situ* cyanidial transcriptional behaviors should also be monitored over shorter time scales, so that more detailed diurnal UV-VIS effects on cyanidial gene expression can be examined. Diurnal differences for *in situ* transcriptional patterns could then be compared across seasons that vary with respect to UV-VIS intensity, allowing for a more comprehensive understanding of adaptive strategies employed by these algae in response to changing UV-VIS. Diurnal sampling and analysis from *in situ* microbial mat communities does present its own challenges, including the necessitation of low biomass sampling (to prevent excessive impact on the spring environment). To obtain enough mRNA for microarray hybridization, diurnal samples would need to be linearly amplified, which has been suggested to introduce experimental variation (11). Recent work, however, has suggested that variability is not increased, and that in fact, linear-amplified mRNA transcriptional profiles show a high degree of reproducibility (13). Collectively, the results of this new research thrust should provide

novel insights into diurnal transcriptional patterns, and a detailed framework by which the Cyanidiales respond to UV-VIS irradiance.

## Effects of High Temperature Stress on Cyanidial Mat Decline

Results of the direct regression analysis presented in chapter 3 failed to identify a statistically significant relationship between water temperature and the expression of any of the genes included in the microarray design, suggesting water temperature was not a contributing factor to cyanidial mat decline. Though seasonal changes in water temperature were somewhat modest, peak summer temperatures recorded in July and August approached the upper temperature limit for the Cyanidiales. Future studies should assess the effects of water temperature on Cyanidiales population dynamics as it relates to mat decline. Previous studies have shown that phototrophic bacterial strains collected and cultivated from hot spring environments often have significantly lower optimal temperatures for photosynthesis and growth than environmental temperatures measured at the time of collection (8, 9). It is therefore possible that seasonal increases in water temperature are selecting for cyanidial populations with higher optimum growth temperatures, and thus reductions in cyanidial viability during the mat decline period are due to dramatic shifts in community structure. This is consistent with previous cyanidial population dynamics analyses, which found seasonal shifts in cyanidial population size corresponded to changes in 18s rDNA clone library composition and chloroplast short sequence repeat (cpSSR) profiles (5).

If additional monthly samplings were conducted at Lemonade Creek, cpSSR analysis could be performed on *in situ* mat community DNA to track population diversity

and dynamics across seasons with varied water temperature levels. SSR mat community amplicons could be matched to pure culture YNP cyanidial isolates, both in terms of electrophoretic patterns as well as sequence identity. [$^{14}$C]-photo incorporation assays along with temperature-related growth studies could then be conducted on pure culture cyanidial isolates which matched dominant seasonal SSR mat community amplicons. Optimal photosynthetic carbon uptake rates and growth temperature could be used to determine if cyanidial populations with increased optimal temperatures are dominant during periods of the year with maximal water temperatures and noticeable mat decline effects.

## The Role of Repetitive DNA Elements In UV-VIS Stress Response

The research presented in chapter 3 uncovered the noticeable repression of a large number of genes coding for repetitive DNA elements (RE) in response to both UV-VIS exposure as well as seasonal shifts in UV-VIS intensity. Though RE's are known to regulate genomic change in response to environmental signals, until the current study they were not yet associated with UV-VIS-related stress response. Initial studies should be conducted to examine the genome location of RE's apparently repressed in response to UV-VIS, as RE's are often localized within the flanking regions of the gene they exert transcriptional control over (12). Microarray data sets could then be used to assess whether genes directly adjacent to a repressed RE on the genome were significantly up- or down-regulated in response to either UV-VIS exposure or seasonal shifts in UV-VIS intensity. Following the identification of several candidate genes, spring-relevant culture isolates could be used in RNA-based laboratory experiments to compare the expression

response of these genes as a function of UV-VIS irradiance and of sampling time. RNA extracted and converted to cDNA could be used for gene expression analyses, potentially resulting in the identification of several genes involved in mediating cyanidial UV-VIS stress response. Down-stream pure culture work would then aim to identify the regulatory mechanism by which the adjacent RE suppresses or induces the expression of the candidate gene under UV-VIS-related stress conditions. Results could potentially provide a working model for how RE's are regulating the expression of genes important in UV-VIS-related stress response.

### Influence of Habitat Humidity on Endolithic Cyanidiales Assemblages

The culture-dependent work presented in chapter 4 focused on assessing major physical factors shaping Cyanidiales assemblages in non-aqueous acidic geothermal environments. Samples collected and cultivated from soil habitats showed a strong correlation to soil moisture level, demonstrating increased colonization and viability in relation to increased gravimetric soil moisture content. Viability rates in the endolithic habitats sampled, however, did not co-vary with gravimetric moisture content, and were instead suggested to be influenced by habitat humidity. In the future, the cyanidial endolithic microclimate should be tested for relative humidity levels, to determine whether humidity is shaping the distribution and abundance of endolithic cyanidial populations. If additional field studies were conducted, long-term measurements of relative humidity at several endolithic sites could be monitored using a thin-film humidity sensor inserted into the rock (1, 3, 10). Endolithic sites could then be sampled and used to estimate the proportional occurrence of the different Cyanidiales genera in endolithic

habitats that varied naturally in relative humidity. Simultaneous extraction of DNA from the samples could be performed and used for *rbcL* gene sequencing. This would allow for a comparison of the species composition across endolithic sites. Results would provide a more comprehensive analysis of whether or not relative humidity shapes Cyanidiales diversity and proportional occurrence in the endolithic environment.

References

1. Bell, RA. 1993. Cryptoendolithic algae of hot semiarid lands and deserts. J. Phycol. 29:133–139.

2. Ciniglia C, Yoon HS, Pollio A, Pinto G, Bhattacharya D. 2004. Hidden biodiversity of the extremophilic Cyanidiales red algae. Mol. Ecol. 13:1827-1838.

3. Friedmann EI, Kappen L, Meyer MA, and Nienow JA. 1993. Long-term productivity in the cryptoendolithic microbial community of the Ross Desert, Antarctica. Microb. Ecol. 25:51–69.

4. Gifford, SM, Sharma, S, Rinta-Kanto, JM, and Moran, MA. 2011. Quantitative analysis of a deeply sequenced marine microbial metatranscriptome. The ISME Journal. 5:461-472.

5. Lehr, CR, Frank, SD, Norris, TB, D'Imperio, S, Kalinin, AV, Toplin, JA, Castenholz, RW, and McDermott, TR. 2007. Cyanidia (Cyanidiales) population diversity and dynamics in an acid-sulfate-chloride spring in Yellowstone National Park. J. Phycol. 43:3–14.

6. Lizotte, MP, and Sullivan, CW. Rates of photoadaptation in sea ice diatoms from McMurdo Sound, Antarctica. J Phycol. 27:367-373.

7. McMinn, A, and Gademann. Diurnal changes in photosynthesis of Antarctic fast ice algal communities determined by pulse amplitude modulation fluorometry. Marine Biology. 143:359-367.

8. Meeks, JC, and Castenholz, RW. 1971. Growth and photosynthesis in an extreme thermophile, *Synechococcus lividus* (Cyanophyta). Arch. Mikrobiol. 78:25-41.

9. Miller, SR, Wingard, CE, and Castenholz, RW. 1998. Effects of visible light and ultraviolet radiation on photosynthesis in a population of the hot spring cyanobacterium, *Synechococcus* sp., subjected to high-temperature stress. Applied and Environmental Microbiology. 64:3893-3899.

10. Palmer RJ, Friedmann, EI. 1990. Water relations and photosynthesis in the Cryptoendolithic microbial habitat of hot and cold deserts. Microbiol. Ecol. 19:111–18.

11. Patel, OV, Suchyta, SP, Sipkovsky, SS, Yao, J, Ireland, JJ, Coussens, PM, and Smith, GW. Validation and application of a high fidelity mRNA linear amplification procedure for profiling gene expression. Veterinary Immunology and Immunopathology. 105:331-342.

12. Schmidt, AL, and Anderson, LM. 2006. Repetitive DNA elements as mediators of genomic change in response to environmental cues. Biol. Rev. 81:531-543.

13. Stewart, FJ, Ottesen, EA, and DeLong, ED. 2010. Development and quantitative analyses of a universal rRNA-subtraction protocol for microbial metatranscritomics. The ISME Journal. 4:896-907.

14. Walker, JJ, and Pace, NR. 2007. Endolithic microbial ecosystems. Annual Review of Microbiology. 61: 331-347.

# LITERATURE CITED

1. Aguilera, A, Souza-Egipsy, V, Gomez, F, and Amils, R. 2007. Development and structure of eukaryotic biofilms in an extreme acidic environment, Río Tinto (SW, Spain). Microb. Ecol. 53:294-305.

2. Albertano, P, Ciniglia, C, Pinto, G, and Pollio, A. 2000. The taxonomic position of *Cyanidium*, *Cyanidioschyzon*, and *Galdieria*: an update. Hydrobiologia. 433:137-143.

3. Allen, MB. 1959. Studies with *Cyanidium caldarium*, an anomalously pigmented chlorophyte. Arch. Microbiol. 32:270-277.

4. Asimgil, H, Kavakli, IH. 2004. Purification and characterization of five members of photolyase/cryptochrome family from *Cyanidioschyzon merolae*. Plant Science. *In Press*.

5. Azua-Bustos, A. 2010. Chilean Cave *Cyanidium.* In: Seckbach, J. and Chapman, D.J. (eds), Red algae in the genomic age. Springer Science and Business Media, B.V. pp. 425-438.

6. Barbier, G, Oesterhelt, C, Larson, MD, Halgren. RG, Wilkerson, C, Garavito, RM, Benning, C, and Weber, AP. 2005. Comparative genomics of two closely related unicellular thermo-acidophilic red algae, *Galdieria sulphuraria* and *Cyanidioschyzon merolae*, reveals the molecular basis of the metabolic flexibility of *Galdieria sulphuraria* and significant differences in carbohydrate metabolism of both algae. Plant Physiol. 137:460–474.

7. Bell, RA. 1993. Cryptoendolithic algae of hot semiarid lands and deserts. J. Phycol. 29:133–139.

8. Bhatnagar, A, Makandar, MB, Garg, MK, and Bhatnagar, M. 2008. Community structure and diversity of cyanobacteria and green algae in the soils of Thar Desert (India). Journal of Arid Environments. 72:73-83.

9. Bonheyo, G, Frias-Lopez, J, and Fouke, BW. 2005. A test for airborne dispersal of thermophilic bacteria from hot springs, In: Inskeep, WP, and McDermott, TR, (eds), Geothermal Biology and Geochemistry in Yellowstone National Park. Montana State University Publication, Bozeman, pp. 327-340.

10. Botero, LM, Brown, KB, Brumefield, S, Burr, M, Castenholz, RW, Young, M, and McDermott, TR. 2004. *Thermobaculum terrenum* gen. nov., sp. nov.: a non-phototrophic gram-positive thermophile representing an environmental clone group related to the Chloroflexi (green non-sulfur bacteria) and Thermomicrobia. Arch. Microbiol. 181:269–277.

11. Brock, T.D.1967. Life at high temperatures. Science. 158:1012–1018.

12. Brock, TD, and Darland, ML. 1970. Limits of microbial existence: temperature and pH. Science. 169:1316-1318.

13. Brock, TD. 1978. The genus *Cyanidium*, In: Brock, TD (eds), Thermophilic microorganisms and life at high temperatures. Springer Verlag, New York, pp. 255-302.

14. Brock, TD. 1978. Thermophilic microorganisms and life at high temperatures. Springer-Verlag, New York, NY.

15. Castenholz, RW, and McDermott, TR. 2010. The Cyanidiales: ecology, biodiversity, and biogeography, In: Seckbach, J. and Chapman DJ (eds), Red algae in the genomic age. Springer Science and Business Media, B.V. pp. 357-371.

16. Chapin, SF. 2002. Carbon input to terrestrial ecosystems. In: Matson, PA, and Mooney, HA., (eds) Principles of Terrestrial Ecosystem Ecology. Springer Science and Business Media, Inc, New York.

17. Ciniglia, C, Yoon, HS, Pollio, A, Pinto, G, and Bhattacharya, D. 2004. Hidden biodiversity of the extremophilic Cyanidiales red algae. Mol. Ecol. 13:1827–1838.

18. Ciniglia, C, Valentino, GM, Cennamo, P, DeStefano, M, Stanzione, D, Pinto, G, and Pollio, A. 2005. Influences of geochemical and mineralogical constraints on algal distribution in acidic hydrothermal environments: Pisciarelli (Naples, Italy) as a model site. Arch. Hydrobiol. 162:121–42.

19. Cockell, CS, and Rothschild, LJ. 1999. The effects of UV radiation A and B in diurnal variation in photosynthesis in three taxonomically and ecologically diverse microbial mats. Photochem. Photobiol. 69:203–210.

20. Collen, J, Roeder, V, Rousvoal, S, Collin, O, and Kloareg, CB. 2006. An expressed sequence tag analysis of thallus and regenerating protoplasts of *Chondrus crispus* (gigartinales, rhodophyceae). Journal of Phycology. 42:104-112.

21. Cunningham, FX, Lee, H, and Gantt, E. 2007. Carotenoid biosynthesis in the primitive red alga *Cyanidioschyzon merolae*. Eukaryotic Cell. 6:533-545.

22. Darienko, T, and Hoffman, L. 2010. Subaerial algae and cyanobacteria from the archaeological remains of Carthage (Tunisia), including the record of a species of *Cyanidium* (Rhodophyta). Algological Studies. 135:41-60.

23. De Luca, P, and Moretti, A. 1983. Floridosides in *Cyanidium caldarium*, *Cyanidioschyzon merolae* and *Galdieria sulphuraria* (Rhodophyta, Cyanidiophyceae). J. Phycol. 19:368–369.

24. Diaz, MM, and Maberly, SC. 2009. Carbon-concentrating mechanisms in acidophilic algae. Phycologia. 48:77–85.

25. Doemel, W, and Brock, TD. 1970. The upper temperature limit of *Cyanidium caldarium*. Arch. Microbiol. 72:326-332.

26. Doemel, WN, and Brock, TD. 1971. The physiological ecology of *Cyanidium caldarium*. J. Gen. Microbiol. 67:17–32.

27. Enami, I, Akutsu, H, and Kyogoku, Y. 1986. Intracellular pH regulation in an acidophilic unicellular alga, *Cyanidium caldarium*: $^{31}$P-NMR determination of intracellular pH. Plant Cell Physiol. 27:1351–1359.

28. Enami, I, Hideyuki, A, and Shen, J. 2010. Mechanisms of acido-tolerance and characteristics of photosystems in and acidophilic and thermophilic red alga, *Cyanidium caldarium*. In Seckbach, J, and Chapman, DJ (eds), Red algae in the genomic age. Springer Science and Business Media, B.V. pp. 373-389.

29. Fan, KW, Aki, T, and Chen, F. 2010. Enhanced production of squalene in the thraustochytrid *Aurantiochytrium mangrovei* by medium optimization and treatment with terbinafine. World J. Microbiol. and Biotechnol. 26:1303-1309.

30. Ferris, MJ, Sheehan, KB, Kuhl, M, Cooksey, K, Wiggleworth-Cooksey, B, Harvey, R, and Henson, JM. 2005. Algal species and light microenvironment in a low-pH, geothermal microbial mat community. Appl. Environ. Microbiol. 71:7164-7171.

31. Foissner, W. 2006. Biogeography and dispersal of micro-organisms: a review emphasizing protists. Acta Protozool. 45:111-136.

32. Friedmann, I. 1964. Progress in the biological exploration of caves and subterranean waters in Israel. Int. J. Speleol. 1:29–33.

33. Friedmann EI, Kappen L, Meyer MA, and Nienow JA. 1993. Long-term productivity in the cryptoendolithic microbial community of the Ross Desert, Antarctica. Microb. Ecol. 25:51–69.

34. Fujiwara, T., Misumi, O., Tashiro, K., Yoshida, Y., Nishida, K., Yagisawa, F., Imamura, S., Yoshida, M., Mori, T., Tanaka, K., Kuroiwa, H., and Kuroiwa, T. 2009. Periodic gene expression patterns during the highly synchronized cell nucleus and organelle division cycles in the unicellular red alga Cyanidioschyzon merolae. DNA Res. 16:59–72.

35. Fujiwara, T, Kuroiwa, H., Yagisawa, F., Ohnuma, M., Yoshida, Y., Yoshida, M., Nishida, K., Misumi, S., Watanabe, S., Tanaka, K., and Kuroiwa, T. 2010. The coiled-coil protein VIG1 is essential for tethering vacuoles to mitochondria during vacuole inheritance of *Cyanidioschyzon merolae*. The Plant Cell. 22:772-781.

36. Geitler, L. 1933. Diagnoses neuer blaualgen von den Sunda-Insela. Arch. Hydrobiol. Suppl. 12.

37. Gifford, SM, Sharma, S, Rinta-Kanto, JM, and Moran, MA. 2011. Quantitative analysis of a deeply sequenced marine microbial metatranscriptome. The ISME Journal. 5:461-472.

38. Gross, W, and Schnarrenberger, C. 1995. Heterotrophic growth of two strains of the acido-thermophilic red alga Galdieria sulphuraria. Plant Cell Physiology. 36(4):633-638.

39. Gross, W, Küver, J, Tischendorf, G, Bouchaala, N, and Busch, W. 1998. Cryptoendolithic growth of the red alga *Galdieria sulphuraria* in volcanic areas. Eur. J. Phycol. 33:25–31.

40. Gross, W, and Oesterhelt, C. 1999. Ecophysiological studies on the red alga *Galdieria sulphuraria* isolated from southwest Iceland. Plant Biol. 1:694–700.

41. Gross, W. 2000. Ecophysiology of algae living in highly acidic environments. Hydrobiologia. 433:31-37.

42. Gross, W, and Gross, S. 2001. Physiological characterization of the acidophilic red alga *Galdieria sulphuraria* isolated from a mining area. Nova Hedwigia, Beiheft. 123:523–530.

43. Gross, W, Heilmann, I, Lenze, D, and Schnarrenberger, C. 2001. Biogeography of the Cyanidiaceae (Rhodophyta) based on 18S ribosomal RNA sequence data. Eur. J. Phycol. 36:275–280.

44. Gross, W, Oesterhelt, C, Tischendorf, G, and Lederer, F. 2002. Characterization of a non-thermophilic strain of the red algal genus *Galdieria* isolated from Soos (Czech Republic). Eur. J. Phycol. 37:477–482.

45. Hashimoto, M., Misumi, O., Hujiwara, T., Sakajiri, T., and Kuroiwa, T. 2006. Gene expression profiles by temperature change of the red alga *Cyanidioschyzon merolae*. J. Plant Res. 119 suppl. 74.

46. Heilmann, I, and Gross, W. 2001. Genetic diversity of thermo-acidophilic red algae according to random amplified polymorphic DNA (RAPD) analysis. Nova Hedwigia 123:531–539.

47. Hofmann, BA, Farmer, JD. 2000. Filamentous fabrics in low-temperature mineral assemblages: are they fossil biomarkers? Implications for the search for a subsurface fossil record on the early Earth and Mars. Planet. Space. Sci. 48:1077-1086.

48. Huang, L, McCluskey, MP, Ni, H, and LaRossa, R. 2002. 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(24): 6845-6858.

49. Imamura, S., Kanesaki, Y., Ohnuma, M., Inouye, T., Sekine, Y., Fujiwara, F., Kuroiwa, T., and Tanaka, K. 2009. R2R3-type MYB transcription factor, CmMYB1, is a central nitrogen assimilation regulator in Cyanidioschyzon merolae. Proc. Natl. Acad. Sci. USA. 106(30):12548-12553.

50. Inskeep, WP, and McDermott, TR. 2005. Geomicrobiology of acid-sulfate-chloride springs in Yellowstone National Park. In: Inskeep, WP and McDermott, TR, (eds) Geothermal biology and geochemistry in Yellowstone National Park. Montana State University, Bozeman, pp. 143-162.

51. Jackson, CR, Langner, HW, Donahoe-Christiansen, J, Inskeep, WP, and McDermott, TR. 2001. Molecular analysis of microbial community structure in an arsenite-oxidizing acidic thermal spring. Environ. Microbiol. 3:532-542.

52. Johnson, BA. 1998. Biodiversity and ecology of acidophilic microorganisms. FEMS Microbiol. Ecol. 27:307-317.

53. Kelly, DJA., Budd, K, Lefebvre, DD. 2007. Biotransformation of mercury in pH-stat cultures of eukaryotic freshwater algae. Arch Microbiol. 187:45-53.

54. Kim, DK, Hong, SJ, Bae, JH, Yim, N, Jin, ES, and Lee, CG. 2011. Transcriptomic analysis of *Haematococcus lacustris* during astaxanthin accumulation under high irradiance and nutrient starvation. Biotechnol. Bioprocess Eng. 16:398-705.

55. Kunin, V, Engelbrektson, A, Ochman, H, and Hugenholtz, P. 2010. Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. Environ. Microbiol. 12:118–123.

56. Kura-Hotta, M, and Enami, I. 1981. Light-induced H+ efflux from intact cells of *Cyanidium caldarium*. Plant Cell Physiol. 22:1175–1183.

57. Kuroiwa, T, Nishida, K, Yoshida, Y, Fujiwara, T, Mori, T, Kuroiwa, H, and Misumi, O. 2006. Structure, function and evolution of the mitochondrial division apparatus. Biochim. Biophys. Acta. 1763:510–521.

58. Lehr, CR, Frank, SD, Norris, TB, D'Imperio, S, Kalinin, AV, Toplin, JA, Castenholz, RW, and McDermott, TR. 2007. Cyanidia (Cyanidiales) population diversity and dynamics in an acid-sulfate-chloride spring in Yellowstone National Park. J. Phycol. 43:3–14.

