Winter 2018

POPULATION GENETICS AND DESICCATION STRESS OF PORPHYRA UMBILICALIS KÜTZING IN THE GULF OF MAINE

Yuanyu Cao
University of New Hampshire, Durham

POPULATION GENETICS AND DESICCATION STRESS OF *PORPHYRA UMBILICALIS* KÜTZING IN THE GULF OF MAINE

BY

YUANYU CAO

B.A., Jimei University, Xiamen, Fujian, Peoples Republic of China, 2008

M. S., Jimei University, Xiamen, Fujian, Peoples Republic of China, 2013

DISSERTATION

Submitted to the University of New Hampshire

in Partial Fulfillment of

the Requirements for the Degree of

Doctor of Philosophy

in

Genetics

December 2018

This dissertation has been examined and approved in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Genetics by:

Dissertation Dir. Anita S. Klein, Assoc. Professor of Biological Sciences

Estelle M. Hrabak, Assoc. Professor of Molecular, Cellular, & Biomedical Sci.

Matthew D. MacManes, Asst. Professor of Molecular, Cellular, & Biomedical Sci.

Arthur Mathieson, Professor of Plant Biology

W. Kelley Thomas, Professor of Molecular, Cellular, & Biomedical Sci.

On September 13, 2018

DEDICATION

To my husband, Mengmeng.

To my parents, Heming Cao and Chenmei Wu,

and to Meng's parents.

ACKNOWLEDGEMENTS

I would like to thank my committee for their financial, technical, and manuscript support: Drs Anita S. Klein, W. Kelley Thomas, Estelle M. Hrabak, Arthur Mathieson, and Matthew D. MacManes. I owe many thanks to a number of people who helped me during my Ph.D years. Dr. David J Ashlin and Dr. Andy Hanneman from the Glycomics Center, UNH, provided valuable technical support and manuscript advice for the GC-MS study. Dr. Vernon Reinhold from the Glycomics Center for access to the GC-MS equipment. Dr. Elizabeth Ficko-Blean from Sorbonne Universités provided technical support and manuscript advice for the phylogenetic tree. Dr. Renée L. Eriksen and Dr. Lindsay A. Green-Gavrielidis provided valuable tissue samples, RNA-Seq data and manuscript suggestions for the population study. Dr. Paul C. Tsang and Joseph Miseirvitch provided technical support and equipment for real-time PCR analysis. Dr. Matthew D. MacManes, Anthony S. Westbrook and Dr. Feseha Abebe-Akele for software installation, bioinformatic support and server usage guidance. Dr. Subhash Minocha provided RNA extraction equipment and RNA extraction technical suggestions. Dr. Christopher Neefus provided lab space for DNA extraction. Jobriah Anderson performed some Sanger sequencing and Stephen Simpson at Hubbard Center for Genome Studies at UNH performed the RNA seq library preparation and sequencing. The *Porphyra* RCN provided access to the *Porphyra umbilicalis* genome database prior to publication, and was funded by NSF 05CH11231 (to Susan H. Brawley, Elizabeth Gantt, Arthur R. Grossman and John Stiller). I also want to thank undergraduates David Macomber and Anna Haber for help with DNA extraction and PCR amplification.

Financial support was provided by by the USDA National Institute of Food and Agriculture Hatch Project accession 1004051. Additional support also was provided by the

Leslie S. Hubbard Marine Program Endowment (to L. Green-Gavrielidis) NOAA contract NA060AR4170108 (to SHB) and by the Joint Genome Institute (U. S. Dept. of Energy) under a Community Science Program award from the Office of Science of the U. S. Dept. of Energy under Contract No. DE-AC02-05CH11231 (to Susan H. Brawley, Elizabeth Gantt, Arthur R. Grossman and John Stiller).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

POPULATION GENETICS AND DESICCATION STRESS OF *PORPHYRA UMBILICALIS* KÜTZING IN THE GULF OF MAINE

By

Yuanyu Cao

University of New Hampshire

The red alga *Porphyra umbilicalis* Kützing is an ecologically and economically important marine macroalga in the Northern Atlantic. *Porphyra umbilicalis* has a broad distribution within the North Atlantic. In the Northeast Atlantic, it is dioecious and reproduces both sexually and asexually, while in the Northwest Atlantic only asexual reproduction has been observed. As a high intertidal alga, *P. umbilicalis* regularly experiences desiccation and rehydration cycles with the tidal cycles, so it has high tolerance towards various abiotic stresses. The present work attempts to understand the population structure in asexual populations of *P. umbilicalis* in the Gulf of Maine by applying putative single nucleotide polymorphisms (SNP) markers developed from transcriptome data (Chapter 1). In order to understand the desiccation tolerance of *P. umbilicalis,* the contents of putative compatible solutes were measured and the genes involved in the synthesis of these solutes were analyzed in response to desiccation and rehydration treatments (Chapter 2). In addition, a comparative transcriptomic analysis was performed using *P. umbilicalis* under fresh, dehydrated, desiccated and rehydrated conditions in order to gain insights into the mechanisms of its desiccation tolerance in responses to water loss and water gain (Chapter 3). My work represents the first attempt to develop a suitable bioinformatic pipeline for RNA-seq to detect SNP markers for the red alga *P. umbilicalis* and to apply these SNP markers for population analysis. The compatible solutes study verifies the occurrences of

nanomolar concentrations of trehalose in *P. umbilicalis* for the first time and identifies additional genes, possibly encoding trehalose phosphate synthases. The transcriptome study suggested distinct molecular responses may occur during dehydration and desiccation and confirmed that the rehydration-induced responses play an important role in the mechanisms of desiccation tolerance in *P. umbilicalis*.

INTRODUCTION

The term “seaweed” traditionally includes only macroscopic and multicellular marine algae. Macroalgae consist of over 25,000 diverse species with different morphologies and bioactive properties (Raja et al. 2008, Holdt and Kraan 2011). They are classified according to their pigmentation into brown (Phaeophyta), red (Rhodophyta), and green (Chlorophyta) algae. Macroalgae are of great economic and ecological importance. They are widely used as food, as ingredients in cosmetics and fertilizers, and in hydrocolloid production such as agar and alginate (Chan et al. 2006). The ecological functions of macroalgae include supplying oxygen as one of the primary producers in the marine food chain, removing heavy metals and biomonitoring heavy metal populations (Chan et al. 2006).

Red algae (Rhodophyta) are a large, morphologically diverse eukaryotic lineage of ~6,500 species (Cole 2010). The oldest fossil record of red algae dates back 1.6 billion years, showing that they are one of the most ancient multicellular lineages (Bengtson et al. 2017). The Rhodophyta contain a “primary” plastid surrounded by a double membrane formed by a primary endosymbiotic event(s), in which a heterotrophic protist engulfed a free-living cyanobacterium for food by phagocytosis and they evolved to become an endosymbiont in the cell (Burki et al. 2016). Multicellular red algae also have some the most complex life cycles in living organisms, such as the transition from unicellularity to complex multicellular bodies and trigenetic life cycles. During the trigenetic life cycle, the haploid unicellular meiotic tetraspores germinate to a

male or a female multicellular gametophyte thallus (Cole and Sheath 1995).

The genus *Porphyra* belongs to the family Bangiaceae, in the class Bangiophyceae of the phylum Rhodophyta (Sutherland et al. 2011). There are more than 150 species in *Porphyra sensu lato,* but there are many fewer species in the genus *Porphyra* (Sutherland et al. 2011). The red alga *Porphyra* is distributed from shallow subtidal to high intertidal habitats throughout the world, and are highly abundant in boreal and cold waters. However, some species are also found in tropical regions (Kapraun and Lemus 1987). The reproductive phases among the *Porphyra* species can vary greatly. The sexual heteromorphic life history of *Porphyra* includes a gametophytic phase and a sporophytic conchocelis phase. During the gametophytic stage, diploid zygotospore was formed by fertilizing haploid male and haploid female gametes. Zygotospores settle onto calcareous substrata and germinate into the filamentous sporophyte conchocelis. The conchosporangial branch releases conchospores that divide meiotically to produce new haploid gametophytes.

Porphyra umbilicalis Kützing

The red alga *Porphyra umbilicalis* Kützing is found on rocky substrata from the middle to high intertidal. Problems with the application of the name and concept of the entity as "*Porphyra umbilicalis"* led to confusion over its identity for more than 250 years (Brodie et al. 2008). Due to the lack of distinguishing morphological and ecological characters within the genus *Porphyra*, the identification of *P. umbilicalis* has been difficult. Thanks to the elucidation of DNA sequences of moderately conserved genes (ribulose-bisphosphate carboxylase oxygenase, rbcL_rbcS (Fredericq et al. 1996), a restriction fragment length polymorphism (RFLP) assay was successfully applied to discriminate *P. umbilicalis* from other

Porphyra species *sensu lato* (Teasdale et al. 2002). According to Sutherland et al. (2011), most cultivars in Asia are probably species of *Pyropia*, which is commonly confused with *Porphyra*. The confirmed distribution for *P. umbilicalis* is currently the North Atlantic and its global distribution needs further reviews (Brodie et al. 2008).

Populations of *P. umbilicalis* in the NE Atlantic are known to reproduce both sexually and asexually (Brodie and Irvine 2003); however, according to current data (Blouin et al. 2007, Gantt et al. 2010, Bird and McLachlan 1992), they are only known to reproduce asexually in the NW Atlantic. Sexual reproduction in *P. umbilicalis* follows the typical Bangiophyte alternation of generations as discussed above (Drew 1949, Graham et al. 2016). Asexual reproduction in the NW Atlantic consists of the release of haploid neutral spores from the edge of blades; these neutral spores germinate and grow directly into haploid blades.

In order to understand the evolutionary history of *Porphyra umbilicalis* and the absence of sexual reproduction in the NW Atlantic, two studies were conducted to analyze the genetic variation and population structure of NW and NE Atlantic *P. umbilicalis* populations. However, these two studies produced somewhat contradictory results regarding the genetic variation of *P. umbilicalis* populations along the Gulf of Maine. The first study used Amplified Fragment Length Polymorphism (AFLP) technology and found 41 distinctive genotypes from 45 individuals at two sites in Maine, suggesting high genetic diversity for two geographically adjacent populations (Blouin and Brawley 2012). However, the high genetic diversity in *P. umbilicalis* from these two populations in Maine was still lower than that of a sexually reproducing population from Sidmouth, England (Blouin and Brawley 2012). A second study that used eight expressed sequence tag-simple sequence repeats (EST-SSR) markers identified six genotypes in seven asexual populations of *P. umbilicalis* from Maine to New Hampshire

(Eriksen et al. 2016). The genetic diversity was highest in the southernmost population from New Hampshire, while the lowest genetic diversity was observed in a population that inhabited an unusual estuarine tidal rapid site. While another population that also inhabited an atypical environment showed moderate genetic diversity. The genetic diversity was relatively high in these seven populations. However, the genetic diversity reported by the EST-SSR markers (Eriksen et al. 2016) was much lower than that reported by the AFLP markers (Blouin and Brawley 2012). There are possible explanations for the contradictory patterns from these studies are first, the inclusion of non-target DNA in the AFLP study, specifically the DNA from cryptic contaminants that can interfere with the AFLP analysis, leading to an increase in observed genetic diversity; second, the EST-SSR markers (Eriksen et al. 2016) used in the second study were developed from protein coding regions. Expansion and contraction of SSR regions in the protein coding regions may be more functionally constrained than SNPs in the same region; third, of the eight EST-SSR used only three polymorphic EST-SSR markers were polymorphic in the second study. Eriksen et al. 2016 data were more consistent with expectations for asexual reproduction of clones. However, the limited number of polymorphic markers used may have limited the ascertainment of overall genetic diversity.

Stress tolerance mechanisms

As an intertidal alga, *Porphyra umbilicalis* is exposed to various abiotic stresses, such as daily and seasonally variable temperatures, high levels of irradiance (including UV), and severe desiccation stress (Brawley et al. 2017). Southward (1958) stated that tolerance to desiccation was the most important factor determining the zonation of intertidal organisms and whether an alga can survive in the intertidal zone. The red alga *P. umbilicalis* experiences alternations between hydrated and air-dry environments during tidal cycles so that it should be equipped with

complex response mechanisms at morphological, physiological and molecular levels to survive the severe loss of water. In land plants, reactive oxygen species (ROS) defense, repression of membrane phase transition and formation of cellular glass are the three major desiccation tolerance mechanisms (Liu 2009).

Several previous studies examined distinct aspects of desiccation tolerance in *P. umbilicalis.* One study examined the differences of reactive oxygen species (ROS) defense, the transition of membrane phase and the cytoplasm between the desiccation tolerant species *P. umbilicalis* and the desiccation sensitive species *Pyropia yezoensis* (Ueda) M.S.Hwang & H.G.Choi during desiccation (Liu 2009). Kim et al. (2009) compared the growth rate and nitrate uptake between *P. umbilicalis* and a sublittoral species *Wildemania amplissima* (Kjellman) Foslie (as *Porphyra amplissima* (Kjellman) Setchell & Hus) under desiccation. The drop of grow and nitrate uptake rates were higher in *W. amplissima* compared with those in *P. umbilicalis* under desiccation. A detailed introduction of desiccation tolerance mechanisms in algae follows below.

Molecular studies

The molecular studies of algae started in the 1990s (C. Chan et al. 2006). The expressed sequence tag (EST) approach, which is a cost-efficient and quick method for novel gene discovery, spurred the development of algal research at the molecular level. In red algae, the EST approach has been applied to *Pyropia yezoensis* (as *Porphyra yezoensis*, Nikaido et al. 2000, Asamizu et al. 2003) and *Porphyra umbilicalis* (Chan et al. 2012a). In *P. umbilicalis,* EST libraries were constructed under different abiotic stresses, such as light stress, desiccation stress, etc. (Chan et al. 2012a).

With the development of next-generation sequencing technologies, completed genomes are available for representatives of a number of major groups of organisms and are offering remarkable insights into fields such as comparative evolution, molecular biology, biochemistry, physiology, and developmental biology. Algae have also entered the arena of large-scale genomics. In red algae, the completed nuclear genomes for *Cyanidioschyzon merolae* P.De Luca, R.Taddei & L.Varano (Matsuzaki et al. 2004), *Galdieria sulphuraria* (Galdieri) Merola (Barbier et al. 2005), *Chondrus crispus* Stackhouse (Collén et al. 2013), *Gracilariopsis chorda* (Holmes) Ohmi (Lee et al. 2018), *Porphyridium purpureum* CCMP 1328 (Bhattacharya et al. 2013), *Pyropia yezoensis* strain U-51(Nakamura et al. 2013) and *Porphyra umbilicalis* Kützing (Brawley et al. 2017) have been sequenced.

Aquaculture

In Asia, *Porphyra sensu lato* is an important source of food due to its high protein content, free amino acids and a high ratio of ω3: ω6 fatty acids (Mouritsen 2013). The culture of *Porphyra* began in China more than 1000 years ago. The greatly improved production of *Porphyra* occurred in 1950s after the elucidation of the life history of *Porphyra* by Drew (1949), who demonstrated an alternation between the blade and conchocelis phases in *Porphyra.*

According to Food and Agriculture Organization of the United Nations (2010), *Porphyra* contributes approximately US$1.5 billion to the global trade of aquaculture production. Over the last 69 years, China and Japan have developed advanced techniques of breeding conchocelis and collecting conchospores for *Pyropia* ever since the discovery of complete life history of *Porphyra* (Drew 1949). However, the current *Porphyra* industry, especially for *P. umbilicalis,* mainly relies on the exploitation of natural populations or primitive aquaculture methods

(García-Jiménez and Robaina 2015), or requires extensive and expensive cultivation and maintenance of conchocelis for *Pyropia* (Blouin et al. 2007). Reliable cultivation strategies are needed to support the sustainable growth of the *P. umbilicalis* industry in the North Atlantic, with the widespread recognition of its potential importance to the economy and food security.

Strain selection has been applied in microalgae and macroalgae for biofuels and co-products (Nascimento et al. 2013, Wensel 2018), and sugar (Enquist-Newman et al. 2014). The red alga *Pyropia* is considered to be the most well-established, high-value cultivated genus of macroalgae (Robinson et al. 2013). The selection of *Porphyra* spp. has been practiced since 1949. Specifically in Asia, strain selection and the cold-storage of seeded nets have enabled an extended growing season and also have improved the yields and quality of harvested *Pyropia*. The approach for strain improvement has been based on selection of a few strains from the wild and/or mutation/hybridization of these strains to deliver novel or improved attributes (as *Porphyra yezoensis* Niwa et al. 2004, Zhang et al. 2011). For example, *Pyropia yezoensis* has been selected for high temperature tolerance (Zhang et al. 2011). However, the lack of knowledge of existing genetic diversity in such approaches focuses on short-term improvement instead of addressing the long-term opportunity (Robinson et al. 2013). Thus, more genetic information is needed for better strategic approaches for strain selection.

Integrated multi-trophic aquaculture (IMTA) involves the farming of species from different trophic positions or nutritional levels in the same system. Eutrophic inputs of nitrogen and phosphorus from fish farming can be reduced during combined aquaculture of marine macroalgae and fish (Folke et al. 1994, Krom et al. 1995, Fei et al. 1998, Day et al. 2009). Implementation of IMTA can improve profits and help the environment. In China, IMTA has

been practiced for many decades and shows promising results to improve economics and environmental impacts of aquaculture (Fang et al. 2016).

For classic commercial production, nets are seeded with *Pyropia* conchospores derived from conchocelis. The seeded nets go through the nursery period until the germlines reach a certain size and then the nets are harvested for up to 120 days, in coastal waters. Another approach has been proposed for *Porphyra* species with an asexual life history, such as *P. umbilicalis* in the NW Atlantic. Using asexual strains may significantly reduce the time as the economically important gametophytic blade of *P. umbilicalis* only grows for 3-5 months each year (Blouin et al. 2007). Blouin et al. (2007) successfully used asexual neutral spores from *P. umbilicalis* in IMTA culture of a salmon species in Maine to demonstrate the feasibility of neutral spore-seeding of nori co-cultured with salmon.

However, shortening the culture period while neglecting the importance of maintaining diverse genetic resources may negatively affect economic activity. For *Gracilaria chilensis* C.J.Bird, McLachlan & E.C.Oliveira, bypassing sexual reproduction without a sustainable strategy resulted in homogeneous strains that could not be genetically improved without hybridizing with other cultivated strains or wild varieties (Guillemin et al. 2008). Homogeneous strains or stains with reduced genetic diversity are vulnerable to diseases and may limit future genetic improvement. The choices of genetic materials, sustainable strategies, and breeding programs are important to maintain genetic diversity when only using a vegetative system for reproduction for strain selection.

Algae and microbe interaction

Algal-bacterial associations are essential, as algae synthesize and release large amounts of organic compounds, serving as chemo-attractants and/or nutrient source for microbes (Haas et al. 2011). The microbial community living on the surface and inside algae may be selected by secondary metabolites and/or exudates of algae (Goecke et al. 2010, Persson et al. 2011). A range of symbiotic, pathological and opportunistic interactions between macroalgae and bacteria have been discovered (Goecke et al. 2010). It is suggested that marine macroalgae harbor species-specific and temporally adapted epiphytic bacterial biofilms on their surfaces (Lachnit et al. 2011). Microbes associated with nitrogen fixation, mineralization of organic substrates, and supplying carbon dioxide are beneficial microbe interactions (Cole 1982). For example, some algae are known to acquire vitamin B12 through a symbiotic relationship with bacteria (Croft et al. 2005). The red alga, *Laurencia dendroidea,* releases secondary metabolites, such as terpenes, that have high biocidal and anti-epibiosis activity (Oliveira et al. 2012).

Technologies, such as 16S rRNA gene library analysis, 16 S rDNA amplicon sequencing, and denaturing gradient gel electrophoresis (DGGE), have been applied to analyze the microbial communities associated with the green alga *Ulva* spp. (Longford et al. 2007, Tujula et al. 2010, Burke et al. 2011, Lachnit et al. 2011), with brown algae (Staufenberger et al. 2008, Lachnit et al. 2011) and with some red algae (Longford et al. 2007, Lachnit et al. 2011). More recently, metagenomic approaches have been used in microalgae (Toulza et al. 2012). In *Porphrya umbilicalis,* large bacterial communities were found associated with both the surface and inside the cell wall, by high-throughput 16S rDNA amplicon pyrosequencing (Miranda et al. 2013). Identification of microbial communities suggested that some bacteria likely provided essential morphogenetic and beneficial nutritive factors to the alga and may have effects upon evolution and function of the alga (Miranda et al. 2013).

Abiotic stress

Adaptation to unfavorable environmental conditions is crucial for the survival of all living organisms. Intertidal algae are periodically exposed to air where they experience a variety of stressful environmental conditions, such as high ultraviolet (UV) irradiation, temperature, desiccation and osmotic stress. As macroalgae do not possess an impermeable cuticle, they cannot avoid desiccation but only tolerate it (Lüning 1990). Under extreme abiotic stress, organic molecules such as lipids, proteins, and nucleic acids are damaged and/or degraded (Miranda 2011). The vertical distribution of species in the intertidal zone corresponds to their distinctive stress tolerance (Davison and Pearson 1996).

In green algae, many studies have focused on stress tolerance mechanisms, i.e. Charophyte green algae (Holzinger and Pichrtová 2016) and in the unicellular green alga *Dunaliella salina* Teod (Ramos et al. 2011). In Charophyte green algae, aggregation of cells, flexible cell walls, mucilage production and accumulation of osmotically active compounds are the most common desiccation tolerance strategies (Holzinger and Pichrtová 2016). However, only a few studies have focused on desiccation tolerance strategies in red algae. Unravelling the mechanisms underlying desiccation tolerance is crucial to understand the distribution of individual algal species within the intertidal zone.

Morphological and cellular alteration

In the red alga *Pyropia columbina* (Montagne) W.A.Nelson (as *Porphyra columbina* Montagne), it was observed that a loss of ~96% of the water content resulted in morphological and cellular alterations (Contreras-Porcia et al. 2010). There was no difference in the water loss rate between the desiccation sensitive red alga *Pyropia yezoensis* (as *Porphyra yezoensis*) and

desiccation tolerant red alga *Porphyra umbilicalis*. Both species lost about 95% of their water in the first two hours of desiccation and contained similar final relative water content (Liu 2009). Thus, water loss rate cannot account for the high desiccation tolerance in *P. umbilicalis.*

The cellular components that are most sensitive to desiccation stress are the membranes (Oliver et al. 1998). The ability to maintain the integrity of selectively permeable membranes and to retain cellular solutes within cells has been used as a measure of tolerance to dehydration/desiccation in terrestrial plants and in intertidal algae (Hurd and Dring 1991, Davison and Pearson 1996). During desiccation, terrestrial plants contain some special molecules in the phospholipid head groups in the membrane, which helps to prevent the formation of the gel-state (Vertucci and Farrant 1995). In cyanobacteria, the presence of compound 20:3ω3 contributes to membrane fluidity in *Nostoc commune* UTEX 584 (Olie and Potts 1986).