59. Li, H, and Durbin R. 2009. Fast and accurate short read alignment with burrows-wheeler transform. Bioinformatics. 25:1754-60.

60. Li, Z, Wakao, S, Fischer, BB, and Niyogi, KK. 2009. Sensing and responding to excess light. Annu. Rev. Plant Biol. 60:239-260.

61. Lizotte, MP, and Sullivan, CW. Rates of photoadaptation in sea ice diatoms from McMurdo Sound, Antarctica. J Phycol. 27:367-373

62. Macur, RE, Langner, HW, Kocar, BD, and Inskeep, WP. 2004. Linking geochemical processes with microbial community analysis: successional dynamics in an arsenic-rich, acid-sulphate-chloride geothermal spring. Geobiology. 2:163-177.

63. Mason, CA, Hamer, G, and Bryers, JD. 1986. The death and lysis of microorganisms in environmental processes. FEMS Microbiol. Rev. 39:373-401.

64. Mathur, J, Bizzco, RW, Ellis, DG, Lipson, DA, Poole, AW, Levine, R, and Kelley, ST. 2007. Effects of abiotic factors on the phylogenetic diversity of bacterial communities in acidic thermal springs. Appl. Environ. Microbiol. 73:2612-2623.

65. Matsuzaki, M, Misumi, O, Shin-i, T, Maruyama, S, Takahara, M, Miyagishima, SY, Mori, T, Nishida, K, Yagisawa, F, Nishida, K, Yoshida, Y, Nishimura, Y, Nakao, S, Kobayashi, T, Momoyama, Y, Higashiyama, T, Minoda, A, Sano, M, Nomoto, H, Oishi, K, Hayashi, H, Ohta, F, Nishizaka, S, Haga, S, Miura, S, Morishita, T, Kabeya, Y, Terasawa, K, Suzuki, Y, Ishii, Y, Asakawa, S, Takano, H, Ohta, N, Kuroiwa, H, Tanaka, K, Shimizu, N, Sugano, S, Sato, N, Nozaki, H,

Ogasawara, N, Kohara, Y, and Kuroiwa, T. 2004. Genome sequence of the ultra small unicellular red alga *Cyanidioschyzon merolae* 10D. Nature. 428:653–657.

66. McMinn, A, and Gademann. Diurnal changes in photosynthesis of Antarctic fast ice algal communities determined by pulse amplitude modulation fluorometry. Marine Biology. 143:359-367.

67. Meeks, JC, and Castenholz, RW. 1971. Growth and photosynthesis in an extreme thermophile, *Synechococcus lividus* (Cyanophyta). Arch. Mikrobiol. 78:25-41.

68. Meneghini, G. 1839. Nuova specie di alga descritta dal sig. dott. giuseppe meneghini di padova. Nuovo Giorn. Lett. 39:67–68.

69. Merola, A., Castaldo, R., De Luca, P., Gambardella, R., Musacchio, A. and Taddei, R. 1981. Revision of *Cyanidium caldarium*: three species of acidophilic algae. Giorn. Bot. Ital. 115:189–195.

70. Miller, SR, Wingard, CE, and Castenholz, RW. 1998. Effects of visible light and ultraviolet radiation on photosynthesis in a population of the hot spring cyanobacterium, *Synechococcus* sp., subjected to high-temperature stress. Applied and Environmental Microbiology. 64:3893-3899.

71. Miller, SR. 2007. Diversity of the cosmopolitan thermophile *Mastigocladus laminosus* at global, regional, and local scales. In: J. Seckbach (ed) Algae and cyanobacteria in extreme environments. Springer, Dordrecht. pp. 401–410.

72. Miller, SR, Castenholz, RW, and Pedersen, D. 2007. Phylogeography of the thermophilic cyanobacterium *Mastigocladus laminosus*. Applied and Environ. Microbiol. 73(15):4751-4759.

73. Minoda, A., Nagasawa, K., Hanaoka, M., Horiuchi, M., Takahashi, H., and Tanaka, K. 2005. Microarray profiling of plastid gene expression in a unicellular red alga, *Cyanidioschyzon merolae*. Plant Molecular Biology. 59:375-385.

74. Minoda, A, Weber, APM, Tanaka, K, and Miyagishima, S. 2010. Nucleus-independent control of the rubisco operon by the plastid-encoded transcription factor Ycf30 in the red alga *Cyanidioschyzon merolae*. Plant. Physiol. 154:1532-1540.

75. Mohanty, BK, and Kushner, SR. 2010. Bacterial/archaeal/organellar polyadenylation. Wiley Interdisciplinary Reviews: RNA. 2:256-276.

76. Moreno-Paz, M., Gomez, M.J., Arcas, A., and Parro, V. 2010. Environmental transcriptome analysis reveals physiological differences between biofilm and

planktonic modes of life of the iron oxidizing bacteria *Leptospirillum spp.* in their natural microbial community. BMC Genomics. 11:404.

77. Nagasaka, S., Nishizawa, NK., Mori, S. and Yoshimura, EY. 2004. Metal metabolism in the red alga *Cyanidium caldarium* and its relationship to metal tolerance. Biometals. 17:177–181.

78. Nordstrom, DK, Ball, JW and McCleskey, RB. 2005. Ground water to surface water: chemistry and thermal outflows in Yellowstone National Park. In: Inskeep WP, and McDermott TR, (eds) Geothermal biology and geochemistry in Yellowstone National Park. Montana State University, Bozeman, pp. 73-94.

79. Norris, TV, McDermott, TR, and Castenholz, RW. 2002. The long-term effects of UV exclusion on the microbial composition and photosynthetic competence of bacteria in hot-spring microbial mats. FEMS Microbiol. Rev. 39:193-209.

80. Novaes, E, Drost, DR, Farmerie, WG, Pappas Jr, GJ, Grattapaglia, D, Sederoff, R, and Kirst, M. 2008. High-throughput gene and SNP discovery in *Eucalyptus grandis,* an uncharacterized genome. BMC Genomics. 9:312.

81. Nozaki, H, Takano, H, Misumi, O, Terasawa, K, Matsuzaki, M, Maruyama, S, Nishida, K, Yagisawa, F, Yoshida, Y, Fujiwara, T, Takio, S, Tamura, K, Chung, SJ, Nakamura, S, Kuroiwa, H, Tanaka, K, Sato, N, and Kuroiwa, T. 2007. A 100% complete sequence reveals unusually simple genomic features in the hot-spring red alga *Cyanidioschyzon merolae.* BMC Biol. 5:28.

82. Oesterhelt, C, Schmalzlin, E, Schmitt, JM, and Lokstein, H. 2007. Regulation of photosynthesis in the unicellular acidophilic red alga *Galdieria sulphuraria*. The Plant Journal. 50:500-511.

83. Ohta N, Sato N, Kuroiwa T. 1998. Structure and organization of the mitochondrial genome of the unicellular red alga *Cyanidioschyzon merolae* deduced from the complete nucleotide sequence. Nucleic Acids Res. 26:5190–5298.

84. Ohta N, Matsuzaki, M, Misumi, O, Miyagishima, S, Nozaki, H, Tanaka, K, Shin-I, T, Kohara, Y, and Kuroiwa, T. 2003. Complete sequence and analysis of the plastid genome of the unicellular red alga *Cyanidioschyzon merolae.* DNA Res. 10:67–77.

85. Palmer RJ, Friedmann, EI. 1990. Water relations and photosynthesis in the Cryptoendolithic microbial habitat of hot and cold deserts. Microbiol. Ecol. 19:111–18.

86. Park, S, Jung, G, Hwang, Y, and Jin, E. 2010. Dynamic response of the transcriptome of a psychrophilic diatom, Chaetoceros neogracile, to high irradiance. Planta. 231:349-360.

87. Patel, OV, Suchyta, SP, Sipkovsky, SS, Yao, J, Ireland, JJ, Coussens, PM, and Smith, GW. Validation and application of a high fidelity mRNA linear amplification procedure for profiling gene expression. Veterinary Immunology and Immunopathology. 105:331-342.

88. Pattanaik, B, Schuman, R, and Karsten, U. 2007. Effects of ultraviolet radiation on cyanobacteria and their protective mechanisms. In: Seckbach, J. (eds) Algae and Cyanobacteria in Extreme Environments. Springer, Dordrecht. pp. 31-45.

89. Phelps, D, and Buseck, PR. 1980. Distribution of soil mercury and the development of soil mercury anomalies in the Yellowstone geothermal area, Wyoming. Econ. Geol. 75:730-741.

90. Pinto, G, Ciniglia, C, Cascone, C, and Pollio, A. 2007. Species composition of Cyanidiales assemblages in Pisciarelli (Campi Flegrei, Italy) and description of *Galdieria Phlegrea* sp. nov. In: J. Seckbach (ed) Algae and cyanobacteria in extreme environments. Springer, Dordrecht. pp. 487–502.

91. Qin, J, Lehr, CR, Yuan, C, Le, X.C, McDermott, TR, and Rosen, BP. 2009. Biotransformation of arsenic by a Yellowstone thermoacidophilic eukaryotic alga. Proc. Natl. Acad. Sci. U.S.A. 106:5213-5217.

92. Raven, JA, Beardall, J, and Johnston AM. 1982. Inorganic carbon transport in relation to H+ transport at the plasmalemma of photosynthetic cells. In: Marme, D, Marre, E, Hertel, R, (eds), Plasmalemma and tonoplast: their functions in the plant cell. Elsevier Biomedical Press BB, Amsterdam, The Netherlands. pp. 41–47.

93. Redman, RS, Litvintseva, A, Sheehan, KB, Henson, JM, and Rodriguez, RJ. 1999. Fungi from geothermal soils in Yellowstone National Park. Appl. Environ. Microbiol. 65:5193-5197.

94. Reeb, V, and Bhattacharya, D. 2010. The thermo-acidophilic Cyanidiophyceae (Cyanidiales), In: Seckbach, J. and Chapman DJ, (eds) Red algae in the genomic age. Springer Science and Business Media, B.V, pp. 410-426.

95. Ryder, NA. 1990. Inhibition of squalene epoxidase and sterol side-chain methylation by allyamines. Biochem. Soc. Trans. 18:45-46.

96. Sakajiri, T., Misumi, O., Mori, T., Fujiwara, T., Hashimoto, M., and Kuroiwa, T. 2006. Monitoring expression profiles of *Cyanidioschyzon merolae* under salt and pH stress using DNA microarrays. J. Plant Res. 119 suppl. 75.

97. Sagi, M, Davydov, O, Orazova, S, Yesbergenova, Z, Ophir, R, Stratmann, JW, and Fluhr, R. 2004. Plant respiratory burst oxidase homologs impinge on wound responsiveness and development in *Lycopericon esculentum*. The Plant Cell. 16:616-628.

98. Sakajiri, T., Asano, K., Hirooka, S., Tashiro, K., Misumi, O., Fujiwara, T., and Kuroiwa, T. 2008. Microarray analysis reveals S-adenosylmethionine (SAM) synthetase involvement in salt tolerance of *Cyanidioschyzon merolae*. Cytologia. 73:341–368.

99. Saravanamuthu SS, Von Gotz F, Salunkhe P, Chozhavendan R, Geffers R, Buer J, Tummler B, and Steinmetz EJ. 2004. Evidence for polyadenylated mRNA in *Pseudomonas aeruginosa*. J Bacteriol. 186:7015–7018.

100. Schmidt, AL, and Anderson, LM. 2006. Repetitive DNA elements as mediators of genomic change in response to environmental cues. Biol. Rev. 81:531-543.

101. Schroda, M, Kropat, J, Rudiger, OW, Vallon, O, Wollman, FA, and Beck CF. 2001. Possible role for molecular chaperones in assembly and repair of photosystem II. Biochemical Society Transactions. 29:413-418.

102. Seckbach, J, Ikan, R, Nagashima, H, and Fukuda, I. 1993. New phylogenetic status for acid hot spring algae. In: Sato, S, Ishida, H, and Ishikawa, H. (eds), Endocytobiology. V. Tòbingen University Press, Tobingen, pp. 241–254.

103. Seckbach, J. 1994. The natural history of *Cyanidium* (Geitler 1933): past and present perspectives, In: Seckbach, J. (eds), Evolutionary pathways and enigmatic algae: *Cyanidium caldarium* (Rhodophyta) and related cells. Kluwer, Dordrecht, The Netherlands, pp. 99–112.

104. Seckbach, J. 1994. Evolutionary pathways and enigmatic algae: *Cyanidium caldarium* (Rhodophyta). Kluwer Academic Publishers, Dordrecht, The Netherlands. pp. 349-350.

105. Seckbach, J. 1999. The Cyanidiophyceae: hot spring acidophilic algae. In: Seckbach, J, (eds) Enigmatic microorganisms and life in extreme environments. Kluwer Academic Publishers, Dordrecht, The Netherlands. pp. 425–435.

106. Shapiro, JA, and Von Sternberg, R. 2005. Why repetitive DNA is essential to genome function. Biol. Rev. 80:227-250.

107. Skowronski, T, Szubinska, S, Pawlik, B, and Jakubowski, M. 1991. The influence of pH on cadmium toxicity to the green alga *Stichococcus bacillaris* and on cadmium forms present in the culture medium. Environmental Pollution. 74:89-100.

108. Skuja, H. 1970. Alghe cavernicole nelle zone illuminate delle Grotte di Castellana (Murge di Bari). Le Grotte d'Italia 4:193–202.

109. Sloth, J, Wiebe, M, and Eriksen, N. 2006. Accumulation of phycocyanin in heterotrophic and mixotrophic cultures of the acidophilic red alga *Galdieria sulphuraria*. Enzyme Microb. Technol. 38:168-175.

110. Smith, DW, and Brock, TD. 1973. The water relations of the alga *Cyanidium caldarium* in soil. J. Gen. Microbiol. 79:219-231.

111. Souffreau, S, Vanormelingen, P, Verleyen, E, Sabbe, K, and Vyverman. 2010. Tolerance of benthic diatoms from temperate aquatic and terrestrial habitats to experimental desiccation and temperature stress. Phycologia. 49(4):309-324.

112. Stauffer, RE, Jenne, EA, and Ball, JW. 1980. Chemical studies of selected trace elements in hot-spring drainages of Yellowstone National Park. US Geological Survey Professional Paper 1044-F.

113. Stewart, FJ, Ottesen, EA, and DeLong, ED. 2010. Development and quantitative analyses of a universal rRNA-subtraction protocol for microbial metatranscritomics. The ISME Journal. 4:896-907.

114. Sturn, A, Quackenbush, J, and Trajanoski, Z. 2002. Genesis: cluster analysis of microarray data. Bioinformatics. 18:207-208.

115. Takacs-Vesbach, MK, Jackson-Weaver, O, and Reysenbach AL. 2008. Volcanic calderas delineate biogeographic provinces among Yellowstone thermophiles. Environ. Microbiol. 10:1681-1689.

116. Taniguchi, Y, Choi, PJ, Li, GW, Chen, H, Babu, M, Hearn, J, Emili, A, and Xie, XS. 2011. Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. Science. 329:533-537.

117. Toplin, JA, Norris, TB, Lehr, CR, McDermott, TR, and Castenholz, RW. 2008. Biogeographic and phylogenetic diversity of thermoacidophilic Cyanidiales in Yellowstone National Park, Japan, and New Zealand. Appl. Environ. Microbiol. 74:2822-2833.

118. Tortell, P. 2000. Evolutionary and ecological perspectives on carbon acquisition in phytoplankton. Limnol. and Oceanogr. 45:744-750.

119. Trauger, SA, Kalisak, E, Kalisak, J, Morita, H, Weinberg, MV, Menon, AL, Poole, FL, Adams, MWW, and Siuzdak, G. 2008. Correlating the transcriptome, proteome and metabolome in the environmental adaptation of a hyperthermophile. Journal of Proteome Research. 7:1027-1035.

120. Turner, L, Houghton, JD, and Brown, SB. 1997. Purification and identification of apophycocyanin α and β subunits from soluble protein extracts of the red alga *Cyanidium caldarium*. Light exposure is not a prerequisite for biosynthesis of the protein moiety of this photosynthetic accessory pigment. Planta. 201:78-83.

121. Walker, JJ, Spear, JR, and Pace, NR. 2005. Geobiology of a microbial endolithic community in the Yellowstone geothermal environment. Nature. 434:1011-1013.

122. Walker, JJ, and Pace, NR. 2007. Endolithic microbial ecosystems. Annual Review of Microbiology. 61: 331-347.

123. Weber, AP, Oesterhelt, C, Gross, W, Brautigam, A, Imboden, LA, Krassovskaya, I, Linka, N, Truchina, J, Schneidereit, J, Voll, H, Voll, LM, Zimmermann, M, Jamai, A, Riekhof, WR, Yu, B, Garavito, RM, and Benning, C. 2004. EST-analysis of the thermo-acidophilic red microalga *Galdieria sulphuraria* reveals potential for lipid A biosynthesis and unveils the pathway of carbon export from rhodoplasts. Plant Mol. Biol. 55:17–32.

124. Whitaker, RJ, Grogan, DW, and Taylor, JW. 2003. Geographic barriers isolate endemic populations of hyperthermophilic archaea. Science 301:976–978.

125. Wiegert, RG, and Mitchell, R. 1973. Ecology of Yellowstone thermal effluent systems: intersects of blue-green algae, grazing flies (*Paracoenia*, Ephydridae) and water mites (*Partnuniella*, Hydrachnellae). Hydrobiologia. 41:251-271.

126. Wilson, KE, and Huner, PA. 2000. The role of growth rate, redox-state of the plastoquinone pool and the trans-thylakoid ΔpH in photoacclimation of *Chlorella vulgaris* to growth irradiance and temperature. Planta. 212:93-102.

127. Yokthongwattana, K, Chrost, B, Behrman, S, Casper-Lindley, C, and Melis, A. 2001. Photosystem II damage and repair cycle in the green alga, Dunaliella salina: involvement of a chloroplast-localized Hsp70. Plant Cell Physiol. 42:1389-1397.

128. Yoon HS, Hackett JD, Pinto G, Bhattacharya D. 2002. The single, ancient origin of chromist plastids. Proc. Natl. Acad. Sci. U.S.A. 99:15507–15512.

129. Yoon, HS, Hackett, JD, and Bhattacharya, D. 2002. A single origin of the peridinin- and fucoxanthin- containing plastids in the dinoflagellates through tertiary endosymbiosis. Proc. Natl. Acad. Sci. U.S.A. 99(18):11724-11729.

130. Yoon, HS, Hackkett, JD, Ciniglia, C, Pinto, G, and Bhattacharya, D. 2004. A molecular timeline for the origin of photosynthetic eukaryotes. Mol. Biol. Evol. 21:809-818.

131. Yoon, HS, Ciniglia, C, Wu, M, Comeron, JM, Pinto, G, Pollio, A, and Bhattacharya, D. 2006. Establishment of endolithic populations of extremophilic Cyanidiales (Rhodophyta). BMC Evol. Biol. 6:1-12.

132. Yoon, HS., Muller, KM, Sheath, RG., Ott, FD., and Bhattacharya, D. 2006. Defining the major lineages of red algae (Rhodophyta). J. Phycol. 42:482-492.

133. Yoshimura, E, Nagasaka, S, Satake, K, and Mori, S. 2000. Mechanism of aluminum tolerance in *Cyanidium caldarium*. Hydrobiologia. 433:57-60.

134. Zenvirth, D, Volokita, M, and Kaplan, A. 1985. Photosynthesis and inorganic carbon accumulation in the acidophilic alga *Cyanidioschyzon merolae*. Plant Physiology 77:237–239.

# APPENDICES

# APPENDIX A

## SUPPLEMENTAL INFORMATION FOR CHAPTER 3: CYANIDIALES TRANSCRIPTOMICS IN ACIDIC GEOTHERMAL ENVIRONMENTS

Table A3.1 Aqueous geochemistry of Lemonade Creek water at four time points. The average concentration is reported as the mean ± 1 SD for the four samples collected. B/D, below detection.

| Constituent | November 2009 | January 2010 | April 2010 | July 2010 | Average Concentration (mM unless otherwise noted) |
|---|---|---|---|---|---|
| *pH* | 2.78 | 3 | 2.43 | N/A | 2.74 ± 0.29 |
| *Element* | | | | | |
| Na | 2.25 | 2.08 | 2.27 | 2.01 | 2.15 ± 0.12 |
| Si | 4.69 | 4.48 | 4.73 | 4.23 | 4.53 ± 0.23 |
| K | 0.69 | 0.66 | 0.70 | 0.63 | 0.67 ± 0.03 |
| Ca | 0.212 | 0.196 | 0.192 | 0.203 | 0.201 ± 0.002 |
| Mg | 83 µM | 78 µM | 81 µM | 78 µM | 80.0 ± 2.5 µM |
| Al | 0.307 | 0.307 | 0.361 | 0.439 | 0.35 ± 0.06 |
| Fe | 57 µM | 52 µM | 54 µM | 56 µM | 54.8 ± 2.2 µM |
| As | B/D | B/D | B/D | B/D | B/D |
| P | B/D | B/D | B/D | B/D | B/D |
| Mn | 2.29 µM | 2.20 µM | 3.38 µM | 2.20µM | 2.52 ± 0.57 µM |
| Zn | 1.97 µM | 2.28 µM | 1.20 µM | 2.34 µM | 1.95 ± 0.52 µM |
| *Ion* | | | | | |
| Cl- | 0.88 | 0.56 | 0.46 | 0.664 | 0.64 ± 0.18 |
| SO42- | 7.16 | 8.18 | 7.2 | 7.53 | 7.52 ± 0.47 |
| F- | 17.3 µM | 17.2 µM | 17.3 µM | 27.6 µM | 19.9 ± 5.2 µM |

Table A3.2 Expression summary for UV-sensitive genes in January 2010 at Lemonade Creek.