It is suggested that the cause of *Porphyra umbilicalis*'s excellent tolerance to desiccation is related to its ability to maintain plasma membrane integrity when compared with desiccation sensitive alga *Pyropia yezoensis* (as *Porphyra yezoensis,* Liu 2009). Liu (2009) measured the leakage of electrolytes and amino acids from cells during rehydration after exposure to different levels of desiccation in desiccation-sensitive *P. yezoensis* and in desiccation-tolerant *P. umbilicalis*. The plasma membrane of *P. yezoensis* was damaged at low water content, while no clear relationship between low water content and amino acid leakage was observed in *P. umbilicalis*. The thylakoid membranes in *P. yezoensis* lost constant spacing and appeared fused in places, whereas those of *P. umbilicalis* still maintained parallel and normal during rehydration after severe water loss. The transition of the phospholipid components of the membranes from a liquid crystalline phase to a gel-like state during desiccation increased the loss of solutes (Crowe et al. 1984). However, both *P. umbilicalis* and *P. yezoensis* membranes remained in a liquid

crystalline state when desiccated, and thus inhibition of membrane phase transition was not the cause for the difference in desiccation tolerance between the two species (Liu 2009). Instead, the author showed that the cytoplasm of *P. umbilicalis* formed a more stable glass with a stronger hydrogen bonding network, which prevented membrane movement and fusion (Liu 2009).

Some green algae have unique cell wall structure and composition, such as polysaccharides that may help prolong water-holding capacities when the alga is exposed to desiccating conditions (Domozych et al. 2012). For example, the formation of mucilage layers from pectic proteins in green algae can increase water holding capacities and it is ecologically relevant in habitats with fluctuating water regimes (Fuller 2013). The mucilage layer has also been suggested to play a significant role during UV stress (Lütz et al. 1997). Callose, which increases significantly upon long-term desiccation stress, is a cell wall polysaccharide that imparts flexibility to the cell wall. Callose can protect cell walls from desiccation-induced damage (Sørensen et al. 2011, Herburger and Holzinger 2015, Holzinger and Pichrtová 2016). Callose is localized in terminal cell walls or in the corners between cells when exposed to desiccation in green algae (Herburger and Holzinger 2015). In *Porphyra umbilicalis,* the extracellular matrix (ECM) component of the cell wall is a highly hydrated agar, porphyrin, which limits desiccation and contributes to the flexibility of the cell during osmotic stress (Ficko-Blean et al. 2015).

Photosynthesis

Photosynthesis is the process by which light energy is converted to chemical energy whereby carbon dioxide and water are converted to organic molecules. The photosynthetic activity of macroalgae has been extensively studied (Neale 1987, Long et al. 1994) under

different patterns of seasonal acclimation and under various stress conditions (Larkum and Barrett 1983, Smith and Melis 1987, Lüning 1990).

Photosynthesis is well known to be affected by desiccation in algae. The photosynthetic rate usually goes up upon short-term dehydration. The net photosynthesis of the tropical intertidal seaweed *Ahnfeltiopsis concinna* (J. Ag.) Silva et DeCew increased during short term water loss compared with that when it is fully hydrated (Beach and Smith 1997). Similar findings were reported by Johnson et al. (1974) and Quadir et al. (1979), who found that upper intertidal algae, such as *Iridaea flaccida* (Setchell & N.L.Gardner) P.C.Silva*, Pyropia perforate* (J.Agardh) S.C.Lindstrom (as *Porphyra perforata* J.Agardh)*, Fucus distichus* Linnaeus*,* and *Endocladia muricata* (Endlicher) J.Agardh) had higher photosynthetic rates in air than in water. During long term desiccation, however, photosynthetic rate was found to decrease in aeroterrestrial green algae (Holzinger and Karsten 2013), the red alga *Pyropia haitanensis* T.J.Chang & B.F.Zheng (as *Porphyra haitanensis,* (T.J.Chang & B.F.Zheng) N.Kikuchi & M.Miyata, Zou and Gao 2002), *Pyropia columbina* (Contreras-Porcia et al. 2010), *Chondrus crispus* and *Mastocarpus stellatus* (Stackhouse) Guiry (Mathieson and Burns 1971). The increase of photosynthetic rate upon short-term desiccation, followed by a drop during long term desiccation, is similar to the photosynthetic changes during light stress. Under light stress, a photoprotection phase occurs first with non-photochemical fluorescent quenching and de-epoxidation of violaxanthin, but without any decrease in the rate of net photosynthesis, followed by an inhibition phase during which de-activation of PSII is observed as a decrease in the rate of net photosynthesis (Harker et al. 1999).

In photosynthetic organisms, including cyanobacteria and algae, rapid changes of photon capture, electron fluxes, and redox potentials during photosynthesis cause reactive oxygen

species (ROS) to be released (Latifi et al. 2009). Photorespiration is also a major source of photosynthesis-associated ROS (Dietz et al. 2016). To compensate for the negative effects of reactive oxygen species (ROS), the regulation of photosynthetic electron transport rates and protection through photoinhibition are crucial for desiccation tolerant organisms (Roach and Krieger-Liszka 2014, Fernandez-Marin et al. 2016). More detailed information about ROS is given below.

Chlorophyll a fluorescence is often used as an indicator of cellular stress and for evaluating photosynthetic activity under various environmental stresses (Van Heerden et al. 2003, Mishra et al. 2011). Chlorophyll a is a ubiquitous photosynthetic pigment that functions in photosystems, as well as a light-harvesting antenna. In all algae, chlorophyll is accompanied by supplementary pigments, which may absorb a portion of the incident light and play other important functions. Phycobiliproteins, which are a family of light-harvesting pigment protein complexes, are widely found in the chloroplasts of red algae and cyanobacteria (MacColl 1998). These proteins contain covalently bound phycocyanin and phycoerythrin. Phycobiliproteins absorb light at 450 nm to 650 nm, where chlorophyll a absorbs poorly, and transfer energy to Photosystem II (Redlinger and Gantt 1982). In the red alga *Pyropia columbina*, the levels of phycocyanin and phycoerythrin both increased during desiccation when compared with fully hydrated samples (Contreras-Porcia et al. 2010).

Red algae do not appear to have a xanthophyll cycle, which is involved in eliminating excess light energy in several different types of algae and vascular plants (Li et al. 2009). Instead, carotenoids (β-carotene) function to dissipate excess light absorbed by the photosynthetic pigments in red algae (Öquist and Fork 1982). Carotenoids are not directly involved in photochemical reactions, so their main function is probably to protect photosynthetic

machinery from oxidative damage by acting as sunscreen pigments and antioxidants (Wada et al. 2013). In Photosystem II, carotenoids play a central role in the deactivation of triplet chlorophyll and singlet oxygen, and the reduction of ROS formation (Jahns and Holzwarth 2012). High concentrations of β-carotene and other carotenoid pigments were found to accumulate in response to various abiotic stresses, especially osmotic stress in the unicellular green alga *Dunaliella salina* (Cowan et al. 1992, Pick 1998, Ramos et al. 2011). The amounts of carotenoids were also demonstrated to increase by UV-A exposure in the aquatic cyanobacterium *Oscillatoria* sp. (Wachi et al. 1995).

Accumulation of organic osmolytes

Algae can synthesize and accumulate various substances that increase the cellular osmotic value and cause a negative osmotic potential (Bisson and Kirst 1995). Organic osmolytes can act as compatible solutes and protect cells in response to various stresses, including desiccation. Compatible solutes are highly soluble and can accumulate to high concentration in cells without interfering with cellular functions (Bisson and Kirst 1995, Yancey 2005). The accumulation of compatible solutes functions in osmotic balance, and acting as thermostabilizing chemical chaperones that help to increase the stability of proteins (Roberts 2005), and as antioxidants (Amor et al. 2005). The most common types of compatible solutes include polyols, sugars, amino acids and their respective derivatives, betaines, ectoines and occasionally some peptides (Galinski 1995). There seems to be some functional specialization of the different classes of compatible solutes, which may contribute to their preferential occurrence in organisms that settle in different niches and are tolerant to adverse environmental stresses (Erdmann and Hagemann 2001). Polyols (Reed 1990), trehalose (Karsten et al. 2005), digeneaside (Karsten et al. 2005) and heterosides (Kremer and Vogl 1975, Bondu et al. 2009)

have been observed in red algae and have been hypothesized to function as compatible solutes for these organisms. Even though organic osmolytes were shown to accumulate in algae during osmotic, desiccation and freezing stress, the detected concentration was not high enough to fully explain the alga's stress tolerance (Hawes 1990, Kaplan et al. 2012).

Trehalose

The disaccharide trehalose is a non-reducing sugar, which accumulates in response to abiotic stresses, such as heat, cold or osmotic stress. Trehalose depresses the gel-to-liquid crystalline phase transition temperature in the dehydrated lipid to a temperature at or near that of the hydrated lipid, and trehalose forms a carbohydrate glass with a relatively high glass transition temperature, leading to inhibition of fusion between vesicles (Jain and Roy 2009). Trehalose also plays a major role in osmotic adjustment by stabilizing membranes and replacing water (Jain and Roy 2009). Protein stabilizer properties of trehalose are much better than many other compatible solutes because of their ability to alter the water environment surrounding a protein and stabilize the protein in its native conformation (Kaushik and Bhat 2003, Magazù et al. 2005). High levels of trehalose have been found in some desiccation-tolerant plants, such as the moss *Selaginella lepidophylla* (Hook. & Grev.) Spring and the resurrection plant *Myrothammus flabellifolius* Welw. (Goddijn and Van 1999, Pampurova et al. 2014). Trehalose was found to be insufficient for desiccation tolerance as a compatible solute in other organisms, such as yeast (Ratnakumar and Tunnacliffe 2006), the green alga *Klebsormidium nitens* (Kützing) Lokhorst and the green alga *Hormidiopsis crenulata* (Kützing) Heering (Kaplan et al. 2012). In the Florideophyceae, trehalose was detected using ^{13}C-nuclear magnetic resonance spectroscopy and high-performance liquid chromatography only in several members of the Ceramiales but not in the Bangiophyceae *Porphyra* (Karsten et al. 2007).

Trehalose biosynthetic pathway(s) are present in eubacteria, archaea, plants, fungi and animals. In bacteria, there are five different biosynthetic routes, whereas there is only one in fungi, plants and animals (Avonce et al. 2006). Among five known trehalose biosynthetic pathways, the pathway that involves two enzymatic steps catalyzed by trehalose-6-phosphosphate synthase (TPS) and trehalose-6-phosphate phosphatase (TPP) is the most widely distributed and has been found in eubacteria, archaea, fungi, insects, and plants (Avonce et al. 2006). The transfer of glucose from UDP-glucose to glucose 6-phosphate is catalyzed by TPS and forms trehalose 6-phosphate (T6P) and UDP, while TPP dephosphorylates T6P to trehalose and inorganic phosphate (Elbein et al. 2003). Trehalose is hydrolyzed into glucose by trehalase. In bacteria, the TPS and TPP genes are clustered, whereas in eukaryotes these domains are fused in a single protein.

In *Pyropia yezoensis,* the presence of a putative trehalose-6-phosphsphate synthase (TPS) gene was confirmed by RACE-PCR and the sequences of this TPS gene were compared with nine other algae, which showed that the sequences were highly conserved both in nucleotide composition and in amino acid composition, and there was no intron in this putative TPS gene in any of the ten algae studied by Wang et al. (as *Porphyra yezoensis* 2010). The comparative genomic study of two closely related unicellular red algae showed that the genes important for trehalose metabolism were found in both *Galdieria sulphuraria* and *Cyanidioschyzon merolae* genomes (Barbier et al. 2005).

Floridoside and isofloridoside

The heterosides are composed of a sugar linked to a polyol; examples include glucosylglycerol (GG) and galactosylglycerols (floridoside and isofloridoside). Glucosylglycerol is mainly found in cyanobacteria, while the galactosylglycerols, floridoside and isofloridoside,

are widespread in red algae except the *Ceramiales* (Kirst 1990). Floridoside and isofloridoside are found in several *Porphyra* species, including *P. umbilicalis* (Wiencke and Läuchli 1981, Meng et al. 1987, Meng and Srivastava 1993, Karsten et al. 1993, Karsten 1999). Wiencke and Läuchli (1981) used various concentrations of artificial seawater medium to impose osmotic stress; they found that the total amount of floridoside and isofloridoside in *P. umbilicalis* increased at high hyperosmotic stress, but then decreased. Wjencke and Läuchli (1981) measured the total galactosylglycerol concentration (floridoside+ isofloridoside) under salt stress; they did not measure how floridoside alone changed during desiccation and rehydration.

According to Pade et al. (2015), the biochemical pathways for the two-step synthesis of trehalose and (iso)floridoside are very similar. In the floridoside synthesis pathway, floridoside phosphate synthase (FPS) first catalyzes the formation of floridoside phosphate from glycerol-3-phosphate (G3P) and UDP-galactose. Then floridoside phosphate is dephosphorylated to floridoside by (iso)floridoside phosphate phosphatase. It is possible that the putative trehalose-6-phosphate synthase gene found in *P. yezoensis* (Deng et al. 2004, Wang et al. 2010) and *P. umbilicalis* may not be involved in the synthesis of trehalose but rather in floridoside, as previously proposed for the red alga *Galdieria sulphuraria* (Pade et al. 2015). According to Brawley et al. (2017), the *Porphyra umbilicalis* genome has been previously screened for candidate genes involved in trehalose and (iso)floridoside biosynthesis, but only two of the possible TPS/FPS genes have been identified; these were suggested to be involved in floridoside and isofloridoside synthesis.

Reactive Oxygen Species

The generation of reactive oxygen species (ROS) is triggered by different kinds of environmental stresses, such as high or low temperatures, high light (UV-B radiation), drought,

desiccation, salinity, low CO_2 availability, nutrient deficiency and pathogen attack (Collén and Davison 1999). ROS are produced directly by excitation of O_2 and subsequent formation of singlet oxygen ($^1O_2$), or by the transfer of one, two, or three electrons to O_2, which results in the formation of superoxide radicals (O_2^-), hydrogen peroxide (H_2O_2), or hydroxyl radicals ($HO^{\cdot}$), respectively (Baker and Orlandi 1995, Tripathy and Oelmüller 2012). Living organisms in the oxidizing environment constantly produce ROS in cells during metabolic processes, such as photosynthesis and respiration.

Under normal conditions, antioxidant systems of cells minimize the perturbations caused by ROS, maintaining the balance between oxidation and reduction. When the production of ROS exceeds the buffering capacity, ROS can be harmful and this is associated with oxidative stress, which can cause damage to lipids, proteins, polysaccharides and DNA, and ultimately lead to cell death (Deighton et al. 1999, Rustérucci et al. 1999, Thoma et al. 2003). The prevention of oxidative stress is an essential process in all aerobic organisms. Photosynthetic organisms show various responses in order to avoid ROS over-production and damages of biomolecules. Organisms protect themselves by the following mechanisms: shading light through sunscreen pigments, controlling photosynthetic electron transport by favoring photosystem I or photosystem II, and/or activation of the antioxidant systems quenching the ROS produced (Latifi et al. 2009). Plants and other living organisms have evolved a host of water-soluble and lipid-soluble antioxidants and anti-oxidative enzymes to dissipate ROS (Foyer et al. 1997, Noctor and Foyer 1998, Smirnoff 2000, Tripathy and Oelmüller 2012). Antioxidant enzymes include superoxide dismutase (SOD), catalase (CAT), ascorbate peroxidase (APX), guaiacol peroxidase (GPX), glutathione reductase (GR), monodehydroascorbate reductase (MDHAR), and dehydroascorbate reductase (DHAR). The non-enzymatic antioxidants include ascorbic acid

(AA), reduced glutathione (GSH), α-tocopherol, carotenoids, flavonoids, and the osmolyte proline (Gill and Tuteja 2010, Miller et al. 2010).

Some ROS may function as important signaling molecules that mediate many different processes in cells. These include the regulation of metabolic pathways and physiological processes, the activation of acclimation responses to abiotic stress, the activation of defense responses against different pathogens and pests, the activation of different developmental programs, and the coordination of systemic plant responses to different environmental stimuli (Inupakutika et al. 2016). Recent studies have linked ROS signaling with different plant hormones, such as abscisic acid, salicylic acid, and nitric oxide, suggesting that ROS can also act as a bridge between different plant hormones, redox biology, and changes in cellular calcium levels and/or protein phosphorylation events (Inupakutika et al. 2016).

According to Contreras-Porcia et al. (2010), an overproduction of reactive oxygen species (ROS) was induced by desiccation and two peaks of H_2O_2 production were detected at 1h and 3h of desiccation in the red alga *Pyropia columbina* (as *Porphyra columbina)*. The ROS level returned to base line after rehydration. The activity of antioxidant enzymes was increased in response to desiccation. The activity of antioxidant enzymes, except catalase (CAT), diminished to near basal levels during rehydration. Similar results have been found in the intertidal red alga *Hypnea musciformis* (Wulfen) J.V.Lamouroux in response to desiccation. Increasing levels of lipid peroxidation, H_2O_2, antioxidants (such as catalase), reduced glutathione and acetylsalicylic acid, were observed during peak summer time with elevated temperature, UV and desiccation stresses in *H. musciformis*. Algae having higher intertidal distribution displayed greater antioxidant enzymatic activity of catalase and ascorbate peroxidase after simulated emersion stress (Flores-Molina et al. 2014). However, Liu (2009) suggested that ROS defense is

not key to the difference between the desiccation tolerant *P. umbilicalis* and the desiccation sensitive *Pyropia yezoensis* because: 1) desiccation in the dark did not alleviate the ROS damage in *P. yezoensis*; 2) there was a small decrease in ascorbate content in *P. umbilicalis* but not *P. yezoensis* after desiccation; and 3) neither *P. umbilicalis* nor *P. yezoensis* showed increasing levels of membrane peroxidation after desiccation. Such findings were consistent with these of Sampath-Wiley et al. (2008). Sampath-Wiley et al. (2008) suggested that even emersion leads to photoinhibitory damage and ROS generation at PSII. However, the antioxidant levels during emersion in *P. umbilicalis* only increased during summer months when irradiance levels and temperatures were high, while antioxidant levels were similar during winter submersion and emersion.

Plant hormones

Plant hormones (phytohormones) are molecules produced in very low concentration. They play important roles in mediating plant growth as well as signaling environmental changes, initiating stress responses, and serving as indicator molecules in the regulation of almost all phases of plant development (Kojima et al. 2009). In green alga *Klebsormidium crenulatum,* almost the complete pathway for cytokinin signaling, abscisic acid (ABA) signaling, and ethylene response were found in its genome (Holzinger and Becker 2015). In the *K. flaccidum* genome, genes involved in the pathways of auxin, ABA, cytokinin, salicylic acid, and jasmonic acid were found (Hori et al. 2014). In *K. crenulatum,* expression of genes controlling ABA signaling, PP2C and SnRK2, significantly increased in response to desiccation (Holzinger and Becker 2015). Endogenous phytohormone profiles of auxins, abscisic acid, and gibberellins were identified qualitatively and quantitatively in the red alga *Sarconema filiformae*, the green alga

Ulva rigida C.Agardh, and brackish water green alga *Ulva lactuca* Linneaus using GC/MS techniques (Shoubaky and Salem 2016).

Among the five characteristic phytohormones, ABA is the most commonly known "stress hormone" that responds to a variety of biotic and abiotic stress (Zhang 2014). ABA also helps control many development and growth characteristics of plants. ABA is synthesized from carotenoids and it is found in more than 60 species of algae (Shoubaky and Salem 2016). Pryce (1972) first hypothesized that algae and liverworts were distinct from higher plants and found that algae and liverworts did not contain ABA. Later on, ABA was found to be present in various groups of algae, including *Porphyra* (Hartung 2010, Hirsch et al. 1989, Tominaga et al. 1993). However, the ABA content in algal cells is significantly lower than that of unstressed terrestrial plants (Hartung 2010). Takezawa et al. (2011) suggested that ABA may not play a signaling role in response to water stress in algae due to the lack of a core component (group A protein phosphatase type 2C) essential in the induction of ABA signaling and the fact that algae have low sensitivity to exogenous ABA (Komatsu et al. 2013). Even after the discovery of important ABA signaling components, protein phosphatase 2C (PP2C) and type 2 SNF1-related protein kinases (SnRK2) genes in green alga *K. crenulatum* (Holzinger and Becker 2015), it is still unclear if ABA functions as a signaling molecule in algae.

Ethylene response appears to be highly conserved as a plant hormone for the past 450 million years (Ju et al. 2015). The effects of ethylene in plants include ripening of fruits, inhibition of stem and root elongation, promotion of seed germination and flowering, senescence of leaves and flowers, and sex determination (Bleecker and Kende 2000). The ethylene synthesis pathway via 1-aminocylopropane-1-acrylic acid (ACC) and ACC oxidase was confirmed in the marine green macroalga *Ulva intestinalis* Linnaeus (Plettner et al. 2005). Ethylene production

was observed in *U. intestinalis* and its level increased substantially in response to high light. In addition, ethylene was shown to reduce chlorophyll levels by 30% in *U. intestinalis*. Timme and Delwiche (2010) also confirmed the presence of ethylene biosynthesis and signaling pathway genes in the green algal charophytes. In red algae, ethylene was detected and shown to play an important role in the regulation of gamete formation and protection against stress-induced damage in *Pyropia yezoensis* (Uji et al. 2016). The genomic data for *Porphyra umbilicalis* also supports the presence of ethylene-mediated regulation in *Porphyra*, because the *Porphyra* genome contains genes encoding several proteins involved in ethylene biosynthesis (Brawley et al. 2017).

Transcriptome analysis under desiccation

Transcriptome sequencing can be performed without prior knowledge of the genome of interest and allows various applications such as: 'de novo' reconstruction of the transcriptome, identification of the expression of isoforms and unknown transcripts, evaluation of methylation patterns and so on (Costa-Silva et al. 2017). With the rapid development of next generation sequencing (NGS) technologies in recent years, RNA-seq provides a cost-effective and high-throughput approach with a wide variety of applications, which can assess the whole transcriptomes, including measuring gene expression levels under different treatments.