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | +UV:-UV Expression Ratio |
|---|---|---|---|---|
| CMK002Z | Pseudo-gene of trefoil factor | | | 0.1 |
| CMT364X | Repetitive element | | | 0.1 |
| CMO064X | Repetitive element | | | 0.2 |
| CMQ311X | Repetitive element | | | 0.2 |
| CMQ092X | Repetitive element | | | 0.2 |
| CMJ035X | Repetitive element | | | 0.2 |
| CMO010X | Repetitive element | | | 0.2 |
| CMF183X | Repetitive element | | | 0.2 |
| CMH068X | Repetitive element | | | 0.2 |
| CMT483X | Repetitive element | | | 0.2 |
| CMF010X | Repetitive element | | | 0.2 |
| CMQ299X | Repetitive element | | | 0.2 |
| CML017X | Repetitive element | | | 0.2 |
| CMD166X | Repetitive element | | | 0.2 |
| CMI053X | Repetitive element | | | 0.2 |
| CMF005X | Repetitive element | | | 0.2 |
| CMR342X | Repetitive element | | | 0.2 |
| CMB161Z | Pseudo-gene of NADPH:protochlorophyllide oxidoreductase | | | 0.2 |
| CMJ127X | Repetitive element | | | 0.2 |
| CMT148X | Repetitive element | | | 0.2 |
| CMT011X | Repetitive element | | | 0.2 |
| CMI305X | Repetitive element | | | 0.2 |
| CMH217X | Repetitive element | | | 0.2 |
| CMQ377X | Repetitive element | | | 0.2 |
| CMO251X | Repetitive element | | | 0.2 |
| CMS382X | Repetitive element | | | 0.2 |
| CMJ007X | Repetitive element | | | 0.2 |

Table A3.2 Continued…

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | +UV:-UV Expression Ratio |
|---|---|---|---|---|
| CML337X | Repetitive element | | | 0.2 |
| CMI222X | Repetitive element | | | 0.2 |
| CMR482X | Repetitive element | | | 0.2 |
| CMT149X | Repetitive element | | | 0.2 |
| CMI011X | Repetitive element | | | 0.2 |
| CML278X | Repetitive element | | | 0.2 |
| CMM101X | Repetitive element | | | 0.2 |
| CMI109C | Similar to glutenin low molecular weight chain precursor | | Fibrinogen binding protein | 0.2 |
| CMK057X | Repetitive element | | | 0.2 |
| CMH086X | Repetitive element | | | 0.2 |
| CMQ428X | Repetitive element | | | 0.2 |
| CMQ108X | Repetitive element | | | 0.2 |
| CMN189X | Repetitive element | | | 0.2 |
| CMO143X | Repetitive element | | | 0.2 |
| CMM133X | Repetitive element | | | 0.2 |
| CMT632X | Repetitive element | | | 0.2 |
| CMS017X | Repetitive element | | | 0.2 |
| CMT039X | Repetitive element | | | 0.2 |
| CME174X | Repetitive element | | | 0.2 |
| CML241X | Repetitive element | | | 0.2 |
| CMK094X | Repetitive element | | | 0.2 |
| CMK218X | Repetitive element | | | 0.2 |
| CMM107X | Repetitive element | | | 0.2 |
| CMD128X | Repetitive element | | | 0.2 |
| CMQ400X | Repetitive element | | | 0.2 |
| CMM283X | Repetitive element | | | 0.2 |
| CMT179X | Repetitive element | | | 0.3 |
| CMS055X | Repetitive element | | | 0.3 |
| CMB163Z | Possible transcribed region | | | 0.3 |

Table A3.2 Continued…

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | +UV:-UV Expression Ratio |
|---|---|---|---|---|
| CMS069X | Repetitive element | | | 0.3 |
| CML275X | Repetitive element | | | 0.3 |
| CMD109C | Hypothetical protein, conserved | | Protein of unknown function (DUF1350) | 0.3 |
| CMT002X | Repetitive element | | | 0.3 |
| CMI115X | Repetitive element | | | 0.3 |
| CMA018X | Repetitive element | | | 0.3 |
| CMT295Z | Pseudo-gene of 3-oxo-5-alpha-steroid 4-dehydrogenase | | | 0.3 |
| CMS266X | Repetitive element | | | 0.3 |
| CMM303Z | Pseudo-gene of 3-oxo-5-alpha-steroid 4-dehydrogenase | | | 0.3 |
| CMM327X | Repetitive element | | | 0.3 |
| CMP264X | Repetitive element | | | 0.3 |
| CMP250X | Repetitive element | | | 0.3 |
| CMD189X | Repetitive element | | | 0.3 |
| CMB068X | Repetitive element | | | 0.3 |
| CMB067X | Repetitive element | | | 0.3 |
| CMR063X | Repetitive element | | | 0.3 |
| CMH033X | Repetitive element | | | 0.3 |
| CMM333X | Repetitive element | | | 0.3 |
| CMJ192X | Repetitive element | | | 0.3 |
| CMM047X | Repetitive element | | | 0.3 |
| CMP317X | Repetitive element | | | 0.3 |
| CMQ135X | Repetitive element | | | 0.3 |
| CMK310X | Repetitive element | | | 0.3 |
| CMG038X | Repetitive element | | | 0.3 |
| CMP197X | Repetitive element | | | 0.3 |
| CML062X | Repetitive element | | | 0.3 |

Table A3.2 Continued…

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | +UV:-UV Expression Ratio |
|---|---|---|---|---|
| CMO295X | Repetitive element | | | 0.3 |
| CMJ307X | Repetitive element | | | 0.3 |
| CMQ025X | Repetitive element | | | 0.3 |
| CMN281X | Repetitive element | | | 0.3 |
| CMQ439X | Repetitive element | | | 0.3 |
| CMH196X | Repetitive element | | | 0.3 |
| CMF172X | Repetitive element | | | 0.3 |
| CML208X | Repetitive element | | | 0.3 |
| CMM234X | Repetitive element | | | 0.3 |
| CMG055X | Repetitive element | | | 0.3 |
| CMP196X | Repetitive element | | | 0.3 |
| CMS003X | Repetitive element | | | 0.3 |
| CMP048X | Repetitive element | | | 0.3 |
| CMI003X | Repetitive element | | | 0.3 |
| CMN277X | Repetitive element | | | 0.3 |
| CMP082X | Repetitive element | | | 0.3 |
| CMT507X | Repetitive element | | | 0.3 |
| CMI090X | Repetitive element | | | 0.3 |
| CMH089X | Repetitive element | | | 0.3 |
| CMC073X | Repetitive element | | | 0.3 |
| CMO158X | Repetitive element | | | 0.3 |
| CMP163X | Repetitive element | | | 0.3 |
| CMR166X | Repetitive element | | | 0.3 |
| CMQ202X | Repetitive element | | | 0.3 |
| CMB137X | Repetitive element | | | 0.3 |

Table A3.2 Continued…

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | +UV:-UV Expression Ratio |
|---|---|---|---|---|
| CML233X | Repetitive element | | | 0.3 |
| CMT005X | Repetitive element | | | 0.3 |
| CMS016X | Repetitive element | | | 0.3 |
| CMM168X | Repetitive element | | | 0.3 |
| CMF008X | Repetitive element | | | 0.3 |
| CMR430X | Repetitive element | | | 0.3 |
| CMQ262X | Repetitive element | | | 0.3 |
| CMF013X | Repetitive element | | | 0.3 |
| CMT007X | Repetitive element | | | 0.3 |
| CMB160X | Repetitive element | | | 0.3 |
| CMP242X | Repetitive element | | | 0.3 |
| CMM312Z | Possible transcribed region | | | 0.3 |
| CMT509X | Repetitive element | | | 0.3 |
| CMT557X | Repetitive element | | | 0.3 |
| CMI005X | Repetitive element | | | 0.3 |
| CMN020X | Repetitive element | | | 0.3 |
| CMT310X | Repetitive element | | | 0.3 |
| CMS072X | Repetitive element | | | 0.4 |
| CMO002X | Repetitive element | | | 0.4 |
| CMR481X | Repetitive element | | | 0.4 |
| CMO148C | Transposon | | | 0.4 |
| CMO156X | Repetitive element | | | 0.4 |
| CMP279X | Repetitive element | | | 0.4 |
| CMG208C | Retroelement | | Reverse transcriptase | 0.4 |
| CMP121C | Retroelement, dead | | Reverse transcriptase | 0.4 |

Table A3.2 Continued…

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | +UV:-UV Expression Ratio |
|---|---|---|---|---|
| CMM301Z | Possible transcribed region | | | 0.4 |
| CMN015X | Repetitive element | | | 0.4 |
| CMT348Z | Possible transcribed region | | | 0.4 |
| CMP153X | Repetitive element | | | 0.4 |
| CMP132X | Repetitive element | | | 0.4 |
| CMI235X | Repetitive element | | | 0.4 |
| CMQ036X | Repetitive element | | | 0.4 |
| CMD064X | Repetitive element | | | 0.4 |
| CMH098X | Repetitive element | | | 0.4 |
| CMS186X | Repetitive element | | | 0.4 |
| CMT189X | Repetitive element | | | 0.4 |
| CMI269X | Repetitive element | | | 0.4 |
| CMJ003X | Repetitive element | | | 0.4 |
| CMF145X | Repetitive element | | | 0.4 |
| CMN093T | Hypothetical transcript | | | 0.4 |
| CMC187X | Repetitive element | | | 0.4 |
| CMI190X | Repetitive element | | | 0.4 |
| CMT040X | Repetitive element | | | 0.4 |
| CMS196X | Repetitive element | | | 0.4 |
| CME003X | Repetitive element | | | 0.4 |
| CMK001C | Similar to hedgehog protein | | Hint module | 0.4 |

Table A3.2 Continued…

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | +UV:-UV Expression Ratio |
|---|---|---|---|---|
| CMR016X | Repetitive element | | | 0.4 |
| CMD191C | Similar to trefoil factor | | Trefoil (P-type) domain | 0.5 |
| CMA143X | Repetitive element | | | 0.5 |
| CMR195Z | Possible transcribed region | | | 0.5 |
| CMH104Z | Possible transcribed region | | | 0.5 |
| CMB162C | Similar to trefoil factor | | Trefoil (P-type) domain | 0.5 |
| CME002Z | Pseudo-gene of NADPH:protochlorophyllide | | | 0.5 |
| CML281X | Repetitive element | | | 0.5 |
| CMF002Z | Pseudo-gene of trefoil factor | | | 0.5 |
| CMP120X | Repetitive element | | | 0.5 |
| CMK165C | Retroelement, alive | | Reverse transcriptase | 0.5 |
| CMT044X | Repetitive element | | | 0.5 |
| CMN009C | Hypothetical protein | | | 0.5 |
| CMP241Z | Possible transcribed region | | | 0.5 |
| CMD178X | Repetitive element | | | 0.5 |

Table A3.2 Continued…

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | +UV:-UV Expression Ratio |
|---|---|---|---|---|
| CMV192C | 30S ribosomal protein S10 | | Ribosomal protein S10p/S20e | 2.0 |
| CMV153C | Pyruvate dehydrogenase E1 component alpha | | Dehydrogenase E1 component | 2.1 |
| CMW025C | NADH-ubiquinone oxidoreductase chain 5 | Oxidative Phosphorylation | NADH-Ubiquinone/plastoquinone (complex I), various chains | 2.1 |
| CMV113C | Naphthoate synthase | | Enoyl-CoA hydratase/isomerase family | 2.1 |
| CMV100C | Molybdopterin biosynthesis MoeB protein | | ThiF family | 2.3 |
| CMV117C | 4-(2'-carboxyphenyl)-4-oxybutyric acid synthase | | | 2.6 |
| CMW013C | 50S ribosomal protein L6 | | Ribosomal protein L6 | 2.6 |
| CMV012C | Thioredoxin type m | | Thioredoxin | 2.7 |
| CMW037C | 50S ribosomal protein L14 | | Ribosomal protein L14p/L23e | 2.8 |
| CMI008T | Hypothetical transcript | | | 2.9 |
| CMV020C | Probable ABC transporter ATP-binding protein YhbG | | ABC transporter | 2.9 |
| CMV127C | Photosystem II 10 kDa phosphoprotein | Photosynthesis | Photosystem II 10 kDa phosphoprotein | 3.2 |
| CMW060C | NADH-ubiquinone oxidoreductase chain 4L | Oxidative Phosphorylation | NADH-ubiquinone/plastoquinone oxidoreductase | 3.5 |
| CMW027C | ATP synthase A chain (protein 6) | Oxidative Phosphorylation | ATP synthase A chain | 3.7 |

Table A3.3 Expression summary for UV-sensitive genes in July 2010 at Lemonade Creek.

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | +UV:-UV Expression Ratio |
|---|---|---|---|---|
| CMR353C | Similar to respiratory burst oxidase protein | | FAD-binding domain | 2.2 |
| CME043C | Hypothetical protein, conserved | | CCT motif | 2.4 |
| CMK034C | Hypothetical protein | | Helix-loop-helix DNA-binding domain | 3.1 |

Table A3.4 Expression summary of UV-sensitive genes observed in April 2010 at Lemonade Creek. Included are fold change values for those genes shared between April 2010 and either January 2010 or July 2010. Heat map color depiction visually illustrates a statistically significant positive (induction, red) or negative (repression, green) expression ratios that were ± 2.

| | | KEGG Functional | | +UV:-UV Expression Ratio | | |
|---|---|---|---|---|---|---|
| Gene ID | Annotation | Group | Pfam Domain | Apr. | Jul. | Jan. |
| CMW053C | 30S ribosomal protein S12 (rps12) | Ribosome | Ribosomal protein S12 | 0.3 | | |
| CMK262C | Hypothetical protein | | | 0.4 | | |
| CMV009C | 30S ribosomal protein S4 (rps4) | Ribosome | Ribosomal protein S4/S9 N-terminal domain | 0.4 | | |
| CMV132C | Acyl carrier protein | | Phosphopantetheine attachment site | 0.4 | | |
| CMV062C | Glucosamine-fructose-6-phosphate aminotransferase | | Glutamine amidotransferases class-II | 0.4 | | |
| CMV007C | Initiation factor 2 (infB) | | Elongation factor Tu domain 2 | 0.4 | | |
| CMV046C | 50S ribosomal protein L32 (rpl32) | Ribosome | Ribosomal L32p protein family | 0.4 | | |
| CMV176C | 50S ribosomal protein L5 (rpl5) | Ribosome | Ribosomal protein L5 | 0.5 | | |
| CMP311C | TBP-1 interacting protein | | Syntaxin | 0.5 | | |
| CMV221C | ATP synthase CF0 C chain (subunit III) | Oxidative phosphorylation | ATP synthase subunit C | 0.5 | | |
| CMF163C | Similar to heterogeneous nuclear ribonucleoprotein H3, isoform a | | RNA recognition motif | 2.0 | | |
| CMI009T | Hypothetical transcript | | | 2.0 | | |
| CMS309C | Similar to microsomal glutathione S-transferase | Glutathione metabolism | MAPEG family | 2.2 | | |
| CML084Z | Possible transcribed region | | | 2.2 | | |
| CMR353C | Similar to respiratory burst oxidase protein | | FAD-binding domain | 2.3 | 2.2 | |
| CMS349C | Hypothetical protein | | | 2.4 | | |
| CMK034C | Hypothetical protein | | Helix-loop-helix DNA-binding domain | 2.7 | 3.2 | |

Table A3.5 Expression summary for UV-sensitive genes observed in November 2009 at Lemonade Creek. Included are fold change values for those genes shared between November 2009 and either January 2010 or July 2010. Heat map color depiction visually illustrates a statistically significant positive (induction, red) or negative (repression, green) expression ratios that were ± 2.

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | +UV:-UV Expression Ratio | | |
|---|---|---|---|---|---|---|
| | | | | Nov. 2009 | Jul. 2010 | Jan. 2010 |
| CMV139C | Cytochrome b6-f complex subunit V (petG) | Photosynthesis | Cytochrome B6-F complex subunit 5 | 0.1 | | |
| CMV213C | 30S ribosomal protein S18 (rps18) | Ribosome | Ribosomal protein S18 | 0.1 | | |
| CMW019C | Cytochrome c oxidase polypeptide I (cox1) | Oxidative phosphorylation | Cytochrome C and Quinol oxidase polypeptide I | 0.1 | | |
| CMW010C | Cytochrome c oxidase polypeptide II (cox2) | Oxidative phosphorylation | Cytochrome C oxidase subunit II, periplasmic domain | 0.1 | | |
| CMW007C | Cytochrome B (cytB) | Oxidative phosphorylation | Cytochrome b(C-terminal)/b6/petD | 0.1 | | |
| CMW002C | Succinate dehydrogenase iron-sulfur protein (sdhB) | Oxidative phosphorylation | 2Fe-2S iron-sulfur cluster binding domain | 0.2 | | |
| CMW021C | NADH-ubiquinone oxidoreductase chain 1 (nad1) | Oxidative phosphorylation | NADH dehydrogenase | 0.2 | | |
| CMV038C | Remnant of ycf55 | | | 0.2 | | |
| CMK169Z | Possible transcribed region | | | 0.3 | | |
| CMV004C | Cell-cycle protein mesJ family | | PP-loop family | 0.3 | | |
| CMV037C | Probable tetrapyrrole-binding protein GUN4 | | GUN4-like | 0.4 | | |
| CMO262Z | Possible transcribed region | | | 0.4 | | |
| CMC014Z | Possible transcribed region | | | 0.4 | | |
| CMK168C | Hypothetical protein | | | 0.4 | | |
| CMV147C | Cytochrome b6-f complex subunit VII (petM) | Photosynthesis | Cytochrome oxidase maturation protein cbb3-type | 0.5 | | |

Table A3.5 Continued…

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | +UV:-UV Expression Ratio Nov. 2009 | Jul. 2010 | Jan. 2010 |
|---|---|---|---|---|---|---|
| CMH112Z | Possible transcribed region | | | 0.5 | | |
| CMD116C | Hypothetical protein, conserved | | | 0.5 | | |
| CMT581C | Hypothetical protein, conserved | | Nucleotidyltransferase domain | 2.0 | | |
| CML304C | 60S ribosomal protein L35A (rpl35A) | Ribosome | Ribosomal protein L35Ae | 2.0 | | |
| CMT589Z | Possible transcribed region | | | 2.0 | | |
| CML311C | Hypothetical protein, conserved | | Conserved region of Rad21 / Rec8 like protein | 2.0 | | |
| CMK273C | 60S ribosomal protein L23A (rpl23A) | Ribosome | Ribosomal protein L23 | 2.0 | | |
| CMP339C | Similar to Golgi vesicular membrane trafficking protein Bet1p | SNARE Interactions | SlyX | 2.0 | | |
| CMI152C | Hypothetical protein, conserved | | Anaphase-promoting complex, cyclosome, subunit 3 | 2.0 | | |
| CMK009C | Hypothetical protein | | NUMOD3 motif (2 copies) | 2.1 | | |
| CMS212C | Mitochondrial ribosomal protein S17 precursor (rpsQ) | | Ribosomal protein S17 | 2.1 | | |
| CML015C | Similar to ferredoxin | Photosynthesis | 2Fe-2S iron-sulfur cluster binding domain | 2.1 | | |
| CMR465C | Sulfate ABC transporter ATP-binding subunit cysA | ABC transporters | ABC transporter | 2.1 | | |
| CML303Z | Possible transcribed region | | | 2.1 | | |
| CMO184C | Similar to hedgehog protein | | Hint module | 2.1 | | |
| CMS027C | Hypothetical protein, conserved | | CBS domain | 2.2 | | |
| CMI218C | Hypothetical protein, conserved | | 26S proteasome subunit RPN7 | 2.2 | | |

Table A3.5 Continued…

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | +UV:-UV Expression Ratio Nov. 2009 | Jul. 2010 | Jan. 2010 |
|---|---|---|---|---|---|---|
| CMO152C | Signal recognition particle recognition component SRP54 | Protein Export | SRP54-type protein, GTPase domain | 2.2 | | |
| CMF159C | AAA-metalloprotease FtsH, mitochondrial precursor | | ATPase family associated with various cellular activities (AAA) | 2.2 | | |
| CMQ072C | Hypothetical protein | | | 2.2 | | |
| CME134C | Hypothetical protein | | Myb-like DNA-binding domain | 2.2 | | |
| CMP056Z | Possible transcribed region | | | 2.2 | | |
| CMF025C | Hypothetical protein | | | 2.2 | | |
| CMP290C | Hypothetical protein | | | 2.2 | | |
| CMT167C | Hypothetical protein, conserved | | AhpC/TSA family | 2.2 | | |
| CMT390C | Hypothetical protein, conserved | | RNA pseudouridylate synthase | 2.3 | | |
| CMP189C | Probable selenophosphate synthetase | Selenocompound metabolism | AIR synthase related protein, N-terminal domain | 2.3 | | |
| CMA031C | Hypothetical protein | | | 2.3 | | |
| CMA044C | Cryptochrome DASH | | | 2.3 | | |
| CMT304C | Hypothetical protein, conserved | | Aldo/keto reductase family | 2.3 | | |
| CMN046C | Sporulation-induced transcript 4-associated protein SAPLb | | SIT4 phosphatase-associated protein | 2.3 | | |
| CMM075T | Hypothetical transcript | | | 2.3 | | |
| CMK106C | Hypothetical protein, conserved | | E2 (early) protein, N terminal | 2.3 | | |
| CMP134C | Similar to transcriptional adaptor like protein | | Myb-like DNA-binding domain | 2.3 | | |
| CMK262C | Hypothetical protein | | | 2.3 | | |
| CMD051C | Hypothetical protein | | | 2.4 | | |
| CME043C | Hypothetical protein, conserved | | CCT motif | 2.4 | 2.4 | |
| CML294C | Hypothetical protein | | 2',5' RNA ligase family | 2.4 | | |
| CMT237Z | Possible transcribed region | | | 2.4 | | |
| CMB097C | Hypothetical protein | | | 2.4 | | |