Some studies have used molecular techniques to identify the functional genes and the molecular pathways behind desiccation tolerance in algae. The studies suggest that there is great difference between the genes that were differently expressed in algae and in plants. The highest differentially expressed genes (DEG) in response to desiccation in *Klebsormidium crenulatum* do not show any similarity to know proteins (Holzinger et al. 2014). Thus, the transcriptome data will be useful to understand the desiccation tolerance mechanisms in algae.

In the green alga *Klebsormidium* (Holzinger et al. 2014), the genes up-regulated during desiccation were related to photosynthesis, energy production and ROS metabolism. Also, genes coding for late embryogenesis abundant proteins (LEA) or osmolytes were up-regulated. In contrast, transcripts related to cell division, DNA replication, cofactor biosynthesis, and amino acid biosynthesis were strictly down-regulated as a consequence of desiccation stress. In the green alga *Zygnema circumcarinatum* Czurda*,* similar results were found by transcriptome study. Photosynthesis was repressed while ROS scavenging, DNA repair and expression of chaperones and aquaporins were induced in response to desiccation (Rippin et al. 2017). In red algae, desiccation response genes were identified in *Pyropia tenera* by transcriptomic study (Im et al. 2017). The study showed that DEGs under desiccation are related to functions such as carbohydrate metabolism, membrane perturbation and compatible solutes. The study also suggested that *Pyropia* may possess novel genes that differ from green plants. Genome-wide expression profiles of *Pyropia haitanensis* (Kjellman) N.Kikuchi, M.Miyata, M.S.Hwang & H.G.Choi have been studied in response to dehydration/rehydration stress using deep sequencing technology (Wang et al. 2015).

Presently, there is no consensus about the most appropriate protocol for identifying differentially expressed genes from RNA-Seq data. The appropriate number of biological replicates is a trade-off between cost and precision. Earlier studies in red algae were limited to two replications per treatment or less, which is now generally considered too few replicates to achieve accurate statistical power to discover DEGs. It is suggested that at least six biological replicates should be used for RNA-seq experiments to identify DEGs (Schurch et al. 2016).

Goals

The thesis focused on two topics: 1) the genetic diversity and population structure among populations of *P. umbilicalis* from the Gulf of Maine; and 2) the desiccation tolerance mechanisms in *P. umbilicalis.*

A computational pipeline was developed to find single nucleotide polymorphism (SNP) markers from transcriptome data from *P. umbilicalis.* A portion of the discovered SNP markers were validated by Sanger sequencing. Five of the validated polymorphic SNP markers were applied in a pilot study of genetic diversity and population structure of seven *P. umbilicalis* populations within the Gulf of Maine (Chapter 1).

Compatible solutes play important roles in the desiccation tolerance mechanisms. The amounts of (iso)floridoside and trehalose, which are common compatible solutes in plants and in algae, were measured in response to desiccation and rehydration treatments; these conditions are similar to the tidal cycles *P. umbilicalis* experiences in its natural habitats. Genes involved in the synthesis of (iso)floridoside and trehalose were screened from the recently completed *Porphyra* genome (Brawley et al. 2017). Phylogenetic and conserved domain analyses of the genes, which may be involved in the synthesis of (iso)floridoside and trehalose in *P. umbilicalis,* were conducted in order to distinguish the functions of genes that were involved in trehalose, floridoside, and isofloridoside synthesis. Quantitative real-time PCR (qPCR) was used to determine the expression profile of four *TPS/FPS* genes at different times before, during and after desiccation in *P. umbilicalis* (Chapter 2).

In order to understand the mechanisms of desiccation tolerance in *P. umbilicalis* at the molecular level, a comparative transcriptomic analysis was performed using *P. umbilicalis* under fresh, dehydrated, desiccated and rehydrated conditions. Differences in expression patterns under different conditions were used to understand: 1) the importance of desiccation protection and

rehydration recovery in desiccation tolerance mechanisms in *P. umbilicalis*; 2) the functions of the DEGs under different conditions (Chapter 3).

CHAPTER I

A pilot study of genetic structure of *Porphyra umbilicalis* Kützing in the Gulf of Maine using SNP markers from RNA-Seq

Abstract

The red alga *Porphyra umbilicalis* Kützing has a broad distribution within the North Atlantic. In the Northeast Atlantic, it is dioecious and reproduces both sexually and asexually, while in the Northwest Atlantic only asexual reproduction has been observed. In this study, transcriptomes were mined to identify putative single nucleotide polymorphisms (SNP) markers. A computational pipeline was developed that accounts for the specific characteristics of transcriptome dataset, filtered against the available red alga *Chondrus* genome and against a *P. umbilicalis* EST library to eliminate microbial contamination. Five hundred forty-nine putative SNPs were detected within a single population (Schoodic Point, ME). Five of the validated SNP markers were applied in a pilot study of genetic diversity and population structure of seven *P. umbilicalis* populations within the Gulf of Maine. Novel genotypes were found in the open coastal populations at Reid State Park, Schoodic Point and the estuarine tidal rapid population at Wiscasset. Our study represents the first attempt to develop suitable bioinformatic pipeline for RNA-seq to detect SNP markers for the red alga *Porphyra umbilicalis* and to successfully use these SNP markers for population studies.

Introduction

The identification of genomic variation helps to clarify the relationship between genotype and phenotype. Single nucleotide polymorphisms (SNPs) are single base differences between DNA sequences of individuals or strains. SNP markers represent the most common type of variation across a genome (Kwok 2001). Studies of SNP markers facilitate marker-aided selection (MAS), detect alleles associated with disease, analyze population history and are used to produce genetic maps (Goya et al. 2010, Cardon and Bell 2001, Flint-Garcia et al. 2003, Nakitandwe et al. 2007). Before the advent of next generation sequencing, SNP markers were identified by Sanger Sequencing and common SNPs were then incorporated into commercial probe-based SNP array platforms (Gupta et al. 2008). Commercial SNP arrays are restricted to model organisms or to biological species of medical, agronomic or commercial importance, which have large sequence databases. The rapid development and reduced expense of next-generation sequencing technologies have provided unprecedented opportunities for researchers to find SNP markers, especially for non-model organisms.

Transcriptome sequencing techniques (i.e., RNA-seq) can detect variants in the coding regions of the genome. RNA-seq provides large amounts of data with reduced costs (Cloonan et al. 2008, Wilhelm et al. 2008). SNPs in noncoding regions may alter transcript levels by disrupting functional *cis*-regulatory elements, while those in coding regions of the genome may be silent or responsible for altered forms of proteins (Chepelev et al. 2009). RNA-seq is commonly employed to quantify gene expression levels under different conditions, and to detect alternative splicing, allele-specific expression, gene fusions, and RNA editing (Wang et al. 2009,

Rapaport et al. 2013, Eswaran et al. 2013, Crowley et al. 2015, Piskol et al. 2013, Conesa et al. 2016). Intrinsic complexities in transcriptomes (e.g., alternative splicing) lead to some challenges when using RNA-seq data to identify genomic variants in computational analysis steps. Several methods have been developed to call variants from RNA-Seq data with the availability of complete genome information (Piskol et al. 2013, Deelen et al. 2015, Quinn et al. 2013). Studies suggested that imposing strong variant filtering criteria, having sufficient coverage, using relevant tissue, imposing suitable quality control screens and having additional whole-exome sequencing (Seo et al. 2012, Cirulli et al. 2010) can increase the accuracy of variant identification using transcriptome data. Recently, RNA-seq has been employed to systematically identify variants in transcribed regions in different species, mostly in humans (Quinn et al. 2013, Cirulli et al. 2010, Chepelev et al. 2009, Kim et al. 2014) and plants (Shearman et al. 2015, Paritosh et al. 2013).

The marine red alga *Porphyra umbilicalis* Kützing is found from the Northeast and Northwest Atlantic, while its global distribution requires further review (Brodie et al. 2008). Sutherland et al. (2011) showed evidence from molecular data that the name of *Porphyra umbilicalis* had been applied to more than one species. *Porphyra umbilicalis* is an important food source with high level of proteins, free amino acids and a high ratio of ω3: ω6 fatty acids (Mouritsen 2013). *Porphyra umbilicalis* could be used as a biofilter in integrated multi-trophic aquaculture (IMTA) and as a partial replacement for fishmeal (Day et al. 2009; Walker et al. 2009*)*. *Porphyra umbilicalis* that is native to Northwest Atlantic could be a good target to develop domesticated strains for aquaculture and for components of IMTA (Blouin et al. 2007). Amplified Fragment Length Polymorphism (Blouin and Brawley 2012) and Simple Sequence Repeats (SSR or microsatellite) markers (Eriksen et al. 2016) have been applied to analyze the

genetic diversity and variation of the Northwest Atlantic populations of *P. umbilicalis*. However, the limited number of SSRs markers used hinders the direct comparison of the microsatellite and AFLP assessment of *P. umbilicalis* genetic variation. Furthermore, there are large microbial communities associated with *P. umbilicalis*, living on the surface and inside the cell wall of the alga (Miranda et al. 2013). Potential bacterial DNA contamination of the *Porphyra* DNA pool may have generated more variable AFLP profiles. It is important to have more genetic markers that can be firmly tied to the *Porphyra* genome to analyze the genetic diversity and structure of Northwest Atlantic populations, and to aid in strain selection for integrated multi-trophic aquaculture.

In the Northeast Atlantic, *P. umbilicalis* can reproduce both sexually and asexually (Brodie and Irvine 2003); however, for NW Atlantic populations only asexual reproduction has been observed (Blouin et al. 2007, Gantt et al. 2010, Bird and McLachlan 1992). Northwest Atlantic *P. umbilicalis* populations are found over a wide latitudinal range, and occupy both rocky, open coastal habitats and estuarine tidal rapids (Eriksen et al. 2016). How these asexual populations adapt to different habitats is still unclear. According to Ingolfsson (1992), the NW Atlantic rocky coast species were extirpated by the previous glacial period, and the current populations are assumed to be the descendants of the Northeast Atlantic species that were introduced from glacial refugia via post-glacial trans-Atlantic currents. Assessing the genetic diversity and variation of *P. umbilicalis* in the NW Atlantic could assist with understanding how *P. umbilicalis* colonized the Gulf of Maine after the last glacial maximum, and how asexual populations survived in different environments, such as open coastal and estuarine tidal rapid habitats.

In this study, a computational pipeline was developed to identify SNP markers from RNA-seq data in *P. umbilicalis*. The pipeline accounts for the specific characteristics of these RNA-Seq experiments, the biological characteristics of *P. umbilicalis* (microbial contamination), and the absence of a genome assembly of *P. umbilicalis* at the time this study was carried out. Twenty-five of these SNP markers were validated and five were then used in a pilot study to examine genetic diversity and population structure of *P. umbilicalis* within the Gulf of Maine.

Material and Methods

Species identification, gametophyte culture, RNA isolation and sequencing

Four "individuals" (each defined as thalli connected to a single holdfast) were collected randomly along the rocky shore during low tide at Schoodic Point, ME (44°20'11.3"N 68°03'23.3"W). Because morphological identification of *P. umbilicalis* is error prone (Klein et al. 2003), the *rbc*L-*rbc*S genes and intergenic spacer were amplified using primers described in Teasdale et al. (2002) and sequenced for those individuals to confirm species identity. Individuals confirmed as *P. umbilicalis* were put into culture (see below). For population diversity studies, species identification was confirmed by amplification of *rbc*L-*rbcS* region and restriction fragment length polymorphism comparison (Teasdale et al. 2002).

Neutral spores release was induced from the four wild *P. umbilicalis* blades from Schoodic Point, ME. Neutral spores were isolated and cultured in the lab-controlled environment according to Redmond et al. (2014). Blade materials from each culture were pooled, and then each pool was subjected to one of five different treatment conditions (Table 1). The total RNA was extracted and assessed for quality and quantity using a NanoDrop 2000c spectrophotometer (ThermoFisher Scientific) and an Agilent 2100 Bioanalyzer. The cDNA libraries were prepared

by Illumina Truseq RNA Preparation Kit (Illumina, San Diego, USA) and then sequenced on Illumina HiSeq 2000 at the Hubbard Center for Genome Studies in University of New Hampshire. The raw reads are deposited in the Sequence Read Archive (SRA) BioProject PRJNA496324 (https://www.ncbi.nlm.nih.gov/Traces/study/?acc=PRJNA496324).

Reference transcriptomes construction

Five libraries of transcriptome sequences from *P. umbilicalis* were generated under five environmental treatments; these transcriptomes were used to build the reference transcriptome assembly. Reads were first error corrected by BLESS (Heo et al. 2014) with kmerlength 25 and then trimmed with Trimmomatic (Bolger et al. 2014) at Phred score of 2 (MacManes 2014) to get rid of low quality reads and adapters. Subsequently, Trinity (Grabherr et al. 2011) was applied for *de novo* assembly with digital normalization. Extensive bacterial contamination was found on the surface and inside the cell walls of *P. umbilicalis* (Miranda et al. 2013). Therefore the raw *de novo* assembly was screened by BLAST analysis against the genome of the red alga *Chondrus crispus* Stackhouse (Collén et al. 2013), as well as the *P. umbilicalis* EST reference (Chan et al. 2012a), to eliminate potential bacterial contamination. Only contigs that were similar to either the *C. crispus* genome or the *P. umbilicalis* EST reference with E-values higher than $1e^{-10}$ were retained, resulting in a contamination-free, partial transcriptome reference for *P. umbilicalis*. The number of "genes", the number of "transcripts", the GC content, and the N50 value were used to checke the quality of the contamination-free reference. The completeness of the contamination-free reference was checked using the Core Eukaryotic Genes Mapping Approach (CEGMA, Parra et al. 2007).

Single nucleotide polymorphism markers detection

Reads from all five libraries were first trimmed at a Phred score of 20 to eliminate low quality reads and then mapped to the reference transcriptome assembly by BWA (Li and Durbin 2010). Samtools (Li et al. 2009) was used to filter alignments with mapping quality lower than 30. Picard (The Broad Institute; http://picard.sourceforge.net) was used to mark duplicates and sort. SNP calling was performed using the Genome Analysis Toolkit (McKenna et al. 2010) with Split'N'Trim. Indels and large structural variants were not analyzed. The raw putative SNPs were further filtered out based on the following criteria: (1) sequence depth at the SNP position ≥ 20; (2) FisherStrand (FS) ≤ 60; (3) RMSMappingQuality (MQ) ≥ 40; (4) MappingQualityRankSumTest (MQRankSum) ≥ -12.5; (5) ReadPosRankSumTest (ReadPosRankSum) ≥ -8.0; (6) three SNPs were within 35 bp of each other. After filtering, SNPs that were common to all five libraries were considered *P. umbilicalis* putative population SNPs and used for further validation. The genes containing SNPs that were unique to an individual treatment were annotated using Blast2GO (Conesa et al. 2005). All the scripts used can be found at https://github.com/yov2/SNPfromRNA_Porphyra_umbilicalis.git.

SNP markers validation

Twenty-five putative SNPs identified by computational pipeline were randomly chosen for validation using five individuals collected from five sites: Fort Stark, NH; Dover Point, NH; Lubec, ME; Reid State Park, ME; Nubble Light, ME. The adjacent sequence around target SNP was retrieved from RNA-seq data to design primers. Primer pairs targeting each putative SNP were designed with Primer 3 software (http://primer3.ut.ee/). Polymerase Chain Reaction (PCR) conditions were optimized for each primer pair and then used to amplify various DNA templates for the targeted SNP regions. Amplification of each target region was performed in 25 μL

reaction volumes containing about 25–125 ng genomic DNA, 1x Q5 One *Taq* Standard Reaction Buffer (New England BioLabs, Ipswich, MA, USA), 200 µM dNTPs, 0.2 µM of each forward and reverse primer (Table 2), and 0.75 U One *Taq* Hot Start DNA Polymerase (New England BioLabs). PCR started with an initial denaturation step of 94°C for 3 min and was followed by 35 cycles of 30 s at 95°C, 1 min at a primer-specific annealing temperature (Table 2), and 1 min extension at 72°C. The amplification ended with a final extension at 72°C for 5 min. The amplicon sizes were the same as predicted from the transcriptome sequence. The amplicons were purified using both QIAuick Gel extraction kit and ExoSAP-IT (Affymetrix.com) and then sent for Sanger Sequencing at GENEWIZ® Company, New Jersey, USA. The Sanger sequencing results were aligned using MAFFT (http://mafft.cbrc.jp/alignment/server/) to detect SNPs. Trace files were utilized to further confirm SNPs found by MAFFT.

Genetic diversity and population structure

A total of five SNP markers (2, 8, 13, 18 and 22) that were validated by Sanger sequencing were used in the pilot population study. Genes containing these SNPs were annotated for their functions using Blast2GO (Conesa et al. 2005). A total of 37 individuals from seven sampling sites (Figure 4) were assayed for these five SNP markers. Among these seven populations, there were five open coastal populations and two estuarine tidal-rapid populations. The two estuarine tidal-rapid populations, Dover Point NH and Wiscasset ME, are geographically close to open coastal populations (Fort Stark NH and Reid State Park ME), respectively. PCR amplification conditions were the same as described above.

The major allele frequency, gene diversity, polymorphism information content (PIC) and Nei's genetic distances for each population were calculated using Power Marker 3.25 (Liu and Muse 2005). The SNP data was coded as follows: A=1, C=2, G=3, T=4 and missing data was

coded as 0 as suggested in GenAlEx V6.5 user manual (Peakall and Smouse 2006, 2012). Analysis of molecular variance (AMOVA) and Principle Coordinate Analysis (PCoA) were performed in GenAlEx v. 6.41 (Peakall and Smouse 2006, 2012). In addition, a Mantel test (Mantel, 1967), with 9999 permutations, was conducted using the program GenALEx 6.5 (Peakall and Smouse 2006, 2012) for correlation between Nei's genetic distance (Nei 1972) and geographic distance. Because neutral spores travel along currents, geographic distance was calculated according to Eriksen (2014).

Results

Reference Transcriptome

Comparisons between the raw Trinity transcriptome and the transcriptome reference corrected for contamination are summarized in Table 3. There were 182,905 contigs in the raw reference directly from Trinity output, with GC content of 56.47% and N50 of 512 bp. There were many fewer contigs in the contamination-free transcriptome reference after clean-up step (42,802 reads), with higher N50 length (1090bp) and higher GC contents (62.28%). However, according to CEGMA result (Table 3), the transcriptome reference without contamination was 4% less complete than the raw Trinity transcriptome reference.

SNP calling and functional annotation of library specific SNP-containing genes

With stringent filtration, about 90% of the raw SNPs were filtered out. There were 549 putative SNPs in common to all five libraries (Figure 1). The functional annotation containing unique putative SNPs in each library is showed in Figure 2. There were fewer SNPs unique to the stress stimulus in library C (air dried, frozen and then cultured for two weeks) compared to other libraries.

SNP validation

Primer pairs were designed for 25 gene transcripts containing putative SNPs. Target DNAs were amplified successfully for 13 of the 25 primer pairs. The primer pairs were used to amplify DNA from five individuals, each from one of the five different populations. The resulting amplicons were sent for Sanger Sequencing at GENEWIZ®. Products of primer pair 4 did not contain the designed SNP so this marker was dropped from further analysis. Amplicons from primer pairs 3 and 17 exhibited double peaks in the trace file from some individuals, even when sequenced from both ends. Among the remaining 10 primers pairs, seven were polymorphic based on screening five individuals (SNP 2, 8, 13, 14, 15, 18 and 22). Two of the primer pairs (SNP 10 and SNP 16) produced monomorphic amplicons for the five *P. umbilicalis* individuals tested. SNP 19 was monomorphic in the predicted SNP position but was polymorphic at another position. The true SNP detection rate was likely higher than 70% since only five individuals were used to validate each putative SNP.

Pilot population study

Five SNP markers from the seven validated polymorphic SNP loci were used for the population genetic study. The functions of these five amplicons were 6-phosphogluconate dehydrogenase decarboxylating (SNP 8), mRNA export factor/elongation factor (SNP 13), translocation protein 3Ec63 homolog (SNP 18), and unknown (SNP 2 and 22). There were more SNPs in the sequenced regions besides those inferred by bioinformatics of RNA-seq libraries and an additional eight SNPs were identified in comparing two or more amplicons from the same primers. Thus, 13 polymorphic SNPs were characterized in this study and the detailed haplotype information of these 13 SNPs is given in Supplemental Table 1. The gene diversity and Polymorphic Information Content (PIC) value for each population are shown in Table 4. The

gene diversity ranged from 0 to 0.16 and the PIC value ranged from 0 to 0.131. Dover Point, Nubble Light, Reid Park and Wiscasset had the lowest genetic diversity and PIC values. The highest genetic diversity occurred from populations collected from Lubec and Schoodic Point.

Population structure

Genetic distance and genetic differentiation across the sampling sites are summarized in Table 5. The genetic distances were highest between Reid State Park and the rest of the populations, ranging from 0.171 to 0.308. Based on result from Mantel test, there was no evidence of isolation by distance between geographic distance and genetic distance (R_{xy} = -0.344, P=0.613). Based on genetic distance, Principle Coordinate Analysis (PCoA) clustered three open coast populations (Fort Stark, Lubec and Schoodic Point) into a central group, and Nubble Light and Dover Point into another group (Figure 4). Reid State Park and Wiscasset were both on the right side of the PCoA but separated by PCoA2. AMOVA showed that there was more variation among populations (59%) than within populations (41%) indicating some level of colonality within populations.

Figure 4 shows the frequency of genotypes in each population. Besides in the Schoodic Point population (Genotype 4 - Genotype 9), two other unique genotypes were found in Wiscasset (Genotype 3) and Reid State Park (Genotype 1). A common genotype (G2) was found in all populations except Reid State Park NH and Wiscasset ME. Schoodic Point ME had the highest genotypic diversity, followed by Lubec ME and Fort Stark NH.

Discussion

SNP discovery and validation

Several bioinformatic pipelines have been developed to detect SNPs using genome information. Pabinger et al. (2014) previously reviewed these methods However, most existing pipelines were designed with the availability of a good reference genome and thus are species-specific, such as SOAPsnp (Li et al. 2009), SNPdetector (Zhang et al. 2005) and SNPiR (Piskol et al. 2013). Several other pipelines have been successfully applied to both model and non-model species using only transcriptome data without a genome reference (Romiguier et al. 2014, Van Belleghem et al. 2012, Piskol et al. 2013, Li and Godzik 2006, Lopez-Maestre et al. 2016). In green algae, SNP markers have been identified using only transcriptome data (Li et al. 2014). The present study is the first attempt to detect SNP markers from red algal species using only transcriptome data.