Table A3.5 Continued…

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | +UV:-UV Expression Ratio Nov. 2009 | Jul. 2010 | Jan. 2010 |
|---|---|---|---|---|---|---|
| CMA025C | Hypothetical protein | | | 2.5 | | |
| CMN167C | RIO-like serine/threonine protein kinase (Rio1) | Ribosome biogenesis in eukaryotes | RIO1 family | 2.6 | | |
| CMS343C | Molecular chaperone of Hsp110 family | | Hsp70 protein | 2.6 | | |
| CMT223C | Similar to translation initiation factor IF-3 | | Translation initiation factor IF-3, C-terminal domain | 2.6 | | |
| CMQ412C | Similar to psbB mRNA maturation factor Mbb1 | | SecE/Sec61-gamma subunits of protein translocation complex | 2.6 | | |
| CMI216C | Probable mitochondrial intermembrane space complex subunit Tim8 | | Tim10/DDP family zinc finger | 2.6 | | |
| CME019C | Dynamin-related protein involved in mitochondrial division CmDnm1/DRP3 | Endocytosis | | 2.6 | | |
| CMS360C | Hypothetical protein, conserved | | | 2.6 | | |
| CMM200C | Probable DNA binding protein | | Metallopeptidase family M24 | 2.6 | | |
| CMM132C | Mitochondrial propionyl-CoA carboxylase, beta subunit, precursor | Glyoxylate and dicarboxylate metabolism | Carboxyl transferase domain | 2.7 | | |
| CMR305C | Similar to pleiotropic regulator 1 (PRL1) | | WD domain, G-beta repeat | 2.8 | | |
| CMG129C | MYB-related protein | | Myb-like DNA-binding domain | 2.8 | | |
| CMT321C | Probable eukaryotic translation initiation factor (eIF-1) | RNA transport | Translation initiation factor SUI1 | 2.8 | | |
| CMN119T | Hypothetical transcript | | | 2.9 | | |
| CMT033Z | Possible transcribed region | | | 2.9 | | |
| CMJ046C | Similar to kinetoplast-associated protein | | CHASE4 domain | 2.9 | | |
| CMP277C | Mitochondrial ribosomal protein L2 precursor | | Ribosomal Proteins L2, RNA binding domain | 2.9 | | |
| CMI252C | DNA topoisomerase I | | ABC transporter | 3.1 | | |
| CMI042C | Tyrosine--tRNA ligase | Aminoacyl-tRNA biosynthesis | Glutathione S-transferase, C-terminal domain | 3.1 | | |

Table A3.5 Continued…

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | +UV:-UV Expression Ratio Nov. 2009 | Jul. 2010 | Jan. 2010 |
|---|---|---|---|---|---|---|
| CMB111Z | Possible transcribed region | | | 3.2 | | |
| CMJ237C | Hypothetical protein, conserved | | PurA ssDNA and RNA-binding protein | 3.7 | | |
| CMA029C | Hypothetical protein | | | 4.0 | | |
| CMP012C | 60S ribosomal protein L34 (rpl34) | Ribosome | Ribosomal protein L34e | 8.2 | | |

Table A3.6 Cluster A: Lemonade Creek gene expression patterns as a function of UV-exposure at four time points. *Red* and *green* colors depict significant expression changes, either an increase or decrease in the +UV:-UV expression ratios under the UV-exposed cyanidial mats. Heat map color depiction visually illustrates a statistically significant positive (induction, red) or negative (repression, green) expression ratios that were ± 2.

| | | | | +UV:-UV Expression Ratio | | | |
|---|---|---|---|---|---|---|---|
| **Gene ID** | **Annotation** | **KEGG Functional Pathway** | **Pfam Domain** | **Nov. 2009** | **Jan. 2010** | **Apr. 2010** | **Jul. 2010** |
| CMV139C | Cytochrome b6-f complex subunit V | Photosynthesis | Cytochrome B6-F complex subunit 5 | 0.09 | 0.54 | 1.54 | 0.47 |
| CMT148X | Hypothetical protein | | | 0.15 | 0.14 | 2.28 | 1.05 |
| CMR430X | Repetitive element | | | 0.27 | 0.19 | 1.93 | 0.77 |
| CML278X | Repetitive element | | | 0.31 | 0.33 | 3.08 | 0.71 |
| CMD166X | Repetitive element | | | 0.32 | 0.22 | 2.65 | 0.66 |
| CMF145X | Retroelement | | Reverse transcriptase | 0.32 | 0.37 | 2.62 | 0.62 |
| CMQ299X | Repetitive element | | | 0.33 | 0.42 | 2.06 | 0.77 |
| CMQ135X | Repetitive element | | | 0.34 | 0.18 | 2.49 | 0.79 |
| CMN009C | Repetitive element | | | 0.36 | 0.18 | 2.53 | 0.82 |
| CMI109C | Hypothetical protein | | | 0.38 | 0.48 | 2.33 | 0.75 |
| CMM107X | Similar to glutenin low molecular weight chain precursor | | Fibrinogen binding protein | 0.39 | 0.22 | 1.46 | 1.27 |
| CMO262Z | Possible transcribed region | | | 0.39 | 0.72 | 1.48 | 0.76 |
| CMK218X | Repetitive element | | | 0.39 | 0.25 | 2.05 | 0.82 |
| CMP197X | Repetitive element | | | 0.39 | 0.23 | 2.29 | 0.74 |
| CMM047X | Repetitive element | | | 0.40 | 0.25 | 2.43 | 0.70 |
| CMR342X | Repetitive element | | | 0.40 | 0.29 | 2.23 | 0.77 |
| CMS069X | Repetitive element | | | 0.41 | 0.28 | 1.92 | 0.66 |
| CMT011X | Repetitive element | | | 0.41 | 0.25 | 2.03 | 0.81 |
| CML017X | Repetitive element | | | 0.42 | 0.18 | 3.02 | 0.72 |
| CMT348Z | Possible transcribed region | | | 0.42 | 0.39 | 1.30 | 0.94 |
| CMK001C | Similar to hedgehog protein | | Hint module | 0.43 | 0.44 | 1.26 | 0.97 |

Table A3.6 Continued…

| Gene ID | Annotation | KEGG Functional Pathway | Pfam Domain | +UV:-UV Expression Ratio | | | |
|---|---|---|---|---|---|---|---|
| | | | | Nov. 2009 | Jan. 2010 | Apr. 2010 | Jul. 2010 |
| CMT040X | Repetitive element | | | 0.44 | 0.43 | 1.38 | 0.66 |
| CMT179X | Repetitive element | | | 0.44 | 0.17 | 1.91 | 1.04 |
| CMM133X | Repetitive element | | | 0.45 | 0.20 | 1.76 | 1.03 |
| CMJ127X | Repetitive element | | | 0.45 | 0.33 | 2.36 | 0.61 |
| CMF008X | Repetitive element | | | 0.45 | 0.34 | 1.99 | 0.72 |
| CMT632X | Repetitive element | | | 0.45 | 0.26 | 1.43 | 1.05 |
| CMR481X | Possible transcribed region | | | 0.46 | 0.25 | 1.74 | 0.80 |
| CMT149X | Repetitive element | | | 0.46 | 0.25 | 1.79 | 0.61 |
| CMR482X | Repetitive element | | | 0.46 | 0.24 | 1.85 | 0.84 |
| CMT007X | Repetitive element | | | 0.46 | 0.19 | 2.07 | 0.85 |
| CML062X | Repetitive element | | | 0.46 | 0.33 | 1.77 | 0.82 |
| CMQ311X | Repetitive element | | | 0.46 | 0.29 | 2.02 | 0.90 |
| CMH217X | Repetitive element | | | 0.46 | 0.24 | 2.24 | 0.82 |
| CMM283X | Repetitive element | | | 0.47 | 0.23 | 1.94 | 0.87 |
| CMS016X | Repetitive element | | | 0.47 | 0.36 | 3.42 | 0.80 |
| CMO064X | Repetitive element | | | 0.48 | 0.21 | 1.76 | 0.99 |
| CMF172X | Repetitive element | | | 0.48 | 0.21 | 1.95 | 0.97 |
| CMI305X | Repetitive element | | | 0.48 | 0.34 | 1.48 | 0.83 |
| CMT364X | Repetitive element | | | 0.48 | 0.29 | 1.57 | 0.81 |
| CMB068X | Repetitive element | | | 0.48 | 0.16 | 1.89 | 0.91 |
| CMR016X | Repetitive element | | | 0.48 | 0.45 | 1.16 | 0.85 |
| CMJ035X | Repetitive element | | | 0.49 | 0.20 | 2.41 | 0.96 |
| CMT189X | Repetitive element | | | 0.49 | 0.41 | 1.48 | 1.02 |
| CMN277X | Repetitive element | | | 0.49 | 0.25 | 1.81 | 1.11 |

Table A3.6 Continued…

| Gene ID | Annotation | KEGG Functional Pathway | Pfam Domain | +UV:-UV Expression Ratio | | | |
|---|---|---|---|---|---|---|---|
| | | | | Nov. 2009 | Jan. 2010 | Apr. 2010 | Jul. 2010 |
| CMM101X | Repetitive element | | | 0.49 | 0.26 | 2.01 | 0.80 |
| CML241X | Repetitive element | | | 0.50 | 0.24 | 2.11 | 0.87 |
| CMH228X | Repetitive element | | | 0.50 | 0.33 | 1.79 | 0.80 |
| CMJ307X | Repetitive element | | | 0.50 | 0.29 | 2.03 | 0.97 |
| CMA018X | Repetitive element | | | 0.50 | 0.15 | 2.23 | 0.83 |
| CMO002X | Repetitive element | | | 0.51 | 0.29 | 1.76 | 0.89 |
| CMP120X | Repetitive element | | | 0.51 | 0.48 | 1.72 | 0.63 |
| CMI011X | Repetitive element | | | 0.51 | 0.20 | 2.31 | 0.70 |
| CMT556X | Repetitive element | | | 0.51 | 0.15 | 2.74 | 0.87 |
| CML208X | Repetitive element | | | 0.52 | 0.29 | 1.02 | 1.02 |
| CMT002X | Repetitive element | | | 0.52 | 0.17 | 2.32 | 0.81 |
| CMP264X | Repetitive element | | | 0.52 | 0.35 | 2.12 | 0.70 |
| CMH086X | Repetitive element | | | 0.52 | 0.31 | 1.47 | 0.71 |
| CMF183X | Repetitive element | | | 0.52 | 0.22 | 2.27 | 0.76 |
| CMF013X | Repetitive element | | | 0.53 | 0.18 | 1.78 | 0.99 |
| CMK165C | Retroelement, alive | | Reverse transcriptase | 0.53 | 0.48 | 1.76 | 0.73 |
| CMM234X | Repetitive element | | | 0.53 | 0.22 | 1.54 | 0.87 |
| CMN189X | Repetitive element | | | 0.53 | 0.35 | 2.01 | 1.05 |
| CMT507X | Repetitive element | | | 0.53 | 0.31 | 1.37 | 1.18 |
| CMK094X | Repetitive element | | | 0.54 | 0.25 | 1.95 | 0.72 |
| CMB162C | Similar to trefoil factor | | Trefoil (P-type) domain | 0.56 | 0.46 | 1.57 | 0.66 |
| CMQ202X | Repetitive element | | | 0.57 | 0.32 | 1.48 | 0.61 |
| CMR195Z | Possible transcribed region | | | 0.59 | 0.45 | 1.18 | 1.07 |
| CMI235X | Repetitive element | | | 0.60 | 0.40 | 1.93 | 0.74 |
| CME002Z | Pseudo-gene of NADPH:protochlorophyllide oxidoreductase | | | 0.60 | 0.46 | 2.24 | 0.72 |
| CMK310X | Repetitive element | | | 0.61 | 0.28 | 1.10 | 1.01 |
| CMC073X | Repetitive element | | | 0.61 | 0.32 | 1.61 | 0.85 |
| CMS266X | Repetitive element | | | 0.62 | 0.26 | 1.92 | 0.81 |
| CMB137X | Repetitive element | | | 0.62 | 0.33 | 1.32 | 1.01 |
| CMJ007X | Repetitive element | | | 0.63 | 0.21 | 1.48 | 0.90 |
| CMI222X | Repetitive element | | | 0.63 | 0.21 | 1.66 | 0.94 |

Table A3.6 Continued…

| Gene ID | Annotation | KEGG Functional Pathway | Pfam Domain | +UV:-UV Expression Ratio | | | |
|---|---|---|---|---|---|---|---|
| | | | | Nov. 2009 | Jan. 2010 | Apr. 2010 | Jul. 2010 |
| CMF010X | Repetitive element | | | 0.63 | 0.18 | 1.96 | 0.78 |
| CMJ192X | Repetitive element | | | 0.64 | 0.28 | 1.44 | 0.86 |
| CMD191C | Similar to trefoil factor | | Trefoil (P-type) domain | 0.64 | 0.45 | 1.66 | 0.66 |
| CMP163X | Repetitive element | | | 0.64 | 0.32 | 1.40 | 1.01 |
| CMO010X | Repetitive element | | | 0.64 | 0.17 | 2.12 | 0.73 |
| CMD128X | Repetitive element | | | 0.65 | 0.25 | 1.68 | 0.74 |
| CMR166X | Repetitive element | | | 0.65 | 0.32 | 1.14 | 0.94 |
| CMI053X | Repetitive element | | | 0.65 | 0.19 | 1.87 | 0.79 |
| CMS017X | Repetitive element | | | 0.65 | 0.24 | 1.80 | 0.89 |
| CMM333X | Repetitive element | | | 0.65 | 0.28 | 1.49 | 0.84 |
| CME003X | Repetitive element | | | 0.66 | 0.44 | 1.19 | 0.99 |
| CMH196X | Repetitive element | | | 0.66 | 0.29 | 1.48 | 0.96 |
| CMH089X | Repetitive element | | | 0.66 | 0.31 | 1.64 | 0.98 |
| CMQ400X | Repetitive element | | | 0.67 | 0.25 | 1.95 | 1.06 |
| CML281X | Repetitive element | | | 0.67 | 0.47 | 1.43 | 0.84 |
| CMI269X | Repetitive element | | | 0.67 | 0.42 | 1.28 | 0.81 |
| CMO295X | Repetitive element | | | 0.68 | 0.29 | 2.06 | 0.79 |
| CMO251X | Repetitive element | | | 0.68 | 0.20 | 1.64 | 1.25 |
| CMB067X | Repetitive element | | | 0.68 | 0.27 | 1.51 | 0.90 |
| CMB161Z | Pseudo-gene of NADPH:protochlorophyllide oxidoreductase | | | 0.68 | 0.19 | 2.01 | 1.03 |
| CMD178X | Repetitive element | | | 0.68 | 0.49 | 1.08 | 0.86 |
| CMO148C | Transposon | | | 0.68 | 0.36 | 1.44 | 1.16 |
| CMN093T | Hypothetical transcript | | | 0.68 | 0.42 | 1.22 | 0.99 |
| CML275X | Repetitive element | | | 0.69 | 0.25 | 1.71 | 0.86 |
| CMT310X | Repetitive element | | | 0.69 | 0.34 | 2.08 | 0.76 |
| CMI190X | Repetitive element | | | 0.70 | 0.43 | 1.19 | 0.91 |
| CMH098X | Repetitive element | | | 0.70 | 0.41 | 1.52 | 0.88 |
| CMP082X | Repetitive element | | | 0.70 | 0.31 | 1.69 | 1.13 |

Table A3.6 Continued…

| Gene ID | Annotation | KEGG Functional Pathway | Pfam Domain | +UV:-UV Expression Ratio | | | |
|---|---|---|---|---|---|---|---|
| | | | | Nov. 2009 | Jan. 2010 | Apr. 2010 | Jul. 2010 |
| CMT295Z | Pseudo-gene of 3-oxo-5-alpha-steroid 4-dehydrogenase | | | 0.71 | 0.26 | 1.79 | 1.05 |
| CML018X | Repetitive element | | | 0.71 | 0.25 | 1.74 | 1.08 |
| CMC187X | Repetitive element | | | 0.71 | 0.43 | 1.32 | 1.10 |
| CMP279X | Repetitive element | | | 0.72 | 0.37 | 1.86 | 0.90 |
| CMT509X | Repetitive element | | | 0.73 | 0.34 | 1.28 | 1.09 |
| CMM327X | Repetitive element | | | 0.74 | 0.27 | 1.19 | 0.89 |
| CMT039X | Repetitive element | | | 0.74 | 0.24 | 1.41 | 1.01 |
| CMO158X | Repetitive element | | | 0.75 | 0.32 | 1.45 | 0.90 |
| CMQ036X | Repetitive element | | | 0.75 | 0.41 | 1.78 | 0.76 |
| CMD109C | Hypothetical protein, conserved | | Protein of unknown function (DUF1350) | 0.76 | 0.26 | 2.30 | 1.33 |
| CMN020X | Repetitive element | | | 0.76 | 0.34 | 0.99 | 0.99 |
| CMP153X | Repetitive element | | | 0.77 | 0.39 | 1.46 | 0.72 |
| CMQ377X | Repetitive element | | | 0.77 | 0.20 | 1.93 | 0.75 |
| CMM312Z | Possible transcribed region | | | 0.78 | 0.34 | 1.81 | 0.99 |
| CMQ092X | Repetitive element | | | 0.78 | 0.17 | 1.66 | 0.89 |
| CMQ439X | Repetitive element | | | 0.79 | 0.29 | 1.55 | 0.89 |
| CMH033X | Repetitive element | | | 0.79 | 0.28 | 1.44 | 0.99 |
| CME174X | Repetitive element | | | 0.80 | 0.24 | 2.15 | 0.89 |
| CML233X | Repetitive element | | | 0.80 | 0.33 | 1.52 | 0.81 |
| CMP250X | Repetitive element | | | 0.81 | 0.27 | 1.86 | 0.89 |
| CMP241Z | Possible transcribed region | | | 0.82 | 0.48 | 1.61 | 0.93 |
| CMA143X | Repetitive element | | | 0.82 | 0.45 | 1.27 | 1.00 |
| CMP317X | Repetitive element | | | 0.82 | 0.28 | 1.80 | 0.98 |
| CMI005X | Repetitive element | | | 0.83 | 0.34 | 1.72 | 0.62 |

Table A3.6 Continued…

| Gene ID | Annotation | KEGG Functional Pathway | Pfam Domain | +UV:-UV Expression Ratio | | | |
|---|---|---|---|---|---|---|---|
| | | | | Nov. 2009 | Jan. 2010 | Apr. 2010 | Jul. 2010 |
| CMK057X | Repetitive element | | | 0.83 | 0.23 | 1.71 | 0.97 |
| CMP121C | Retroelement, dead | | Reverse transcriptase | 0.84 | 0.37 | 2.09 | 0.74 |
| CMD189X | Repetitive element | | | 0.84 | 0.27 | 1.50 | 0.97 |
| CMD064X | Repetitive element | | | 0.85 | 0.41 | 1.37 | 1.13 |
| CMO156X | Repetitive element | | | 0.85 | 0.37 | 1.42 | 0.95 |
| CMG055X | Repetitive element | | | 0.87 | 0.30 | 1.70 | 0.87 |
| CMK002Z | Pseudo-gene of trefoil factor | | | 0.87 | 0.11 | 2.24 | 0.91 |
| CMQ025X | Repetitive element | | | 0.87 | 0.29 | 2.01 | 0.96 |
| CMS055X | Repetitive element | | | 0.89 | 0.25 | 1.43 | 1.00 |
| CML337X | Repetitive element | | | 0.91 | 0.21 | 1.73 | 1.08 |
| CMT044X | Repetitive element | | | 0.93 | 0.48 | 1.43 | 1.05 |
| CMJ003X | Repetitive element | | | 0.94 | 0.42 | 1.14 | 0.99 |
| CMT557X | Repetitive element | | | 0.97 | 0.34 | 1.29 | 0.85 |
| CMH104Z | Possible transcribed region | | | 0.98 | 0.46 | 1.04 | 0.88 |
| CMN015X | Repetitive element | | | 0.99 | 0.39 | 1.32 | 0.88 |
| CMM301Z | Possible transcribed region | | | 0.99 | 0.39 | 2.15 | 0.78 |
| CMS186X | Repetitive element | | | 1.01 | 0.41 | 1.20 | 1.23 |
| CMI090X | Repetitive element | | | 1.04 | 0.31 | 2.04 | 0.49 |
| CMI003X | Repetitive element | | | 1.07 | 0.31 | 1.51 | 0.86 |
| CMP048X | Repetitive element | | | 1.09 | 0.30 | 1.62 | 1.21 |
| CMP132X | Repetitive element | | | 1.11 | 0.40 | 1.17 | 0.86 |
| CMT005X | Repetitive element | | | 1.13 | 0.33 | 1.53 | 0.88 |
| CMS196X | Repetitive element | | | 1.17 | 0.44 | 1.01 | 1.25 |
| CMF002Z | Pseudo-gene of trefoil factor | | | 1.28 | 0.47 | 0.94 | 0.88 |

Table A3.7 Cluster B: Lemonade Creek gene expression patterns as a function of UV-exposure at four time points. *Red* and *green* colors depict significant expression changes, either an increase or decrease in the +UV:-UV expression ratios under the UV-exposed cyanidial mats. Heat map color depiction visually illustrates a statistically significant positive (induction, red) or negative (repression, green) expression ratios that were ± 2.