The computational pipeline we used accounted for transcriptome's intrinsic complexity and characteristics of *P. umbilicalis,* such as the lack of a reference genome at the time the research was conducted, and the extensive microbial contamination of *P. umbilicalis* tissues (Miranda et al. 2013). Filtering the transcriptome assemblies against the *Porphyra* EST database (Chan et al. 2012a) and *C. crispus* genome (Collén et al. 2013) eliminated microbial contamination and ensured the SNP markers came from *Porphyra.* After filtering, the contamination-free transcriptome reference assembly was estimated to be slightly less complete (83 versus 87%) than the original Trinity reference transcriptome (Table 3).

We found a total of 549 putative SNPs in common to five *Porphyra* RNA-Seq libraries. SNP validation results showed that at least 70% of the SNPs identified by the bioinformatic pipeline were true SNPs. The true detection rate was likely higher than 70% since only five individuals from different algal populations were screened to validate each amplicon. Increasing

the number of individuals or increasing the number of sampling sites would likely increase the true SNP detection rate. The 549 SNPs are good candidate markers for further genetic diversity and population structure analysis, especially to study the evolutionary history of how mutations accumulate in asexual *Porphyra* populations in the NW Atlantic.

Library C went through air-drying and freezing stress conditions, which might be expected to cause more stress-related genes to be expressed. However, there were less unique SNPs detected responding to stress-related stimuli in library C compared to other libraries (Figure 2). Based on annotation results, we suspect that some library-specific SNPs were identified due to differences in transcript coverage in each library rather than due to differential gene expression. We showed that RNA-seq data is suitable to find SNPs without bias towards specific treatments. We also identified two SNPs (3 and 17) that had high background noise in the sequencing files. This high background noise could be the result of polymorphisms from recent gene duplications; this is a drawback to not having a reference genome with information about underlying gene family structure.

Population genetics analysis

Thirteen polymorphic SNPs were used to investigate in a pilot study of the genetic diversity and population structure of *P. umbilicalis* from seven populations within the Gulf of Maine. A total of 11 genotypes were found. Genotype 2 was present in a wide range of environmental conditions from the northern border of Maine to New Hampshire, as well as the estuarine environment at Dover Point. A "general purpose genotype" (Baker 1965) that can confer broad environmental tolerance will rapidly increase its frequency in a population without the selection of a locally adapted variant. Genotype 2 may represent a "general purpose genotype". However, this "general purpose genotype" may not be able to deal with all changing

environments (Selander and Hudson 1976) as Genotype 2 was not detected in Wiscasset ME and Reid State Park ME, which have unique environmental characteristics.

Schoodic Point had the highest genotypic diversity, with each individual examined having a unique genotype; this population also had the highest genetic diversity. Such results suggested there may be population ascertainment bias as the SNP markers usedwere originally discovered using cultures established from several individuals from Schoodic Point. Dover Point, Nubble Light, Reid State Park and Wiscasset populations had no genetic diversity within populations with this set of SNPs. Fort Stark and Lubec populations had intermediate levels of genetic diversity. The genetic diversity of asexual populations depends on the number of possible mutations that occur over time and the proportion of these mutations that persist through time (Good et al. 2012). The low (or absent) allelic diversity in Dover Point, Nubble Light, Wiscasset and Reid State Park may be a result of the small number of individuals tested, or may result from recent population bottlenecks. The latter would limit the ability of populations to adapt to the changing environments and thus impact their long-term survival potential, especially under stressful conditions (Markert et al. 2010).

In addition to seven unique genotypes found within the Schoodic Point population, there were two other novel genotypes. One was from the estuarine population at Wiscasset ME and the other was from the near-by coastal population at Reid State Park ME. The unique genotype from Wiscasset (Genotype 3) was only one locus different from Genotype 2 and may be the result of genetic drift. Another possible explanation is that the low salinity in the estuarine environment induces the unique genotype as suggested by Ram and Hadany (2014); stress-induced mutagenesis can help generate a better adaptive genotype in an extreme environment. The other unique genotype (Genotype 1) was only present in the Reid State Park population, it differs at

three loci from Genotype 2. These three loci were not in the same gene but all reach fixation based on our limited sampling in the population (five individuals). Besides the possibility of genetic drift, it is also likely that these three fixed mutations occurred in succession as three hard selective sweeps events or represent one clone resulting from the effect of clonal interference. As suggested by Lang et al. (2011), clonal interference is far more likely to happen compared to selective sweeps in an asexual population. It is likely that we missed other clones in the Reid State Park population. The genetic distance and differentiation (Table 3) also showed that significantly higher genetic differentiation existed between Reid State Park and the rest of the populations. It is not clear why Reid State Park has such a high genetic differentiation versus the other populations. Bottlenecks or extirpation events in the past followed by subsequent recolonization could lead to high genetic differentiation. Another explanation for the observed high genetic differentiation may be the unique environment in Reid State Park: this site has more sandy habitat than the typical rocky terrain of the Maine coastline. However, *P. umbilicalis* in Reid State Park is different from true "sand-loving" (i.e., psammophytic) species that grow on/in unconsolidated sediments or on rocky substrata impacted by sand scouring and are restricted to such habitats. Many psammophytic species are tough and wiry, such as *Ahnfeltia plicata* (Hudson) Fries, *Ahnfeltiopsis concinna* (J. Agardh) P.C.Silva & DeCew and *Ahnfeltiopsis linearis* (C.Agardh) P.C.Silva & DeCew. Psammophtyic species have two major reproductive adaptations: regeneration of upright frounds from their bases and incomplete or asexual life history (Daly and Mathieson, 1997). Although *P. umbilicalis* in the NW Atlantic has an asexual life history, the morphology of the alga is different from many psammophytic species metioned above and it lacks the regeneration capacity. The *P. umbilicalis* population in Reid State Park may be affected by the sandy environment and this further might impact its prevalent genotype.

However, larger sample sizes are needed to draw any conclusion for the genetic differentiation of the Reid State Park population compared to the other open coastal populations.

The previous study of genetic diversity of these same populations (Eriksen et al. 2016) looked at a much larger number of individuals (221) using three polymorphic SSR markers. A total of six genotypes were identified by SSR, compared to the 11 genotypes found in this study that sampled 37 individuals. One explanation for the lower level of polymorphisms identified by SSR loci is that SSR markers used by Eriksen et al. (2016) were developed from protein coding regions. Expansion and contraction of SSR regions in the protein coding regions may be more functionally constrained than SNPs in the same region. Different genotype patterns were also observed by Eriksen et al. (2016); they reported two or more SSR genotypes in each population, with the largest number of genotypes (4) observed for Fort Stark NH population. By contrast, the present study found four populations that had single genotypes (Reid State Park ME, Wiscasset ME, Nubble Light ME and Dover Point NH) and Schoodic Point with the highest number of genotypes. However, I found much fewer genotypes in my study compared to the AFLP diversity assessed in Blouin and Brawley (2012): they found 41distinct clones in 51 individuals from 2 populations in ME. The differences between my observations and those of Blouin and Brawley (2012) were probably due to contamination by non-target DNA n the AFLP DNA fingerprinting from microbial that embedded in the *Porphyra* cell wall was (Miranda et al. 2013).

With the predominantly asexual reproduction of *P. umbilicalis* in the NW Atlantic (Blouin and Brawley, 2012), genotype diversity among *Porphyra* populations is restricted by the dispersal ability of neutral spores. To date, there is no report on how far *Porphyra* neutral spores can travel. In natural environment with water constantly mixed, spores tend to stay near the surface (Hoffmann and Camus 1989). The rate of spore sinking is low, irrespective of spore size

in other algae species (Hoffmann and Camus 1989). By constrast under lab conditions, *Porphyra* spores tend to settle and attach to substrate within 12 – 24 hours of release into still water (Cao et al. unpub. data). The genetic differentiation among the three geographically close southern populations (Fort Stark, Nubble Light and Dover Point) was low and insignificant (Table 5), possibly because these three populations are within the dispersal range of neutral spores (<30 km apart). It is possible that the genetic differentiation among these populations is mainly driven by disperse. However, we found that genetic variation was slightly higher among populations than within populations based on AMOVA results, suggesting that genetic differentiation could be caused by genetic drift or selection (Excoffier et al. 1992). Genetic drift and selection are more likely to happen to small relatively isolated estuarine populations like Wiscasset ME and Dover Point NH or populations in atypical environments like Reid State Park ME. Although Reid State Park is geographically closest to the estuarine population at Wiscasset, it was genetically most distant to Wiscasset. Genetic drift and selection for specific environments (estuarine environment for Wiscasset and more sandy sediment for Reid State Park) may play an important role in the high genetic differentiation between these two near-by populations (27 km apart).

It is worth mentioning that observed genetic structure and genotype variation for *P. umbilicalis* may also vary year-round and between years. Drenth and coworkers (1994) showed that the genetic structure of asexual fungal populations was different every year as less than 10% of the genotypes survived from year to year. Selection and drift have the ability to change the genetic structure of small populations rapidly (Worrall 2012) and its effects can vary considerably from year to year within a population (Price et al. 1984, Milner et al. 1999). As population size in *P. umbilicalis* drops during the summer due to thermal and UV stress, random genetic drift will increase as the power of random genetic drift is inversely proportional to the

effective population size (Lynch et al. 2016). Environmental impacts on genetic diversity suggests that it is better to collect samples at the same season in order to accurately compare genetic diversity and population structure. The samples used in this study were originally collected by Eriksen et al. (2016). She collected samples at different seasons of the year, and thus my inferences about population structure should be viewed with caution. In future studies, inclusion of a larger number of markers, more samples per population, and attention to seasonal sampling should yield a better estimate of genetic diversity and population structure of *P. umbilicalis* within the Gulf of Maine.

Limitation and Future Directions

In the study, we identified large amounts of putative SNP markers using RNA seq data and applied some of the SNP markers in a pilot population study. During the time the study was conducted, there was no P. umbilicalis genome available for the bioinformatic analysis. Thus, we constructed the de novo reference transcriptome. Now the genome of P. umbilicalis is available, analyzing the RNA Seq data using the genome information may generate a larger set of SNP markers.

Only a limited number of SNP markers were applied to a small number of samples for this pilot population study. More SNP markers and more individuals from more populations are needed to provide a more complete picture of the population structure of P. umbilicalis in the NW Atlantic.