| | | | | +UV:-UV Expression Ratio | | | |
|---|---|---|---|---|---|---|---|
| **Gene ID** | **Annotation** | **KEGG Functional Pathway** | **Pfam Domain** | **Nov. 2009** | **Jan. 2010** | **Apr. 2010** | **Jul. 2010** |
| CMI009T | Hypothetical transcript | | | 1.05 | 1.62 | 2.04 | 1.41 |
| CMV221C | ATP synthase CF0 C chain (subunit III) | Photosynthesis | ATP synthase subunit C | 1.21 | 1.92 | 0.47 | 0.87 |
| CMI008T | Hypothetical transcript | | | 1.45 | 2.9 | 2.05 | 1.44 |
| CMR353C | Similar to respiratory burst oxidase protein | | FAD-binding domain | 1.47 | 1.61 | 2.27 | 2.19 |
| CMV176C | 50S ribosomal protein L5 (rpl5) | Ribosome | Ribosomal protein L5 | 1.48 | 3.55 | 0.46 | 0.85 |
| CMS349C | Hypothetical protein | | | 1.53 | 1.39 | 2.37 | 1.47 |
| CMP311C | TBP-1 interacting protein | | Syntaxin | 1.66 | 1.06 | 0.47 | 0.99 |
| CMF163C | Similar to heterogeneous nuclear ribonucleoprotein H3, isoform a | | RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain) | 1.68 | 1.44 | 2.01 | 1.24 |
| CMK034C | Hypothetical protein | | Helix-loop-helix DNA-binding domain | 1.75 | 1.47 | 2.73 | 3.08 |
| CMT581C | Hypothetical protein, conserved | | Nucleotidyltransferase domain | 2.00 | 1.39 | 1.03 | 1.14 |
| CML304C | 60S ribosomal protein L35A (rpl35A) | Ribosome | Ribosomal protein L35Ae | 2.01 | 1.04 | 0.63 | 1.05 |
| CMT589Z | Possible transcribed region | | | 2.01 | 1.1 | 0.91 | 1.12 |
| CML311C | Hypothetical protein, conserved | | Conserved region of Rad21 / Rec8 like protein | 2.02 | 1.35 | 0.99 | 1.53 |
| CMK273C | 60S ribosomal protein L23A (rpl23A) | Ribosome | Ribosomal protein L23 | 2.02 | 1.37 | 0.7 | 1.4 |
| CMP339C | Similar to Golgi vesicular membrane trafficking protein Bet1p | SNARE Interactions | SlyX | 2.03 | 0.98 | 0.69 | 1.31 |
| CMI152C | Hypothetical protein, conserved | | Anaphase-promoting complex, cyclosome, subunit 3 | 2.04 | 1.18 | 0.85 | 1.16 |

Table A3.7 Continued…

| Gene ID | Annotation | KEGG Functional Pathway | Pfam Domain | +UV:-UV Expression Nov. 2009 | Jan. 2010 | Apr. 2010 | Jul. 2010 |
|---|---|---|---|---|---|---|---|
| CMK009C | Hypothetical protein | | NUMOD3 motif (2 copies) | 2.05 | 1.11 | 0.9 | 1.19 |
| CMS212C | Mitochondrial ribosomal protein S17 precursor | | Ribosomal protein S17 | 2.07 | 1.48 | 0.84 | 1.08 |
| CML015C | Similar to ferredoxin | Photosynthesis | 2Fe-2S iron-sulfur cluster binding domain | 2.1 | 1.07 | 0.84 | 1.13 |
| CMR465C | Sulfate ABC transporter ATP-binding subunit cysA | | ABC transporter | 2.11 | 1.31 | 0.81 | 1.06 |
| CML303Z | Possible transcribed region | | | 2.11 | 1.06 | 0.81 | 1.2 |
| CMO184C | Similar to hedgehog protein | | Hint module | 2.15 | 0.99 | 0.73 | 1.05 |
| CMS027C | Hypothetical protein, conserved | | CBS domain | 2.15 | 1.1 | 0.96 | 1.52 |
| CMI218C | Hypothetical protein, conserved | | 6S proteasome subunit RPN7 | 2.16 | 1.35 | 1.15 | 1.33 |
| CMO152C | Signal recognition particle recognition component SRP54 | Protein Export | SRP54-type protein, GTPase domain | 2.16 | 1.23 | 0.79 | 1.18 |
| CMF159C | AAA-metalloprotease FtsH, mitochondrial precursor | | ATPase family associated with various cellular activities (AAA) | 2.17 | 1.23 | 1.13 | 1.42 |
| CMQ072C | Hypothetical protein | | | 2.17 | 1.19 | 0.76 | 1.21 |
| CME134C | Hypothetical protein | | Myb-like DNA-binding domain | 2.18 | 1.34 | 0.62 | 1.18 |
| CMP056Z | Possible transcribed region | | | 2.2 | 1.24 | 0.84 | 1.37 |
| CMF025C | Hypothetical protein | | | 2.22 | 1.01 | 0.56 | 0.91 |
| CMP290C | Hypothetical protein | | | 2.23 | 1.06 | 0.84 | 1.13 |
| CMT167C | Hypothetical protein, conserved | | AhpC/TSA family | 2.23 | 1.3 | 0.64 | 1.28 |
| CMT390C | Hypothetical protein, conserved | | RNA pseudouridylate synthase | 2.28 | 1.1 | 0.84 | 1.22 |
| CMP189C | Probable selenophosphate synthetase | Selenocompound metabolism | AIR synthase related protein, N-terminal domain | 2.29 | 1.27 | 0.77 | 1.29 |

Table A3.7 Continued…

| Gene ID | Annotation | KEGG Functional Pathway | Pfam Domain | +UV:-UV Expression Nov. 2009 | Jan. 2010 | Apr. 2010 | Jul. 2010 |
|---|---|---|---|---|---|---|---|
| CMA031C | Hypothetical protein | | | 2.29 | 1.13 | 0.74 | 1.19 |
| CMA044C | Cryptochrome DASH | | DNA photolyase | 2.3 | 1.09 | 1.13 | 1.78 |
| CMT304C | Hypothetical protein, conserved | | Aldo/keto reductase family | 2.3 | 1.54 | 1.19 | 1.76 |
| CMN046C | Sporulation-induced transcript 4-associated protein SAPL b | | SIT4 phosphatase-associated protein | 2.32 | 1.19 | 0.94 | 1.09 |
| CMM075T | Hypothetical transcript | | | 2.33 | 1.01 | 1.04 | 1.17 |
| CMK106C | Hypothetical protein, conserved | | E2 (early) protein, N terminal | 2.34 | 1.02 | 0.69 | 1.02 |
| CMP134C | Similar to transcriptional adaptor like protein | | Myb-like DNA-binding domain | 2.34 | 1.02 | 0.84 | 1.15 |
| CMK262C | Hypothetical protein | | | 2.35 | 1.07 | 0.39 | 0.93 |
| CMD051C | Hypothetical protein | | Cse1 | 2.36 | 1.43 | 0.66 | 1.16 |
| CME043C | Hypothetical protein, conserved | | CCT motif | 2.39 | 1.54 | 1.41 | 2.4 |
| CML294C | Hypothetical protein | | 2',5' RNA ligase family | 2.42 | 1.05 | 0.75 | 1.16 |
| CMT237Z | Possible transcribed region | | | 2.42 | 1.17 | 0.78 | 1.23 |
| CMB097C | Hypothetical protein | | Transcription initiation factor TFIID component TAF4 family | 2.45 | 1.13 | 0.79 | 1.22 |
| CMA025C | Hypothetical protein | | Bromodomain | 2.5 | 1.01 | 0.76 | 1.18 |
| CMN167C | RIO-like serine/threonine protein kinase (rio1) | Ribosome biogenesis in eukaryotes | RIO1 family | 2.58 | 1.57 | 0.82 | 1.29 |
| CMS343C | Molecular chaperone of Hsp110 family | | Hsp70 protein | 2.58 | 1.5 | 0.83 | 1.07 |
| CMT223C | Similar to translation initiation factor IF-3 (eIF3) | RNA transport | Translation initiation factor IF-3, C-terminal domain | 2.59 | 1.1 | 0.83 | 1.23 |
| CMQ412C | Similar to psbB mRNA maturation factor Mbb1 | | SecE/Sec61-gamma subunits of protein translocation complex | 2.6 | 1.89 | 1.34 | 1.62 |
| CMI216C | Probable mitochondrial intermembrane space complex subunit Tim8 | | Tim10/DDP family zinc finger | 2.62 | 1.44 | 0.8 | 0.97 |

Table A3.7 Continued…

| Gene ID | Annotation | KEGG Functional Pathway | Pfam Domain | +UV:-UV Expression Nov. 2009 | Jan. 2010 | Apr. 2010 | Jul. 2010 |
|---|---|---|---|---|---|---|---|
| CME019C | Dynamin-related protein involved in mitochondrial division CmDnm1/DRP3 | Endocytosis | Dynamin family | 2.64 | 1.28 | 0.7 | 1.13 |
| CMS360C | Hypothetical protein, conserved | | | 2.64 | 1.15 | 0.66 | 1.16 |
| CMM200C | Probable DNA binding protein | | Metallopeptidase family M24 | 2.65 | 1.05 | 0.63 | 1.24 |
| CMM132C | Mitochondrial propionyl-CoA carboxylase, beta subunit, precursor | Glyoxylate and dicarboxylate metabolism | Carboxyl transferase domain | 2.68 | 1.14 | 1.03 | 1.15 |
| CMR305C | Similar to pleiotropic regulator 1 (PRL1) | | WD domain, G-beta repeat | 2.76 | 1.1 | 0.81 | 1.19 |
| CMG129C | MYB-related protein | | Myb-like DNA-binding domain | 2.77 | 1.83 | 1.54 | 1.98 |
| CMT321C | Probable eukaryotic translation initiation factor (eIF-1) | RNA transport | Translation initiation factor SUI1 | 2.81 | 1.06 | 0.44 | 0.97 |
| CMN119T | Hypothetical transcript | | | 2.85 | 1.11 | 0.53 | 1.04 |
| CMT033Z | Possible transcribed region | | | 2.89 | 1.04 | 2.21 | 1.77 |
| CMJ046C | Similar to kinetoplast-associated protein | | CHASE4 domain | 2.9 | 1.58 | 0.57 | 1.12 |
| CMP277C | Mitochondrial ribosomal protein L2 precursor (rplB) | Ribosome | Ribosomal Proteins L2, RNA binding domain | 2.94 | 1.24 | 0.94 | 1.29 |
| CMV007C | Initiation factor 2 (eIF2) | RNA transport | Elongation factor Tu domain 2 | 3.03 | 2.29 | 0.44 | 0.92 |
| CMI252C | DNA topoisomerase I | | ABC transporter | 3.12 | 1.15 | 0.68 | 1.29 |
| CMI042C | Tyrosine--tRNA ligase | | Glutathione S-transferase, C-terminal domain | 3.12 | 1.37 | 1.01 | 1.41 |
| CMB111Z | Possible transcribed region | | | 3.2 | 1.09 | 1.18 | 1.18 |
| CMJ237C | Hypothetical protein, conserved | | PurA ssDNA and RNA-binding protein | 3.69 | 1.46 | 0.66 | 1.18 |
| CMA029C | Hypothetical protein | | NUC173 domain | 4.03 | 1.15 | 1.14 | 1.17 |
| CMP012C | 60S ribosomal protein L34 (rpl34) | Ribosome | Ribosomal protein L34e | 8.17 | 1.56 | 0.43 | 1.3 |

Table A3.8 Cluster C: Lemonade Creek gene expression patterns as a function of UV-intensity at four time points. Red and green colors depict significant expression changes, either an increase or decrease in the +UV:-UV expression ratios under the UV-exposed cyanidial mats. Heat map color depiction visually illustrates a statistically significant positive (induction, red) or negative (repression, green) expression ratios that were ± 2.

| | | | | +UV:-UV Expression Ratio | | | |
|---|---|---|---|---|---|---|---|
| Gene ID | Annotation | KEGG Functional Pathway | Pfam Domain | Nov. 2009 | Jan. 2010 | Apr. 2010 | Jul. 2010 |
| CMV213C | 30S ribosomal protein S18 (rps18) | Ribosome | Ribosomal protein S18 | 0.12 | 0.64 | 1.07 | 0.93 |
| CMW019C | Cytochrome c oxidase polypeptide I (cox1) | Oxidative phosphorylation | Cytochrome C and Quinol oxidase polypeptide I | 0.12 | 1.61 | 1.82 | 0.51 |
| CMW010C | Cytochrome c oxidase polypeptide II (cox2) | Oxidative phosphorylation | Cytochrome C oxidase subunit II, periplasmic domain | 0.13 | 1.2 | 1.5 | 0.65 |
| CMW007C | Cytochrome B (cytB) | Oxidative phosphorylation | Cytochrome b(C-terminal)/b6/petD | 0.15 | 1.46 | 1.67 | 0.54 |
| CMW002C | Succinate dehydrogenase iron-sulfur protein (sdhB) | Oxidative phosphorylation | 2Fe-2S iron-sulfur cluster binding domain | 0.16 | 1.63 | 1.91 | 0.51 |
| CMW021C | NADH-ubiquinone oxidoreductase chain 1 (nad1) | Oxidative phosphorylation | NADH dehydrogenase | 0.19 | 2.61 | 1.39 | 0.63 |
| CMW027C | ATP synthase A chain (protein 6) | Oxidative phosphorylation | ATP synthase A chain | 0.21 | 3.7 | 1.56 | 0.64 |
| CMV038C | Remnant of ycf55 | | | 0.22 | 1.35 | 1.19 | 0.52 |
| CMK169Z | Possible transcribed region | | | 0.29 | 1.02 | 1.29 | 0.55 |
| CMV004C | Cell-cycle protein mesJ family | | PP-loop family | 0.33 | 1.12 | 1.15 | 0.47 |
| CMV127C | Photosystem II 10 kDa phosphoprotein | Photosynthesis | Photosystem II 10 kDa phosphoprotein | 0.35 | 3.20 | 0.64 | 0.48 |
| CMV037C | Probable tetrapyrrole-binding protein GUN4 | | GUN4-like | 0.37 | 1.28 | 1.15 | 0.58 |
| CMW013C | 50S ribosomal protein L6 (rpl6) | Ribosome | Ribosomal protein L6 | 0.4 | 2.57 | 1.04 | 0.65 |
| CMW025C | NADH-ubiquinone oxidoreductase chain 5 (nad5) | Oxidative phosphorylation | NADH-Ubiquinone/plastoquinone (complex I), various chains | 0.4 | 2.07 | 1.04 | 0.77 |

Table A3.8 Continued…

| Gene ID | Annotation | KEGG Functional Pathway | Pfam Domain | +UV:-UV Expression Ratio | | | |
|---|---|---|---|---|---|---|---|
| | | | | Nov. 2009 | Jan. 2010 | Apr. 2010 | Jul. 2010 |
| CMC014Z | Possible transcribed region | | | 0.4 | 0.95 | 1.05 | 0.72 |
| CMK168C | Hypothetical protein | | | 0.45 | 0.87 | 1.56 | 0.63 |
| CMV147C | Cytochrome b6-f complex subunit VII | Photosynthesis | Cytochrome oxidase maturation protein cbb3-type | 0.45 | 0.83 | 0.92 | 0.57 |
| CMV012C | Thioredoxin type m | | Thioredoxin | 0.46 | 2.68 | 0.49 | 0.66 |
| CMV046C | 50S ribosomal protein L32 (rpl32) | Ribosome | Ribosomal L32p protein family | 0.47 | 1.55 | 0.44 | 0.69 |
| CMH112Z | Possible transcribed region | | | 0.48 | 0.87 | 0.84 | 0.74 |
| CMD116C | Hypothetical protein, conserved | | LysR substrate binding domain | 0.49 | 1.05 | 1.08 | 0.95 |
| CMV153C | Pyruvate dehydrogenase E1 component alpha subunit | | Dehydrogenase E1 component | 0.64 | 2.05 | 0.85 | 0.64 |
| CMW060C | NADH-ubiquinone oxidoreductase chain 4L (nad4L) | Oxidative phosphorylation | NADH-ubiquinone/plastoquinone oxidoreductase chain 4L | 0.71 | 3.48 | 0.80 | 0.94 |
| CMV020C | Probable ABC transporter ATP-binding protein YhbG homolog | | ABC transporter | 0.72 | 2.92 | 0.56 | 0.78 |
| CMV192C | 30S ribosomal protein S10 (rps10) | Ribosome | Ribosomal protein S10p/S20e | 0.75 | 2.03 | 0.53 | 1.02 |
| CML084Z | Possible transcribed region | | | 0.76 | 1.2 | 2.2 | 0.71 |

Table A3.8 Continued…

| Gene ID | Annotation | KEGG Functional Pathway | Pfam Domain | +UV:-UV Expression Ratio | | | |
|---|---|---|---|---|---|---|---|
| | | | | Nov. 2009 | Jan. 2010 | Apr. 2010 | Jul. 2010 |
| CMV151C | 50S ribosomal protein L20 rpl20) | Ribosome | Ribosomal protein L20 | 0.80 | 2.45 | 0.43 | 0.85 |
| CMS309C | Similar to microsomal glutathione S-transferase | Glutathione metabolism | MAPEG family | 0.8 | 1.46 | 2.15 | 1.18 |
| CMV117C | 4-(2'-carboxyphenyl)-4-oxybutyric acid synthase (menC) | Ubiquinone and other terpenoid-quinone biosynthesis | | 0.89 | 2.57 | 0.47 | 0.68 |
| CMW053C | 30S ribosomal protein S12 (rps12) | Ribosome | Ribosomal protein S12 | 0.91 | 4.09 | 0.33 | 0.98 |
| CMV113C | Naphthoate synthase (menB) | Ubiquinone and other terpenoid-quinone biosynthesis | Enoyl-CoA hydratase/isomerase family | 1.01 | 2.09 | 0.39 | 0.68 |
| CMV062C | Glucosamine-fructose-6-phosphate aminotransferase | | Glutamine amidotransferases class-II | 1.04 | 2.08 | 0.43 | 0.87 |
| CMV132C | Acyl carrier protein | | Phosphopantetheine attachment site | 1.06 | 2.37 | 0.40 | 0.84 |
| CMW037C | 50S ribosomal protein L14 (rpl14) | Ribosome | Ribosomal protein L14p/L23e | 1.08 | 2.81 | 0.39 | 1.05 |
| CMV009C | 30S ribosomal protein S4 (rps4) | Ribosome | Ribosomal protein S4/S9 N-terminal domain | 1.11 | 1.65 | 0.39 | 0.67 |

Table A3.9 Summary of gene expression events exclusively sensitive to VIS wavelengths.

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | July:January Expression Ratio VIS Only |
|---|---|---|---|---|
| CMK002Z | Pseudo-gene of trefoil factor | | | 0.08 |
| CMR342X | Repetitive element | | | 0.12 |
| CMO064X | Repetitive element | | | 0.12 |
| CMI053X | Repetitive element | | | 0.13 |
| CMR063X | Repetitive Element | | | 0.13 |
| CMF010X | Repetitive element | | | 0.14 |
| CMO251X | Repetitive element | | | 0.14 |
| CMQ377X | Repetitive element | | | 0.14 |
| CML337X | Repetitive element | | | 0.15 |
| CML275X | Repetitive element | | | 0.16 |
| CMI109C | Similar to glutenin low molecular weight chain precursor | | Fibrinogen binding protein | 0.16 |
| CMP317X | Repetitive element | | | 0.16 |
| CMQ439X | Repetitive element | | | 0.17 |
| CMD128X | Repetitive element | | | 0.17 |
| CMF172X | Repetitive element | | | 0.17 |
| CMD109C | Hypothetical protein, conserved | | Protein of unknown function (DUF1350) | 0.17 |
| CMG038X | Repetitive element | | | 0.18 |
| CMS017X | Repetitive element | | | 0.18 |
| CMP264X | Repetitive element | | | 0.18 |
| CMT557X | Repetitive Element | | | 0.18 |
| CMD189X | Repetitive element | | | 0.18 |
| CMB068X | Repetitive element | | | 0.18 |
| CMQ400X | Repetitive element | | | 0.19 |
| CMO158X | Repetitive element | | | 0.19 |
| CMH033X | Repetitive element | | | 0.19 |
| CML208X | Repetitive element | | | 0.20 |
| CMI003X | Repetitive element | | | 0.20 |

Table A3.9 Continued…

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | July:January Expression Ratio VIS Only |
|---|---|---|---|---|
| CMH089X | Repetitive element | | | 0.21 |
| CMM283X | Repetitive element | | | 0.21 |
| CMT005X | Repetitive element | | | 0.21 |
| CMT002X | Repetitive element | | | 0.21 |
| CMP197X | Repetitive element | | | 0.22 |
| CMS055X | Repetitive element | | | 0.22 |
| CMP048X | Repetitive element | | | 0.22 |
| CML233X | Repetitive element | | | 0.24 |
| CMN314X | Repetitive element | | | 0.24 |
| CMN020X | Repetitive element | | | 0.25 |
| CMT509X | Repetitive element | | | 0.25 |
| CMP082X | Repetitive element | | | 0.25 |
| CMS016X | Repetitive element | | | 0.26 |
| CMM327X | Repetitive element | | | 0.26 |
| CMM168X | Repetitive element | | | 0.27 |
| CMK310X | Repetitive element | | | 0.28 |
| CMH196X | Repetitive element | | | 0.28 |
| CMB137X | Repetitive element | | | 0.28 |
| CMT295Z | Pseudo-gene of 3-oxo-5-alpha-steroid 4-dehydrogenase | | | 0.29 |
| CMO294X | Repetitive element | | | 0.29 |
| CMR166X | Repetitive element | | | 0.30 |
| CMS003X | Repetitive element | | | 0.30 |
| CMC187X | Repetitive element | | | 0.30 |
| CMQ202X | Repetitive element | | | 0.30 |
| CMS186X | Repetitive element | | | 0.31 |
| CMT505X | Repetitive element | | | 0.32 |
| CMT507X | Repetitive element | | | 0.33 |
| CMP121C | Retroelement, dead | | Reverse transcriptase | 0.34 |

Table A3.9 Continued…

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | July:January Expression Ratio VIS Only |
|---|---|---|---|---|
| CMJ003X | Repetitive element | | | 0.34 |
| CML151X | Repetitive element | | | 0.36 |
| CMA143X | Repetitive element | | | 0.37 |
| CMN064C | Hypothetical protein | | | 0.38 |
| CMG150T | Hypothetical transcript | | | 0.38 |
| CMS196X | Repetitive element | | | 0.38 |
| CMP120X | Repetitive Element | | | 0.38 |
| CMR265C | Retroelement | | Reverse transcriptase | 0.39 |
| CMD064X | Repetitive element | | | 0.39 |
| CMH128C | Cyclin dependent kinase, B-type | | Protein kinase domain | 0.39 |
| CMH175X | Repetitive element | | | 0.39 |
| CMJ025T | Hypothetical transcript | | | 0.40 |
| CMI085X | Repetitive element | | | 0.40 |
| CMM183X | Repetitive element | | | 0.40 |
| CMH244T | Hypothetical transcript | | | 0.40 |
| CMK001C | Similar to hedgehog protein | | Hint module | 0.41 |
| CMF002Z | Pseudo-gene of trefoil factor | | | 0.41 |
| CMP241Z | Possible transcribed region | | | 0.41 |
| CMF003Z | Possible transcribed region | | | 0.43 |
| CMO293T | Hypothetical transcript | | | 0.43 |
| CMA137Z | Possible transcribed region | | | 0.45 |
| CMH002Z | Pseudo-gene of trefoil factor | | | 0.45 |
| CMS160C | Hypothetical protein | | PHD-finger | 0.45 |

Table A3.9 Continued…

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | July:January Expression Ratio VIS Only |
|---|---|---|---|---|
| CMT065X | Repetitive element | | | 0.45 |
| CMM302X | Repetitive element | | | 0.45 |
| CMD178X | Repetitive element | | | 0.46 |
| CMR235T | Hypothetical transcript | | | 0.46 |
| CMT455X | Repetitive element | | | 0.46 |
| CMM295C | Hypothetical protein | | | 0.46 |
| CMN093T | Hypothetical transcript | | | 0.47 |
| CMJ274X | Repetitive element | | | 0.47 |
| CMT446X | Repetitive element | | | 0.47 |
| CMN077Z | Possible transcribed region | | | 0.48 |
| CMO271C | Cell cycle protein kinase (cdc7) | | Protein kinase domain | 0.48 |
| CMP132X | Repetitive element | | | 0.48 |
| CMH001C | Similar to hedgehog protein | | Hint module | 0.48 |
| CMS004C | Plastid division protein (ftsZ) | Cell Cycle | Tubulin/FtsZ family, GTPase domain | 0.48 |
| CMP030C | Probable chromatin assembly factor 1 subunit B | | WD domain, G-beta repeat | 0.48 |
| CML207C | Retroelement | | Reverse transcriptase | 0.49 |
| CMN156C | Hypothetical protein, conserved | | PLATZ transcription factor | 0.50 |
| CMM123T | Hypothetical transcript | | | 0.50 |
| CMN340C | Similar to hedgehog protein | | Hint module | 0.50 |

Table A3.9 Continued…

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | July:January Expression Ratio VIS Only |
|---|---|---|---|---|
| CMQ463C | 60S ribosomal protein L17 (rpl17) | Ribosome | Ribosomal protein L22p/L17e | 0.50 |
| CMM027Z | Possible transcribed region | | | 2.00 |
| CMB082C | Similar to JEMMA protein | | PHD-finger | 2.00 |
| CME045C | Hypothetical protein | | | 2.00 |
| CMD148C | ATP-binding cassette, sub-family B (MDR/TAP), member 1 | | ABC transporter | 2.01 |
| CMA019C | RNA polymerase III transcription factor IIIB | | Transcription factor TFIIB repeat | 2.01 |
| CMR236C | Probable tRNA-specific adenosine-34 deaminase | | Cytidine and deoxycytidylate deaminase zinc- | 2.01 |
| CMN058C | Hypothetical protein | | Tom37 C-terminal domain | 2.01 |
| CMV115C | 2-succinyl-6-hydroxy-2, 4-cyclohexadiene-1-carboxylate synthase (menD) | Ubiquinone and other terpenoid-quinone biosynthesis | Thiamine pyrophosphate enzyme, C-terminal TPP binding domain | 2.02 |
| CMP109Z | Possible transcribed region | | | 2.02 |
| CMN327X | Repetitive element | | | 2.02 |
| CMT424C | Hypothetical protein | | Pumilio-family RNA binding repeat | 2.03 |
| CMF024C | Hypothetical protein, conserved | | WLM domain | 2.03 |
| CMB046C | Similar to JEMMA protein | | PHD-finger | 2.03 |
| CMO209C | Hypothetical protein, conserved | | Thioredoxin | 2.03 |

Table A3.9 Continued…

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | July:January Expression Ratio VIS Only |
|---|---|---|---|---|
| CMQ412C | Similar to psbB mRNA maturation factor Mbb1 | | SecE/Sec61-gamma subunits of protein translocation complex | 2.05 |
| CMQ101C | Gamma-glutamyltransferase | Arachidonic acid metabolism | Gamma-glutamyltranspeptidase | 2.05 |
| CMP079Z | Possible transcribed region | | | 2.05 |
| CMS122C | Probable farnesyl-protein transferase beta-subunit | | Prenyltransferase and squalene oxidase repeat | 2.06 |
| CMR361C | Hypothetical protein | | Domain of unknown function (DUF1857) | 2.06 |
| CMS276C | Exportin 1 (xpo1) | Ribosome biogenesis in eukaryotes | Importin-beta N-terminal domain | 2.06 |
| CMV054C | ORF47 | | | 2.08 |
| CML253C | Hypothetical protein | | | 2.09 |
| CMS459C | Similar to RNA export mediator | | GLE1-like protein | 2.10 |
| CMS220C | Probable phosphate transporter | | Phosphate transporter family | 2.10 |
| CMM121C | Hypothetical protein, conserved | | ATP-dependent Clp protease adaptor protein ClpS | 2.10 |
| CMV086C | 60 kDa chaperonin (groEL) | RNA Degradation | TCP-1/cpn60 chaperonin family | 2.13 |
| CMT356C | Dynein light chain | | Dynein light chain type 1 | 2.14 |
| CML121C | Hypothetical protein | | Syntaxin | 2.14 |
| CMQ385C | Probable pre-mRNA splicing factor ATP-dependent RNA helicase PRP16 | | DEAD/DEAH box helicase | 2.15 |

Table A3.9 Continued…

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | July:January Expression Ratio VIS Only |
|---|---|---|---|---|
| CMI075T | Hypothetical transcript | | | 2.15 |
| CMN143C | Similar to t-SNARE SED5 | SNARE Interactions | Syntaxin | 2.15 |
| CMD129C | Similar to BRCA1-binding helicase-like protein BACH1 | | Type III restriction enzyme, res subunit | 2.17 |
| CMI026C | Similar to cyclin M2 | Cell Cycle | CBS domain | 2.17 |
| CMI249C | Similar to syntaxin | SNARE Interactions | Syntaxin | 2.17 |
| CMV153C | Pyruvate dehydrogenase E1 component alpha subunit (odpA) | | Dehydrogenase E1 component | 2.21 |
| CMV080C | OmpR-like transcriptional regulator | | Response regulator receiver domain | 2.21 |
| CMA084Z | Possible transcribed region | | | 2.22 |
| CME144C | Hypothetical protein | | | 2.23 |
| CMW007C | Cytochrome B (cytB) | Oxidative phosphorylation | Cytochrome b(C-terminal)/b6/petD | 2.23 |
| CMQ138C | Hypothetical protein, conserved | | Mitochondrial ribosomal protein L27 | 2.24 |
| CMT207C | Chaperonin containing TCP1, subunit 1 (alpha) (tcp1) | | TCP-1/cpn60 chaperonin family | 2.24 |
| CMV110C | ELIP-like protein | | Chlorophyll A-B binding protein | 2.24 |
| CMP314C | Similar to tyrosine phosphatase Oca1p | | Tyrosine phosphatase family | 2.25 |
| CMR362C | Hypothetical protein, conserved | | Peptidase family M28 | 2.25 |

Table A3.9 Continued…

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | July:January Expression Ratio VIS Only |
|---|---|---|---|---|
| CMC144C | Hypothetical protein, conserved | | yrdC domain | 2.25 |
| CMS198Z | Possible transcribed region | | | 2.26 |
| CMS177C | Hypothetical Protein | | | 2.27 |
| CMK210C | Cysteine--tRNA ligase, mitochondrial or chloroplast | | tRNA synthetases class I (C) catalytic domain | 2.27 |
| CMT371C | Small GTP-binding protein of Ras family | | Ras family | 2.29 |
| CMR354C | Hypothetical protein | | alpha/beta hydrolase fold | 2.31 |
| CMB100Z | Possible transcribed region | | | 2.31 |
| CMC060C | Hypothetical protein, conserved | | THUMP domain | 2.31 |
| CMV046C | 50S ribosomal protein L32 (rpl32) | Ribosome | Ribosomal L32p protein family | 2.34 |
| CME062C | Probable cationic amino acid transporter (ctrA) | Cell Cycle | Amino acid permease | 2.35 |
| CMW019C | Cytochrome c oxidase polypeptide I (cox1) | Oxidative phosphorylation | Cytochrome C and Quinol oxidase polypeptide I | 2.36 |
| CMM041C | Hypothetical protein | | Complex 1 protein (LYR family) | 2.41 |
| CMF055C | GTP cyclohydrolase I (fol2) | Folate biosynthesis | GTP cyclohydrolase I | 2.45 |
| CMH028C | Probable alpha-1,3-glucosyltransferase (alg6) | N-glygan biosynthesis | ALG6, ALG8 glycosyltransferase family | 2.45 |
| CMI186C | Cytochrome c-type biogenesis protein (ccmE) | ABC Transporters | CcmE | 2.46 |

Table A3.9 Continued…

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | July:January Expression Ratio VIS Only |
|---|---|---|---|---|
| CMT380C | TRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase (trmU) | Sulfur relay system | tRNA methyl transferase | 2.47 |
| CMT122C | ATP-dependent Hsl protease ATP-binding subunit (hslU) | | Protein of unknown function (DUF826) | 2.48 |
| CMS244T | Hypothetical transcript | | | 2.48 |
| CMQ259C | Hypothetical protein | | Protein of unknown function (DUF448) | 2.50 |
| CMO033Z | Possible transcribed region | | | 2.51 |
| CML057C | Unknown kinase with aarF domain | | Phosphotransferase enzyme family | 2.56 |
| CMA083C | Hypothetical protein, conserved | | Protein of unknown function (DUF1350) | 2.57 |
| CMP323C | Hypothetical protein, conserved | | Activator of Hsp90 ATPase homolog 1-like protein | 2.64 |
| CMV113C | Naphthoate synthase (menB) | Ubiquinone and other terpenoid-quinone biosynthesis | Enoyl-CoA hydratase/isomerase family | 2.69 |
| CMW025C | NADH-ubiquinone oxidoreductase chain 5 (nad5) | Oxidative phosphorylation | NADH-Ubiquinone/plastoquinone (complex I), various chains | 2.69 |
| CMV117C | 4-(2'-carboxyphenyl)-4-oxybutyric acid synthase (menC) | Ubiquinone and other terpenoid-quinone biosynthesis | | 2.71 |

Table A3.9 Continued…

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | July:January Expression Ratio VIS Only |
|---|---|---|---|---|
| CMS343C | Molecular chaperone of Hsp110 family (hsp105) | Protein Processing in the ER | Hsp70 protein | 2.71 |
| CMI120Z | Possible transcribed region | | | 2.74 |
| CMV193C | Ferredoxin (petF) | Photosynthesis | 2Fe-2S iron-sulfur cluster binding domain | 2.85 |
| CMR086C | Hypothetical protein | | | 2.89 |
| CMO324C | 4-amino-4-deoxychorismate (ADC) synthase | Folate biosynthesis | Glutamine amidotransferase class-I | 2.92 |
| CMV012C | Thioredoxin type m (trxM) | | Thioredoxin | 3.02 |
| CMA011C | Probable aminoacylase I | Arginine and Proline metabolism | Peptidase family M20/M25/M40 | 3.10 |
| CMV127C | Photosystem II 10 kDa phosphoprotein (psbH) | Photosynthesis | Photosystem II 10 kDa phosphoprotein | 3.16 |
| CMW027C | ATP synthase A chain (protein 6) (atp6) | Oxidative phosphorylation | ATP synthase A chain | 3.20 |
| CMW013C | 50S ribosomal protein L6 (rpl6) | Ribosome | Ribosomal protein L6 | 3.43 |
| CMW037C | 50S ribosomal protein L14 (rpl14) | Ribosome | Ribosomal protein L14p/L23e | 3.54 |
| CMA085C | Hypothetical Protein | | | 3.56 |
| CMV003C | Unknown | | Photosystem II complex subunit Ycf12 | 3.67 |
| CMW021C | NADH-ubiquinone oxidoreductase chain 1 (nad1) | Oxidative phosphorylation | NADH dehydrogenase | 3.67 |
| CMW060C | NADH-ubiquinone oxidoreductase chain 4L (nad4L) | Oxidative phosphorylation | NADH-ubiquinone/plastoquinone oxidoreductase chain 4L | 3.82 |

Table A3.10 Summary of genes displaying a statistically significant correlation (positive or negative) to seasonal shifts in VIS irradiance intensity.

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | r | $r^2$ | p-value |
|---|---|---|---|---|---|---|
| CMP356C | DNA replication licensing factor (mcm4) | DNA Replication | MCM2/3/5 family | -0.97 | 0.95 | 0.001 |
| CMT087C | Similar to DNA replication licensing factor (mcm8) | DNA Replication | MCM2/3/5 family | -0.96 | 0.92 | 0.003 |
| CMS145C | Similar to shuttle craft like transcriptional regulator NFX1 | | | -0.92 | 0.85 | 0.019 |
| CMS123C | Hypothetical protein | | | -0.92 | 0.85 | 0.019 |
| CMS361C | Organelle division protein (ftsZ) | Cell Cycle | Tubulin/FtsZ family, GTPase domain | -0.92 | 0.84 | 0.021 |
| CMJ016C | Aspartate kinase | Cysteine and Methionine metabolism | Amino acid kinase family | -0.92 | 0.84 | 0.021 |
| CMB004C | 40S ribosomal protein S18 | Ribosome | Ribosomal protein S13/S18 | -0.91 | 0.83 | 0.023 |
| CMS426C | Hypothetical protein | | Tetraspanin family | -0.91 | 0.82 | 0.023 |
| CMP271C | Similar to cyclophilin B | | Cyclophilin type peptidyl-prolyl cis-trans isomerase/CLD | -0.89 | 0.80 | 0.028 |
| CMP216C | CTP synthase | Pyrimidine metabolism | Glutamine amidotransferase class-I | -0.88 | 0.78 | 0.036 |
| CML190C | Cell division cycle protein cdc23 | Ubiquitin mediated proteolysis | Tetratricopeptide repeat | -0.88 | 0.78 | 0.036 |
| CML214C | Ornithine carbamoyltransferase | Arginine and Proline metabolism | Aspartate/ornithine carbamoyltransferase, Asp/Orn binding domain | -0.87 | 0.76 | 0.041 |

Table A3.10 Continued…

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | r | $r^2$ | p-value |
|---|---|---|---|---|---|---|
| CMG150T | Hypothetical transcript | | | -0.86 | 0.74 | 0.049 |
| CMN259C | Similar to phospholpid scramblase | Phospholipid translocation | Scramblase | 0.86 | 0.74 | 0.049 |
| CML228C | Similar to serine/threonine kinase KIN82 | | Protein kinase domain | 0.86 | 0.74 | 0.049 |
| CMB092C | Flavin-dependent tRNA:m5U methyltransferase trmFO | | Glucose inhibited division protein A | 0.86 | 0.74 | 0.049 |
| CMA132C | Unknown metallopeptidase | | Peptidase M76 family | 0.87 | 0.75 | 0.043 |
| CMI267C | Protein-L-isoaspartate(D-aspartate) O-methyltransferase (pcmT) | Chaperones and folding catalysts | Protein-L-isoaspartate(D-aspartate) O-methyltransferase (PCMT) | 0.87 | 0.75 | 0.043 |
| CMT234C | Cysteine desulfurase IscS, mitochondrial precursor | Sulfur relay system | Aminotransferase class-V | 0.87 | 0.76 | 0.041 |
| CMM121C | ATP-dependent Clp protease adaptor protein (clpS) | Chaperones and folding catalysts | ATP-dependent Clp protease adaptor protein ClpS | 0.87 | 0.76 | 0.041 |
| CMI023C | Hypothetical protein | Cell cycle | ATPase MipZ | 0.87 | 0.76 | 0.041 |
| CME144C | Hypothetical protein | | | 0.87 | 0.76 | 0.041 |
| CMR297T | Hypothetical transcript | | | 0.87 | 0.77 | 0.040 |
| CMS387C | Similar to PRIP-interacting protein with methyltransferase domain | RNA transport | Homeobox prospero-like protein (PROX1) | 0.88 | 0.77 | 0.040 |

Table A3.10 Continued…

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | r | $r^2$ | p-value |
|---|---|---|---|---|---|---|
| CMR078C | Hypothetical protein | | Putative ATPase (DUF699) | 0.88 | 0.77 | 0.039 |
| CMI208T | Hypothetical transcript | | | 0.88 | 0.77 | 0.039 |
| CMO052Z | Possible transcribed region | | | 0.88 | 0.77 | 0.037 |
| CMA106C | Similar to alpha-glucosidase | | Glycosyl hydrolases family 31 | 0.88 | 0.78 | 0.036 |
| CMI120Z | Possible transcribed region | | | 0.89 | 0.79 | 0.034 |
| CMP080C | Hypothetical protein | | | 0.89 | 0.79 | 0.029 |
| CME029C | Chromosome assembly complex Condensin, core subunit C | Cell cycle | RecF/RecN/SMC N terminal domain | 0.89 | 0.80 | 0.028 |
| CMT558C | Chaperonin containing TCP1, subunit 7 (eta) (tcp1) | Chaperones and folding catalysts | TCP-1/cpn60 chaperonin family | 0.89 | 0.80 | 0.028 |
| CMK043C | MYB-related protein | | Myb-like DNA-binding domain | 0.89 | 0.80 | 0.028 |
| CMH079T | Hypothetical transcript | | | 0.90 | 0.80 | 0.028 |
| CMA083C | Hypothetical protein, conserved | | Protein of unknown function (DUF1350) | 0.90 | 0.81 | 0.028 |
| CMR298C | Heat shock protein (dnaJ) | Protein processing in the endoplasmic reticulum | DnaJ domain | 0.90 | 0.81 | 0.024 |
| CMF057C | Hypothetical protein, conserved | | | 0.90 | 0.82 | 0.024 |
| CMK306C | Hypothetical protein, conserved | | PAP fibrillin | 0.91 | 0.82 | 0.023 |
| CMP108T | Hypothetical transcript | | | 0.91 | 0.82 | 0.023 |
| CMT307Z | Possible transcribed region | | | 0.91 | 0.83 | 0.023 |
| CME143C | Smt3-activating enzyme E1 C subunit | | ThiF family | 0.92 | 0.85 | 0.019 |
| CMP109Z | Possible transcribed region | | | 0.93 | 0.87 | 0.015 |

Table A3.11 Summary of gene expression events that are sensitive to a combination of VIS and UV wavelengths. The VIS + UV column denotes the statistically significant July:January expression ratios of genes under the OP-4 filters (VIS + UV exposed), where the VIS column denotes the statistically significant expression ratios of genes under the UF-5 filter (VIS only exposed).

| | | | | July:January Expression Ratio | |
|---|---|---|---|---|---|
| **Gene ID** | **Annotation** | **Group** | **Pfam Domain** | **UV + VIS** | **VIS** |
| CMP356C | DNA replication licensing factor (mcm4) | DNA Replication | MCM2/3/5 family | 0.11 | 0.17 |
| CMO045C | Hypothetical protein | | CsiD | 0.11 | 0.18 |
| CMT084C | Hypothetical protein | | Oxidoreductase NAD-binding domain | 0.13 | 0.14 |
| CMK170C | Hypothetical protein | | | 0.13 | 0.19 |
| CMS347C | Probable mitotic spindle assembly checkpoint protein (mad2) | Cell Cycle | HORMA domain | 0.17 | 0.16 |
| CMN190X | Repetitive element | | | 0.17 | 0.18** |
| CMR481X | Repetitive element | | | 0.18 | 0.08 |
| CMJ261C | DNA replication licensing factor (mcm6) | DNA Replication | MCM2/3/5 family | 0.18 | 0.24 |
| CMS123C | Hypothetical protein | | | 0.18 | 0.24 |
| CMK201T | Hypothetical transcript | | | 0.19 | 0.22 |
| CMI203C | Probable mitotic cyclin a2-type (cyclin A2) | Cell Cycle | Cyclin, N-terminal domain | 0.19 | 0.22 |
| CMR430X | Repetitive element | | | 0.19 | 0.09** |
| CMQ036X | Repetitive element | | | 0.20 | 0.11** |
| CMH085X | Repetitive element | | | 0.23 | 0.16** |
| CMN263C | Beta-tubulin | Phagosome | Tubulin/FtsZ family, GTPase domain | 0.23 | 0.26 |
| CML278X | Repetitive element | | | 0.24 | 0.08** |
| CMF145X | Repetitive element | | | 0.24 | 0.13** |
| CMP242X | Repetitive element | | | 0.25 | 0.12** |
| CMQ135X | Repetitive element | | | 0.25 | 0.12** |

Table A3.11 Continued…

| Gene ID | Annotation | Group | Pfam Domain | July:January Expression Ratio | |
|---|---|---|---|---|---|
| | | | | UV + VIS | VIS |
| CMI243C | Plastid terminal oxidase | | Alternative oxidase | 0.25 | 0.35 |
| CMM234X | Repetitive element | | | 0.25 | 0.10** |
| CMD003Z | Pseudo-gene of NADPH:protochlorophyllide oxidoreductase | Porphyrin and chlorophyll metabolism | | 0.25 | 0.28 |
| CMA142C | Similar to trefoil factor | | Trefoil (P-type) domain | 0.25 | 0.18** |
| CMF004Z | Pseudo-gene of NADPH:protochlorophyllide oxidoreductase | Porphyrin and chlorophyll metabolism | | 0.26 | 0.26 |
| CMI235X | Repetitive element | | | 0.27 | 0.15** |
| CMS072X | Repetitive element | | | 0.27 | 0.14** |
| CMF173C | DNA replication licensing factor (mcm5) | DNA Replication | MCM2/3/5 family | 0.27 | 0.36 |
| CMK218X | Repetitive element | | | 0.28 | 0.10** |
| CMT068C | Hypothetical protein | | E2F/DP family winged-helix DNA-binding domain | 0.28 | 0.38 |
| CMP196X | Repetitive element | | | 0.28 | 0.11** |
| CMT632X | Repetitive element | | | 0.29 | 0.08** |
| CME002Z | Pseudo-gene of NADPH:protochlorophyllide oxidoreductase | Porphyrin and chlorophyll metabolism | | 0.29 | 0.18** |
| CMA036C | Similar to tubulin folding cofactor B | | CAP-Gly domain | 0.29 | 0.33 |
| CMV139C | Cytochrome b6-f complex subunit V | Photosynthesis | Cytochrome B6-F complex subunit 5 | 0.29 | 0.34 |
| CMQ262X | Repetitive element | | | 0.30 | 0.16** |
| CMP311C | TBP-1 interacting protein | | Syntaxin | 0.30 | 0.32 |
| CMQ428X | Repetitive element | | | 0.30 | 0.09** |

Table A3.11 Continued…

| Gene ID | Annotation | Group | Pfam Domain | July:January Expression Ratio | |
|---|---|---|---|---|---|
| | | | | UV + VIS | VIS |
| CMG208C | Retroelement | | Reverse transcriptase (RNA-dependent DNA polymerase) | 0.30 | 0.18** |
| CMI005X | Repetitive element | | | 0.30 | 0.17** |
| CMT010C | Retroelement, alive | | Reverse transcriptase (RNA-dependent DNA polymerase) | 0.30 | 0.27 |
| CMJ159C | Hypothetical protein | | F/Y rich C-terminus | 0.30 | 0.24 |
| CMF037X | Retroelement, dead | | | 0.31 | 0.26 |
| CMN189X | Repetitive element | | | 0.31 | 0.10** |
| CMH099Z | Pseudo-gene of 3-oxo-5-alpha-steroid 4-dehydrogenase | | | 0.31 | 0.35 |
| CML148C | Hypothetical protein | | | 0.32 | 0.34 |
| CMD191C | Similar to trefoil factor | | Trefoil (P-type) domain | 0.32 | 0.22 |
| CMQ425Z | Possible transcribed region | | | 0.32 | 0.38 |
| CMB162C | Similar to trefoil factor | | Trefoil (P-type) domain | 0.32 | 0.23 |
| CMN025X | Repetitive element | | | 0.33 | 0.24 |
| CMC065C | Hypothetical protein | | | 0.33 | 0.38 |
| CMO177C | Hypothetical protein, conserved | | Alpha/beta hydrolase family | 0.33 | 0.29 |
| CMR506C | Similar to hedgehog protein | | Hint module | 0.34 | 0.31 |
| CMN277X | Repetitive element | | | 0.34 | 0.15** |
| CMG028C | Similar to meiotic nuclear division protein (mnd1) | Cell Cycle | HlyD family secretion protein | 0.35 | 0.28 |
| CMS070X | Repetitive element | | | 0.35 | 0.32 |
| CMO295X | Repetitive element | | | 0.35 | 0.13** |
| CMS071C | Retroelement | | Reverse transcriptase (RNA-dependent DNA polymerase) | 0.35 | 0.43 |
| CMO238C | Hypothetical protein | | DNA replication factor CDT1 like | 0.36 | 0.39 |
| CMR234C | DNA replication licensing factor (mcm7) | DNA Replication | MCM2/3/5 family | 0.36 | 0.40 |

Table A3.11 Continued…

| Gene ID | Annotation | Group | Pfam Domain | July:January Expression Ratio | |
|---|---|---|---|---|---|
| | | | | UV + VIS | VIS |
| CMR358C | Hypothetical protein, conserved | | Glycosyl transferase family 8 | 0.36 | 0.33 |
| CMN098Z | Pseudo-gene of 3-oxo-5-alpha-steroid 4-dehydrogenase | | Trehalase | 0.36 | 0.39 |
| CMT454C | Similar to chondroitin 4-sulfotransferase | | Sulfotransferase family | 0.36 | 0.44 |
| CMC073X | Repetitive element | | | 0.36 | 0.13** |
| CMR504Z | Pseudo-gene of NADPH:protochlorophyllide oxidoreductase | Porphyrin and chlorophyll metabolism | | 0.36 | 0.17** |
| CMJ021C | Probable 1-acylglycerol-3-phosphate O-acyltransferase | Glycerolipid Metabolism | Acyltransferase | 0.36 | 0.48 |
| CMN281X | Repetitive element | | | 0.36 | 0.12** |
| CMS044Z | Possible transcribed region | | | 0.37 | 0.28 |
| CMP279X | Repetitive element | | | 0.37 | 0.15** |
| CMM301Z | Possible transcribed region | | | 0.38 | 0.19** |
| CMH282C | Similar to trefoil factor | | Trefoil (P-type) domain | 0.38 | 0.34 |
| CMO292X | Retroelement | | | 0.38 | 0.43 |
| CMQ025X | Repetitive element | | | 0.38 | 0.11** |
| CMQ235T | Hypothetical transcript | | | 0.38 | 0.37 |
| CMH148C | Hypothetical protein | | | 0.38 | 0.39 |
| CML181C | Hypothetical protein | | E2F/DP family winged-helix DNA-binding domain | 0.38 | 0.48 |
| CME184T | Hypothetical transcript | | | 0.38 | 0.50 |
| CMM047X | Repetitive element | | | 0.39 | 0.16** |
| CMQ419X | Repetitive element | | | 0.39 | 0.24** |
| CMP250X | Repetitive element | | | 0.39 | 0.12** |
| CMO126C | DUTP pyrophosphatase | Pyrimidine metabolism | dUTPase | 0.39 | 0.27 |

Table A3.11 Continued…

| Gene ID | Annotation | Group | Pfam Domain | July:January Expression Ratio | |
|---|---|---|---|---|---|
| | | | | UV + VIS | VIS |
| CMK165C | Retroelement, alive | | Reverse transcriptase (RNA-dependent DNA polymerase) | 0.39 | 0.26** |
| CMM101X | Repetitive element | | | 0.39 | 0.11** |
| CMS361C | Organelle division protein (ftsZ) | Cell Cycle | Tubulin/FtsZ family, GTPase domain | 0.39 | 0.37 |
| CMI115X | Repetitive element | | | 0.39 | 0.13** |
| CMO002X | Repetitive element | | | 0.40 | 0.17** |
| CMC005C | Similar to trefoil factor | | Trefoil (P-type) domain | 0.40 | 0.36 |
| CMT011X | Repetitive element | | | 0.40 | 0.09** |
| CMH217X | Repetitive element | | | 0.40 | 0.09** |
| CMM107X | Repetitive element | | | 0.40 | 0.12** |
| CMD190C | NADPH:protochlorophyllide oxidoreductase | Porphyrin and chlorophyll metabolism | short chain dehydrogenase | 0.41 | 0.48 |
| CMT007X | Repetitive element | | | 0.41 | 0.17** |
| CMN015X | Repetitive element | | | 0.42 | 0.18** |
| CMM312Z | Possible transcribed region | | | 0.42 | 0.14** |
| CMS426C | Hypothetical protein | | Tetraspanin family | 0.42 | 0.45 |
| CMI188X | Repetitive element | | | 0.43 | 0.27** |
| CMQ160C | Cytochrome c reductase (Complex III) subunit 10(XI) | Oxidative phosphorylation | Carlavirus putative nucleic acid binding protein | 0.43 | 0.26** |
| CMO156X | Repetitive element | | | 0.43 | 0.17** |
| CMM038C | Hypothetical protein | | | 0.43 | 0.47 |
| CME174X | Repetitive element | | | 0.43 | 0.12** |
| CMQ031C | Hypothetical protein | | | 0.43 | 0.38 |
| CMS069X | Repetitive element | | | 0.43 | 0.14** |
| CMP040C | Hypothetical protein | | | 0.43 | 0.24** |
| CML120C | Hypothetical protein | | | 0.43 | 0.48 |

Table A3.11 Continued…

| Gene ID | Annotation | Group | Pfam Domain | July:January Expression Ratio | |
|---|---|---|---|---|---|
| | | | | UV + VIS | VIS |
| CMA018X | Repetitive element | | | 0.44 | 0.14** |
| CME049Z | Possible transcribed region | | | 0.44 | 0.32 |
| CMT044X | Repetitive element | | | 0.44 | 0.20** |
| CMS214X | Repetitive element | | | 0.45 | 0.23** |
| CML281X | Repetitive element | | | 0.45 | 0.25** |
| CMS101C | Probable proliferating cell nuclear antigen | Base Excision Repair | Proliferating cell nuclear antigen, N-terminal domain | 0.45 | 0.48 |
| CMR239Z | Possible transcribed region | | | 0.45 | 0.40 |
| CMT189X | Repetitive element | | | 0.45 | 0.18** |
| CMH098X | Repetitive element | | | 0.46 | 0.21** |
| CMD025C | DNA replication licensing factor (mcm2) | DNA Replication | MCM2/3/5 family | 0.46 | 0.44 |
| CMP183C | Probable 2-epi-valiolone synthase | Phenylalanine, Tyrosine, and Tryptophan biosynthesis | 3-dehydroquinate synthase | 0.46 | 0.49 |
| CME003X | Repetitive element | | | 0.47 | 0.21** |
| CMS365Z | Possible transcribed region | | | 0.47 | 0.50 |
| CMF013X | Repetitive element | | | 0.47 | 0.22** |
| CMF008X | Repetitive element | | | 0.48 | 0.19** |
| CME173C | Retroelement | | Reverse transcriptase (RNA-dependent DNA polymerase) | 0.48 | 0.39 |
| CMQ118C | 60S ribosomal protein L39 (rpl39) | Ribosome | Ribosomal L39 protein | 0.48 | 0.45 |
| CMI012C | Retroelement, dead | | Reverse transcriptase (RNA-dependent DNA polymerase) | 0.48 | 0.43 |

Table A3.11 Continued…

| Gene ID | Annotation | Group | Pfam Domain | July:January Expression Ratio | |
|---|---|---|---|---|---|
| | | | | UV + VIS | VIS |
| CMI190X | Repetitive element | | | 0.48 | 0.23** |
| CMP153X | Repetitive element | | | 0.48 | 0.26 |
| CMM246X | Repetitive element | | | 0.48 | 0.38 |
| CMK090C | Hypothetical protein, conserved | | DnaJ central domain | 0.48 | 0.42 |
| CMH283C | Similar to hedgehog protein | | Hint module | 0.49 | 0.33** |
| CMH129C | Hypothetical protein | | | 0.49 | 0.39 |
| CMN102X | Repetitive element | | | 0.49 | 0.40 |
| CMQ458T | Hypothetical transcript | | | 0.50 | 0.16** |
| CMR078C | Hypothetical Protein | | | 2.00 | 2.40 |
| CMJ248C | Similar to cytochrome c-type biogenesis protein ccmH | | Cytochrome C biogenesis protein | 2.00 | 2.40 |
| CMH080C | Hypothetical protein | | Endonuclease/Exonuclease/phosphatase family | 2.03 | 2.12 |
| CMR201C | Probable mitochondrial ribosomal protein S21 | Ribosome | Ribosomal protein S21 | 2.06 | 2.13 |
| CMF057C | Hypothetical protein, conserved | | | 2.10 | 2.02 |
| CMR185C | Mitochondrial division protein Mda1 | | WD domain, G-beta repeat | 2.13 | 2.78 |
| CMR244C | Probable dehydrogenase | | Periplasmic binding proteins and sugar binding domain of LacI family | 2.15 | 2.35 |
| CML043C | Hypothetical protein, conserved | | PPR repeat | 2.19 | 2.09 |
| CMM131T | Hypothetical transcript | | | 2.25 | 2.48 |
| CML084Z | Possible transcribed region | | | 2.26 | 3.85 |
| CMG192T | Hypothetical transcript | | | 2.28 | 2.25 |
| CMP326C | Hypothetical protein, conserved | | Protein of unknown function, DUF647 | 2.30 | 2.07 |
| CMC143T | Hypothetical transcript | | | 2.30 | 4.51 |

Table A3.11 Continued…

| Gene ID | Annotation | Group | Pfam Domain | July:January Expression Ratio | |
|---|---|---|---|---|---|
| | | | | UV + VIS | VIS |
| CMM072C | Sigma subunit for chloroplast RNA polymerase | | Sigma-70 factor, region 1.2 | 2.33 | 2.04 |
| CMT192C | Unknown mitchondrial carrier protein | | Mitochondrial carrier protein | 2.35 | 2.43 |
| CMO050Z | Possible transcribed region | | | 2.35 | 2.18 |
| CMQ223Z | Possible transcribed region | | | 2.37 | 2.74 |
| CML187T | Hypothetical transcript | | | 2.39 | 3.25 |
| CMQ366Z | Possible transcribed region | | | 2.40 | 2.32 |
| CMN159C | Unknown kinase with aarF domain | | ABC1 family | 2.53 | 2.84 |
| CMM279C | Similar to tellurium resistance protein TerC | | Integral membrane protein TerC family | 2.59 | 2.92 |
| CMT200Z | Possible transcribed region | | | 2.70 | 2.75 |
| CME143C | Smt3-activating enzyme E1 C subunit | | ThiF family | 2.80 | 5.49 |
| CMN105C | ATP-binding cassette, sub-family B (ATM) | | ABC transporter | 2.86 | 3.11 |
| CMI032C | Hypothetical protein | | | 3.13 | 2.63 |
| CMK124C | Similar to peroxisomal targeting signal type 1 (PTS1) receptor PEX5 | Peroxisome | Tetratricopeptide repeat | 3.16 | 2.54 |
| CMF187C | Similar to ubiquitin-conjugating enzyme E2 variant | | Fatty acid hydroxylase superfamily | 3.18 | 3.95 |
| CMT208C | Hypothetical protein | | Protein of unknown function (DUF1674) | 3.27 | 3.12 |

Table A3.11 Continued…

| Gene ID | Annotation | Group | Pfam Domain | July:January Expression Ratio | |
|---|---|---|---|---|---|
| | | | | UV + VIS | VIS |
| CMA106C | Similar to alpha-glucosidase | Protein Processing in the Endoplasmic Reticulum | Glycosyl hydrolases family 31 | 3.31 | 3.45 |
| CMB078C | Similar to suppressor of lin12 SEL1 | | Sel1 repeat | 3.34 | 2.76 |
| CMM081C | Hypothetical protein, conserved | | FAD dependent oxidoreductase | 3.39 | 2.66 |
| CMT050C | Unknown transporter | | EamA-like transporter family | 3.46 | 2.65 |

Table A3.12 Summary of gene expression events exclusively sensitive to UV wavelengths.

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | July:January Expression Ratio UV only |
|---|---|---|---|---|
| CMQ381C | Hypothetical protein, conserved | | | 0.24 |
| CMV125C | Photosystem II protein T (psbT) | Photosynthesis | Photosystem II reaction centre T protein | 0.32 |
| CMV174C | 50S ribosomal protein L14 | Ribosome | Ribosomal protein L14p/L23e | 0.32 |
| CMV075C | C-type cytochrome biogenesis protein Ccs1 | | ResB-like family | 0.36 |
| CMV089C | Photosystem II protein W (psbW) | Photosynthesis | Psb28 protein | 0.36 |
| CMV021C | Initiation factor 3 | Translation Factors | Translation initiation factor IF-3, C-terminal domain | 0.37 |
| CMG089C | Similar to respiratory burst oxidase protein | | Late Protein L2 | 0.38 |
| CMV057C | ORF193 | | mce related protein | 0.39 |
| CMV051C | Phycobilisome rod-core linker polypeptide (cpcG) | Photosynthesis-Anteanna Proteins | Phycobilisome Linker polypeptide | 0.39 |
| CMK062C | Hypothetical protein | | Chromatin assembly factor 1 subunit A | 0.39 |
| CMV241C | ORF515 | | | 0.39 |
| CMV050C | Lipoate-protein ligase B | Lipoic Acid Metabolism | Biotin/lipoate A/B protein ligase family | 0.40 |
| CMH264C | DNA replication licensing factor (mcm3) | DNA Replication | MCM2/3/5 family | 0.40 |
| CMP059Z | Possible transcribed region | | | 0.41 |

Table A3.12 Continued…

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | July:January Expression Ratio UV only |
|---|---|---|---|---|
| CMI175Z | Possible transcribed region | | | 0.42 |
| CMV078C | ORF197 | | Chloroplast import component protein (Tic20) | 0.42 |
| CMV091C | Remnant of 30S ribosomal protein S1 | Ribosome | Alpha/beta hydrolase of unknown function (DUF915) | 0.42 |
| CMR048Z | Possible transcribed region | | | 0.42 |
| CMV006C | Tryptophan synthase alpha chain | Phenylalanine, Tyrosine, and Tryptophan biosynthesis | Tryptophan synthase alpha chain | 0.42 |
| CMV176C | 50S ribosomal protein L5 | Ribosome | Ribosomal protein L5 | 0.42 |
| CMV158C | Allophycocyanin (APC) alpha chain (apcA) | Photosynthesis-Anteanna Proteins | Phycobilisome protein | 0.43 |
| CMN099X | Repetitive element | | | 0.43 |
| CML051Z | Possible transcribed region | | | 0.43 |
| CMV222C | ATP synthase CF0 B' chain (subunit II) | Oxidative phosphorylation | ATP synthase B/B' CF(0) | 0.43 |
| CMM293C | Hypothetical protein | | Methyltransferase domain | 0.44 |
| CMV134C | Acetyl-CoA carboxylase biotin carrier protein | Fatty Acid Biosynthesis | Biotin-requiring enzyme | 0.44 |
| CMG162C | Similar to ferredoxin component (biphenyl dioxygenase, Rieske iron-sulfur component related protein) | | Rieske [2Fe-2S] domain | 0.44 |
| CMK202T | Hypothetical transcript | | | 0.44 |
| CMV023R | TmRNA | | | 0.45 |
| CMM138C | Similar to cyclin dependent kinase | | Protein kinase domain | 0.45 |

Table A3.12 Continued…

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | July:January Expression Ratio UV only |
|---|---|---|---|---|
| CMV160C | Acetylglutamate kinase | | Amino acid kinase family | 0.45 |
| CMS404T | Hypothetical transcript | | | 0.46 |
| CMO157C | Retroelement, dead | | Reverse transcriptase (RNA-dependent DNA polymerase) | 0.46 |
| CMV172C | 50S ribosomal protein L29 (rpl29) | Ribosome | Ribosomal L29 protein | 0.46 |
| CMO306C | Probable myo-inositol dehydrogenase | | Oxidoreductase family, NAD-binding Rossmann fold | 0.47 |
| CMI008T | Hypothetical transcript | | | 0.47 |
| CMG201C | Hypothetical protein | | Cbb3-type cytochrome oxidase component FixQ | 0.47 |
| CMN213C | Similar to molybdenum cofactor biosynthesis protein C (moaC) | Sulfur Relay System | MoaC family | 0.47 |
| CMC007X | Repetitive element | | | 0.48 |
| CMX008C | Similar to hedgehog protein | | Hint module | 0.48 |
| CMT087C | Similar to DNA replication licensing factor MCM8 | DNA Replication | MCM2/3/5 family | 0.48 |
| CMO051C | Similar to Raf-related MAP kinase kinase kinase (MAP3K) | | Protein tyrosine kinase | 0.48 |
| CML214C | Ornithine carbamoyltransferase | Arginine and Proline metabolism | Aspartate/ornithine carbamoyltransferase, Asp/Orn binding domain | 0.48 |
| CMA002Z | Pseudo-gene of NADPH:protochlorophyllide oxidoreductase | Porphyrin and chlorophyll metabolism | | 0.48 |
| CMV191C | Elongation factor Tu | Translation Factors | Elongation factor Tu GTP binding domain | 0.49 |

Table A3.12 Continued…

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | July:January Expression Ratio UV only |
|---|---|---|---|---|
| CMV007C | Initiation factor 2 | Translation Factors | Elongation factor Tu domain 2 | 0.49 |
| CMN083T | Hypothetical transcript | | | 0.49 |
| CMT412C | Isocitrate dehydrogenase (NAD+) subunit 1, mitochondrial | Citrate Cycle (TCA cycle) | Isocitrate/isopropylmalate dehydrogenase | 0.49 |
| CMN106C | Similar to regulator of nonsense transcript stability | | UvrD/REP helicase | 0.49 |
| CMS189C | 40S ribosomal protein S3A (rps3A) | Ribosome | Ribosomal S3Ae family | 0.49 |
| CMA141Z | Pseudo-gene of NADPH:protochlorophyllide oxidoreductase | Porphyrin and chlorophyll metabolism | | 0.50 |
| CML213C | Hypothetical protein, conserved | | Zinc finger, C3HC4 type (RING finger) | 0.50 |
| CMS378Z | Possible transcribed region | | | 0.50 |
| CMH281Z | Possible transcribed region | | | 0.50 |
| CMT400Z | Possible transcribed region | | | 0.50 |
| CMT062T | Hypothetical transcript | | | 2.01 |
| CMI173C | Hypothetical protein, conserved | | Lipase (class 3) | 2.02 |
| CMN152Z | Possible transcribed region | | | 2.02 |
| CMT284C | Similar to permease involved in the uptake of glycerophosphoinositol Git1p | | Sugar (and other) transporter | 2.03 |
| CMM037C | Hypothetical protein | | Protein of unknown function | 2.04 |
| CMN326C | Similar to MAP kinase phosphatase-1 | | Rhodanese-like domain | 2.06 |
| CMF061C | Hypothetical protein | | | 2.06 |
| CMG019C | Nitrate reductase | Nitrogen metabolism | Cytochrome b5-like Heme/Steroid binding domain | 2.06 |

Table A3.12 Continued…

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | July:January Expression Ratio UV only |
|---|---|---|---|---|
| CMJ233C | Hypothetical protein, conserved | | Beta-lactamase superfamily domain | 2.07 |
| CMR145C | Hypothetical Protein | | | 2.11 |
| CMG122T | Hypothetical transcript | | | 2.11 |
| CMB122C | Hypothetical protein | | | 2.12 |
| CMH258C | Hypothetical protein | | | 2.12 |
| CMF188T | Hypothetical transcript | | | 2.13 |
| CMB053C | Probable xylulose kinase | Pentose and glucuronate conversions | FGGY family of carbohydrate kinases, N-terminal domain | 2.14 |
| CMS388C | Hypothetical protein, conserved | | Ubiquinol-cytochrome C chaperone | 2.15 |
| CMJ099C | Formate--tetrahydrofolate ligase | | Formate--tetrahydrofolate ligase | 2.15 |
| CMT061C | Zeta-carotene desaturase | Carotenoid biosynthesis | Flavin containing amine oxidoreductase | 2.21 |
| CMN331Z | Possible transcribed region | | | 2.23 |
| CMM122C | Similar to 3-oxoacyl-(acyl carrier protein) reductase | | short chain dehydrogenase | 2.26 |
| CMK299C | Hypothetical protein, conserved | | | 2.27 |
| CMK034C | Hypothetical protein | | Helix-loop-helix DNA-binding domain | 2.29 |
| CMJ039C | ATP-binding cassette, sub-family A | | ABC transporter | 2.32 |
| CMP056Z | Possible transcribed region | | | 2.34 |
| CMB043C | Activity of bc1 complex ABC1 | RNA Polymerase | Endoribonuclease L-PSP | 2.35 |

Table A3.12 Continued…

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | July:January Expression Ratio UV only |
|---|---|---|---|---|
| CMH265C | Hypothetical protein | | | 2.36 |
| CMP055T | Hypothetical transcript | | | 2.40 |
| CMS309C | Similar to microsomal glutathione S-transferase | Glutathione metabolism | MAPEG family | 2.43 |
| CMT384C | Adenylate kinase | | Adenylate kinase | 2.48 |
| CMA044C | Cryptochrome DASH (CmPHR5) | | DNA photolyase | 2.53 |
| CMO323C | Hypothetical protein | | | 2.58 |
| CMM306C | Hypothetical protein, conserved | | Protein of unknown function (DUF3529) | 2.65 |
| CMQ436C | Glycolate oxidase, peroxysomal | Glyoxylate and dicarboxylate metabolism | FMN-dependent dehydrogenase | 2.65 |
| CMG099C | Hypothetical protein, conserved | | | 2.76 |
| CMT033Z | Possible transcribed region | | | 3.03 |
| CMT416C | Similar to carbonic anhydrase precursor | Nitrogen metabolism | Eukaryotic-type carbonic anhydrase | 3.88 |

Table A3.13 Summary of genes displaying a statistically significant correlation (positive or negative) to seasonal shifts in UV-A and/or UV-B irradiance intensity.

| | | | | UV-B | | UV-A | |
|---|---|---|---|---|---|---|---|
| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | r | p-value | r | p-value |
| CMO306C | Probable myo-inositol dehydrogenase | | Oxidoreductase family, NAD-binding Rossmann fold | -1.00 | 0.004 | -1.00 | 0.005 |
| CMG201C | Hypothetical protein | | Cbb3-type cytochrome oxidase component FixQ | -0.99 | 0.008 | -0.95 | 0.047 |
| CML214C | Ornithine carbamoyltransferase | Arginine and Proline metabolism | Aspartate/ornithine carbamoyltransferase, Asp/Orn binding domain | -0.98 | 0.017 | -0.95 | 0.047 |
| CMG089C | Similar to respiratory burst oxidase protein | | Late Protein L2 | -0.97 | 0.034 | | |
| CMN099X | Repetitive element | | | -0.95 | 0.049 | | |
| CMS309C | Similar to microsomal glutathione S-transferase | Glutathione metabolism | MAPEG family | 0.95 | 0.046 | | |
| CMF061C | Hypothetical protein | | | 0.98 | 0.016 | 1.00 | 0.004 |
| CMB122C | Hypothetical protein | | | 1.00 | 0.003 | 0.98 | 0.025 |

Table A3.14 Summary of the Chao 1 allelic richness estimates for ten genes identified in the July cDNA pyrosequencing library.

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | # of reads | %ID* | Chao1 (95% CI#) |
|---|---|---|---|---|---|---|
| CMH161C | Hypothetical protein | | | 23,797 | 97 | 212 (157-322) |
| CMQ224C | Heat shock protein of Hsp90 family | Protein Processing in endoplasmic reticulum | Hsp90 protein | 2,104 | 97 | 14 (11-34) |
| CMO220C | Similar to UDP-galactose transporter | | UAA transporter family | 849 | 97 | 38 (25-85) |
| CMR398C | Similar to dehydrogenase | | Oxidoreductase family, NAD-binding Rossmann fold | 596 | 97 | 11 (6-42) |
| CMN115C | Similar to Na+-driven multidrug efflux pump | | MatE | 446 | 97 | 54 (25-166) |
| CMS319C | Cytochrome P450, family 51 | Steroid hormone biosynthesis | Cytochrome P450 | 251 | 97 | 10 (7-30) |
| CMS270C | Chaperonin containing TCP1, subunit 4 (delta) | | TCP-1/cpn60 chaperonin family | 102 | 97 | 3 (3-3) |
| CMX009C | Hypothetical protein | | | 50 | 97 | 2 (2-2) |
| CMJ101C | Cimilar to heat shock protein | Protein Processing in endoplasmic reticulum | Hsp20/alpha crystallin family | 26 | 97 | 1 (1-1) |
| CME119C | Cyclin dependent kinase, A-type | | Protein kinase domain | 12 | 97 | 2 (2-0) |

*ID, Identity

# CI, Confidence Interval

Table A3.15 Summary of the Chao 1 allelic richness estimates for ten genes identified in the January cDNA pyrosequencing library.

| Gene ID | Annotation | KEGG Functional Group | Pfam Domain | # of reads | %ID* | Chao1 (95% CI#) |
|---|---|---|---|---|---|---|
| CMH161C | Hypothetical protein | | | 7,535 | 97 | 188 (112-371) |
| CMB054C | Hypothetical protein | | | 1,147 | 97 | 15 (13-30) |
| CMG213C | Proline | | Proline | 578 | 97 | 47 (22-145) |
| CMS124C | Valosin-containing protein | Ribosome biogenesis in eukaryotes | ATPase family associated with various cellular activities | 411 | 97 | 26 (19-46) |
| CMM197C | Hypothetical protein | | | 350 | 97 | 62 (33-163) |
| CMN038C | Hypothetical protein | | Peptidyl-tRNA | 222 | 97 | 2(2-0) |
| CMS294C | Hypothetical protein | | FAD dependent | 103 | 97 | 10 (8-20) |
| CMJ098C | Proline-tRNA ligase | Aminoacyl-tRNA biosynthesis | tRNA synthetase class II core domain | 56 | 97 | 37 (16-114) |
| CMP139C | Hypothetical protein | | | 25 | 97 | 2(2-0) |
| CMI223C | Retroelement | | Reverse | 12 | 97 | 1 (1-2) |

*ID, Identity

# CI, Confidence Interval

Table A3.16 Summary of July cDNAs not corresponding to the *Cyanidioschyzon merolae* strain 10D genome. The library was constructed from non-filtered cyanidial mats sampled at Dragon Spring. Sequences were identified by a BLASTx search against the NCBI-nr protein database with an (E-value <10-5).

| Domain | Gene ID | Annotation | Taxa | Accession | E-value |
|---|---|---|---|---|---|
| **Bacteria** | | | | | |
| | FQGCU5R02H1EMH | Ferredoxin | *Thiomonas* sp. 3As | CAZ88012 | 2.0E-45 |
| | FQGCU5R02G781B | Hypothetical protein | *Nitrococcus mobilis* | ZP_01126498 | 5.0E-36 |
| | FQGCU5R02GQ7EK | Lytic murein transglycosylase | *Agrobacterium tumefaciens* | EHJ98576 | 8.0E-30 |
| | FQGCU5R02I6F8P | Citrate (Si)-synthase | *Ktedonobacter racemifer* | ZP_06969850 | 3.0E-26 |
| | FQGCU5R02GVDHE | Putative Phosphate starvation-inducible protein psiF precursor | *Thiomonas* sp. 3As | CAZ89895 | 4.0E-18 |
| | FQGCU5R02HXW85 | Hypothetical protein | *Hydra magnipapillata* | CBA31936 | 1.0E-15 |
| | FQGCU5R02JW9WI | Unnamed protein | *Frankia alni* | YP_713537 | 2.0E-14 |
| | FQGCU5R02I7BMG | Conserved hypothetical protein | *Methylococcus capsulatus* | AAU91204 | 2.0E-11 |
| | FQGCU5R02GLIQU | Xanthine/uracil permease | *Amycolatopsis mediterranei* | YP_003762611 | 4.0E-11 |
| | FQGCU5R02GVI1J | Conserved hypothetical protein | *Streptomyces albus* | ZP_06591517 | 4.0E-11 |
| | FQGCU5R02GE0RR | Putative Cell wall-associated hydrolase | *Thiomonas* sp. 3As | CAZ89293 | 1.0E-10 |
| | FQGCU5R02ISPUE | Conserved hypothetical protein | *Clostridium botulinum* | ZP_04824114 | 1.0E-10 |
| | FQGCU5R02IN0CM | Hypothetical protein | Uncultured proteobacteria | ADI19242 | 1.0E-09 |
| | FQGCU5R02H54LG | Pyrimidine utilization transporter G | *Streptomyces sp.* | ZP_07276092 | 2.0E-09 |
| | FQGCU5R02FXXAQ | Unknown | *Lawsonia intracellularis* | CAJ55200 | 3.0E-09 |
| | FQGCU5R02FSEL1 | Putative Bifunctional | *Thiomonas* sp. 3As | CAZ87814 | 3.0E-08 |
| | FQGCU5R02FYG1S | Hypothetical protein | *Meiothermus silvanus* | YP_003686657 | 1.0E-05 |
| | FQGCU5R02I3G8B | Hypothetical protein | *Acidimicrobium ferrooxidans* | YP_003109527 | 7.0E-05 |

Table A3.16 Continued…

| Domain | Gene ID | Annotation | Taxa | Accession | E-value |
|---|---|---|---|---|---|
| **Eukarya** | | | | | |
| | FQGCU5R02IY9O0 | Hypothetical protein | *Schistosoma mansoni* | XP_002570463 | 5.0E-21 |
| | FQGCU5R02ICY20 | Hypothetical protein | *Neurospora tetrasperma* | EGZ67539 | 8.0E-18 |
| | FQGCU5R02GBD4U | Hypothetical protein | *Trichoderma atroviride* | EHK48569 | 1.0E-16 |
| | FQGCU5R02IYHSD | Hypothetical protein | *A. anophagefferens* | EGB07523 | 1.0E-13 |
| | FQGCU5R02JN8MS | Hypothetical protein | *Komagataella pastoris* | CCA41173 | 2.0E-12 |
| | FQGCU5R02GGJ7D | Conserved hypothetical protein | *Perkinsus marinus* | XP_002781113 | 7.0E-12 |
| | FQGCU5R02FMLDL | Hypothetical protein | *Sorghum bicolor* | XP_002489151 | 1.0E-11 |
| | FQGCU5R02GXZ24 | RRNA intron-encoded homing endonuclease | *Medicago truncatula* | XP_003614389 | 2.0E-11 |
| | FQGCU5R02JE0O3 | Hypothetical protein | *Trichophyton equinum* | EGE09619 | 4.0E-11 |
| | FQGCU5R02IBEDK | Hypothetical protein | *Schizophyllum commune* | XP_003038164 | 1.0E-09 |
| | FQGCU5R02IB8Q5 | rRNA promoter binding protein-like | *Oryctolagus cuniculus* | XP_002724111 | 2.0E-09 |
| | FQGCU5R02G3JEK | Hypothetical protein | *Selaginella moellendorffi* | XP_002963859 | 5.0E-09 |
| | FQGCU5R02G1F8X | Vacuolar H[+] ATPase 44kD C subunit, isoform E | *Drosophila melanogaster* | NP_725565 | 1.0E-07 |

Table A3.17 Summary of January cDNAs not corresponding to the *Cyanidioschyzon merolae* strain 10D genome. The library was constructed from non-filtered cyanidial mats sampled at Dragon Spring. Sequences were identified by a BLASTx search against the NCBI-nr protein database with an (E-value <10-5).

| Domain | Gene ID | Annotation | Taxa | Accession no. | E-value |
|---|---|---|---|---|---|
| **Bacteria** | | | | | |
| | FQGCU5R02H3MLI | NADH dehydrogenase (quinone) | *K. racemifer* | ZP_06968724 | 8.0E-31 |
| | FQGCU5R02GJSNY | Alanine dehydrogenase/PNT | *P. acnes* | EGR92525 | 2.0E-25 |
| | FQGCU5R02IOEWI | Putative disulfide reductase | *Thiomonas* sp. | CAZ88011 | 8.0E-23 |
| | FQGCU5R02JDJL7 | Putative two-component response regulator | *Vibrio anguillarum* | YP_004577995 | 5.0E-22 |
| | FQGCU5R02IRGPC | Type IV pilus assembly protein PilB | *Clostridium* sp. | ZP_08129062 | 1.0E-20 |
| | FQGCU5R02FVPVY | rRNA promoter binding protein | *Brugia malayi* | XP_001891797 | 1.0E-17 |
| | FQGCU5R02GIEME | Transposase | *S. acidophilus* | YP_004718564 | 2.0E-13 |
| | FQGCU5R02H6JSZ | Transposase IS4 family protein | *A. acidocaldarius* | AEJ42198 | 5.0E-13 |
| | FQGCU5R02FRSBS | aIF-2BI family translation initiation factor | *Petrotoga mobilis* | YP_001567857 | 9.0E-13 |
| | FQGCU5R02G2N68 | D-alanine/D-alanine ligase | *Thermotogales* sp. | ZP_07577446 | 1.0E-09 |
| | FQGCU5R02I4NVF | Hypothetical protein | Uncultured bacteria | ADI19242 | 3.0E-09 |
| | FQGCU5R02JYS1L | Transposase IS4 family protein | *Thauera* sp. | YP_002353947 | 9.0E-07 |
| | FQGCU5R02FIXGI | Transposase IS4 family protein | *Acidithiobacillus caldus* | ZP_05294015 | 1.0E-05 |
| | FQGCU5R02I8STW | Universal stress family protein | *Aerococcus viridans* | ZP_06808121 | 1.0E-05 |
| | FQGCU5R02JJBKS | HSP20-like chaperone | *Dokdonia donghaensis* | ZP_01050947 | 9.0E-05 |

Table A3.17 Continued…

| Domain | Gene ID | Annotation | Taxa | Accession no. | E-value |
|---|---|---|---|---|---|
| **Eukarya** | | | | | |
| | FQGCU5R02F3P2Q | Hypothetical protein | *Tribolium castaneum* | EFA07645 | 4.0E-29 |
| | FQGCU5R02GIN6E | Leucine-rich repeat-containing protein | *D. fasciculatum* | EGG17880 | 1.0E-21 |
| | FQGCU5R02GF2KD | Hypothetical protein | *Neurospora tetrasperm* | EGZ67539 | 1.0E-20 |
| | FQGCU5R02J2W77 | Glutaminase | *Dugesia japonica* | BAG16389 | 1.0E-19 |
| | FQGCU5R02GXFV3 | Conserved hypothetical protein | *Trichinella spiralis* | XP_003370331 | 1.0E-16 |
| | FQGCU5R02IFLLT | Hypothetical protein | *A. anophagefferens* | EGB07523.1 | 2.0E-14 |
| | FQGCU5R02HR6NQ | RRNA intron-encoded homing endonuclease | *Medicago truncatula* | XP_003614387 | 7.0E-14 |
| | FQGCU5R02HZ2ND | Hypothetical protein | *A. anophagefferens* | EGB07523 | 1.0E-13 |
| | FQGCU5R02HW1RT | Cyp-like protein | *Mytilus trossulus* | ABY21303 | 1.0E-13 |
| | FQGCU5R02FK4HI | Hypothetical protein | *D. purpureum* | XP_003293438 | 8.0E-11 |
| | FQGCU5R02IUWZ2 | Hypothetical protein | *Trichophyton equinum* | EGE09619 | 8.0E-09 |
| | FQGCU5R02GE239 | Hypothetical protein | *S. moellendorffii* | XP_002967759 | 4.0E-07 |
| | FQGCU5R02G2QL4 | Predicted protein | *Nematostella vectensis* | XP_001624693 | 1.0E-06 |
| | FQGCU5R02HBU04 | Glycine-rich RNA-binding protein | *Perkinsus marinus* | XP_002769641 | 9.0E-05 |

# APPENDIX B

## SUPPLEMENTAL INFORMATION FOR CHAPTER 4:

## CYANIDIALES DIVERSITY IN YELLOWSTONE NATIONAL PARK

Table A4.1 Primers used in this study

| Primer | Primer Sequence | Target | Reference/Source |
|---|---|---|---|
| rbcL-90F | CCATATGCYAAAATGGGATATTGG | *rbcL* | (Ciniglia, 2004) |
| rCR | GCWGTTGGTGTYTCHACWAAATC | *rbcL* | (Ciniglia, 2004) |
| MR-rbcLR | TTCAAAGAAACCTTGAGGAAGATCT- | *rbcL* | This Study |
| DBV182- rbcLR | TTCAAAGAACAATCCCTGAGGCAAATTTACA | *rbcL* | This Study |
| Sybil-rbcLR | TCCAAGATAATCGAGTAGTAA | *rbcL* | This Study |

Table A4.2 Sample information and GenBank accession number for the *rbcL* gene sequences used for phylogenetic comparisons

| Taxa | Source | *rbcL* |
|---|---|---|
| Bangiales | | |
| *Bangia atropurpurea* | SAG 33.94 | AY119770 |
| *Porphyra purpurea* | GenBank | NC_000925 |
| Compsopogonales | | |
| *Compsopogon coeruleus* | SAG 36.94 | AF087116 |
| Cyanidiales | | |
| *Cyanidioschyzon merolae* | | |
| | DBV 201 JAVA | AY119765 |
| | DBV 001 NAPS | AY119766 |
| | Pisciarelli-D1 | AY541320 |
| *Cyanidium caldarium* | | |
| | DBV 019 SIPE | AY541297 |
| | MR5-5 | DQ916751 |
| | MR6-C35 | DQ916752 |
| C. sp. — Monte Rotaro | Monte Rotaro | AY391368 |
| C. sp. — Sybil | Sybil cave | AY391369 |
| C. sp. — Pisciarelli | Pisciarelli-C2 | AY541318 |
| *Galdieria daedala* | IPPAS P508 | AY541302 |
| *Galdieria maxima* | IPPAS P507 | AY541302 |
| *Galdieria partita* | IPPAS P500 | AB18008 |
| *Galdieria sulphuraria* -A | | |
| | SAG 108.79 | AY119767 |
| | UTEX 2393 | AF233069 |
| | DBV 011 CEMD | AY541303 |
| | DBV 018 CNASC | AY541304 |
| | DBV 015 NAFG | AY541305 |
| | DBV 017 NASF | AY541306 |
| | DBV 021 MEVU | AY541307 |
| | DBV 074 JAVA | AY541308 |
| | DBV 135 AZUF | AY541309 |
| | Pisciarelli-D15 | AY541322 |
| | Pisciarelli-E11 | AY541324 |
| | MR5-C17 | DQ916746 |
| | SP3-C2 | DQ916749 |
| *Galdieria sulphuraria* -B | | |
| | DBV 009 VTNE | AY119768 |
| | DBV 012 BNTE | AY541310 |
| | DBV 063 AGCS | AY119769 |
| | DBV 002 NAPS | AY541311 |
| | Pisciarelli-B20 | AY541316 |
| Unidentified | Pisciarelli-C16 | AY541319 |

Table A4.2 Continued…

| Taxa | Source | *rbcL* |
|---|---|---|
| Florideophycidae | | |
| *Palmaria palmata* | GenBank | U28421 |
| *Thorea violacea* | SAG 51.94 | AF506271 |
| Porphyridiales | | |
| *Bangiopsis subsimplex* | PR21 | AY119772 |
| *Rhodella_violacea* | SAG 115.79 | AY119776 |
| *Stylonema alsidii* | SAG 2.94 | AY119779 |
| Rhodochaetales | | |
| *Rhodochaete parvula* | UTEX LB 2715 | AY119777 |