Tables

Table 1 Treatments applied to each pooled *P. umbilicalis* culture

Library	Treatment Conditions
A	Air dried for 4 hours, frozen at -20°C for 3 months, then cultured for 3 hours
B	Cultured for 3 months
C	Air dried for 4 hours, frozen at -20°C for 3 months, then cultured for 2 weeks
D	Fully hydrated fresh tissue
E	Air dried for 4 hours

Table 2 Primers used to amplify regions containing SNP markers in *P. umbilicalis*

Primer	Major Allele SNP	Minor Allele SNP	Position In the Amplicon (bp)	Amplicon length	Annealing Tm	Sequence (The first is the forward primer and the second is the reverse primer)
2	C	T	123	599	62°C	5'-TGGGATCGTCACTCTCATCA-3' 5'-GAGGGTGTTCAAGGGAAAGG-3'
3	C	T	291	407	60°C	5'-TCGGGCGAATACTGTCTCTT-3' 5'-CGCCTCACCGATAATACGAG-3'
4	G	A	878	999	60°C	5'-CCTGTCGCTCTCCTCTCTTG-3' 5'-TTGAGAGCCTTTTTCCAGTCG-3'
8	G	T	164	561	62°C	5'-GTATGAGCGGACGGATAAGG-3' 5'-ATGAGCGAGCGTGGATAAAG-3'
10	G	A	293	398	60°C	5'-ATCGCAAAGGCGTTATCAAG-3' 5'-GTGTCGTCAACTGCAATCGT-3'
13	C	T	183	448	62°C	5'-GATGAAATCTCGGACCCTGA-3' 5'-TGAGGTTGAGCTTCGTGTTG-3'
14	T	C	148	752	60°C	5'-CTCTCCATCTGCTCCTTGCT-3' 5'-TTCTGTCACGCTCATTGGTC-3'
15	C	G	430	500	60°C	5'-CAACGTGATCCGCTACTTCC-3'

						5’-CACCCGACTGCATCATCA-3’
16	G	T	278	382	60°C	5’-GGTCAATGCACAGGATGAGA-3’ 5’-CACGAGCAGATTGTGTTTGC-3’
17	A	C	216	395	65°C	5’-GTGCTCGCCCTCTTCACTAA-3’ 5’-AGACAGAAGGTGGTGGGTTG-3’
18	C	G	103	450	60°C	5’-GACGTATGCTTTGTTCCGAGA-3’ 5’-ACCTCCATTGCGTGGTAGAC-3’
19	C	T	305	447	60°C	5’-GTGGTGGATGTAGGCTCGTT-3’ 5’-AGGTGCGTTGACAGGTGACT-3’
22	A	G	120	397	62°C	5’-GTTATTGCGCTCCTCATCGT-3’ 5’-CGCACAGAGACCACTTCGT-3’

Table 3 Assembly result comparisons between *de novo* reference before and after cleaning up contamination

	# of contigs	N50	GC%	248 Core conserved eukaryotic genes (CEGMA)	
				# of protein	% complete
Trinity output	182,905	512	56.47	216	87%
Trinity minus contamination	42,802	1,090	62.28	206	83%

Table 4 Genetic diversity of seven *P. umbilicalis* populations

Sampling site	Abbr.	Sample Size	Major Allele Freq.	Gene Diversity	PIC
Dover Point, NH	D	5	1.000	0.000	0.000
Fort Stark, NH	F	5	0.969	0.037	0.028
Lubec, ME	L	5	0.892	0.160	0.131
Nubble Light, ME	N	5	1.000	0.000	0.000
Reid Park, ME	R	5	1.000	0.000	0.000
Schoodic Point, ME	S	7	0.890	0.144	0.115
Wiscasset, ME	W	5	1.000	0.000	0.000

Table 5 Genetic distance (below diagonal) and genetic differentiation (above diagonal) between seven sample sites of *Porphyra* within the Gulf of Maine

	D	F	L	N	R	S	W
D		0.500	0.071	0.000*	1.000*	0.237	1.000*
F	0.028		0.000	0.500	0.875*	0.318*	0.813*
L	0.058	0.042		0.071	0.639*	0.200*	0.458*
N	0.000	0.028	0.058		1.000*	0.237	1.000*

R	0.231	0.171	0.223	0.231		0.704*	1.000*
S	0.069	0.097	0.127	0.069	0.265		0.517*
W	0.077	0.105	0.135	0.077	0.308	0.146	

* significant with P value < 0.05

Figures

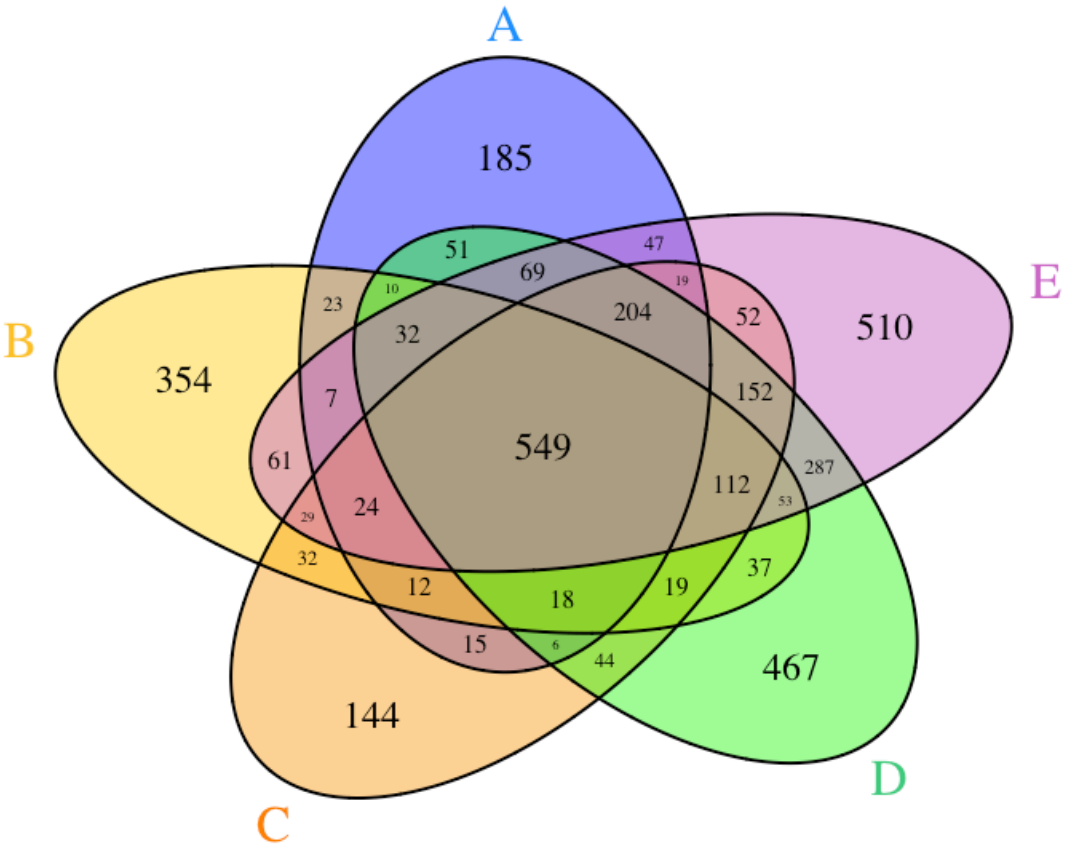


Figure 1 Venn diagram of putative SNPs in each treatment library.

Each oval represents a library. A total of 549 putative SNPs were found in common to all five libraries. Library A-A; Library B-B; Library C-C; Library D-D; Library E-E

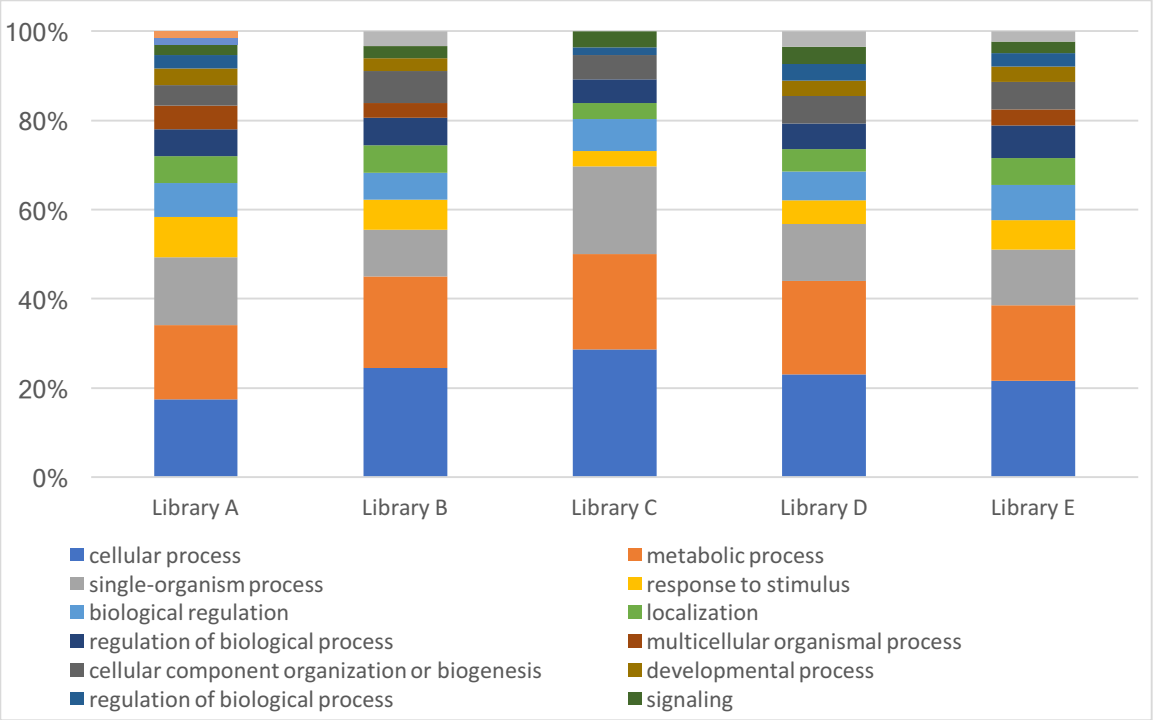


Figure 2 Comparison of functional annotation of genes containing library-specific SNPs.

Y-axis = percentage of library-specific genes.

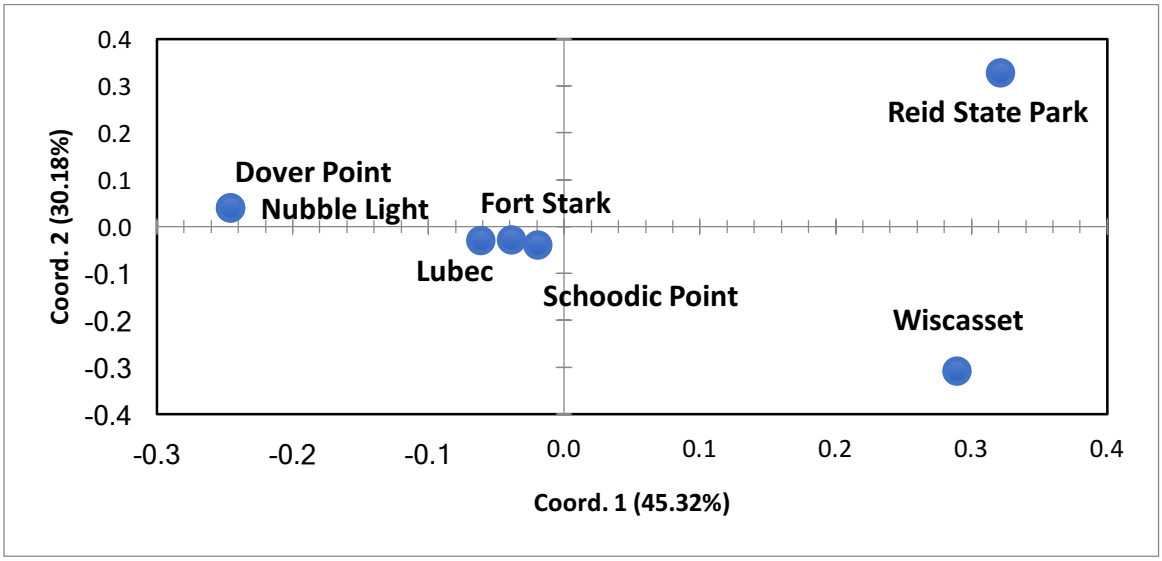


Figure 3 Principle Coordinate Analysis (PCoA) based on genetic distance among seven populations

Dover Point and Nubble Light populations mapped to the same PCoA coordinate. Percentage of the total variation explained by each axes was shown.

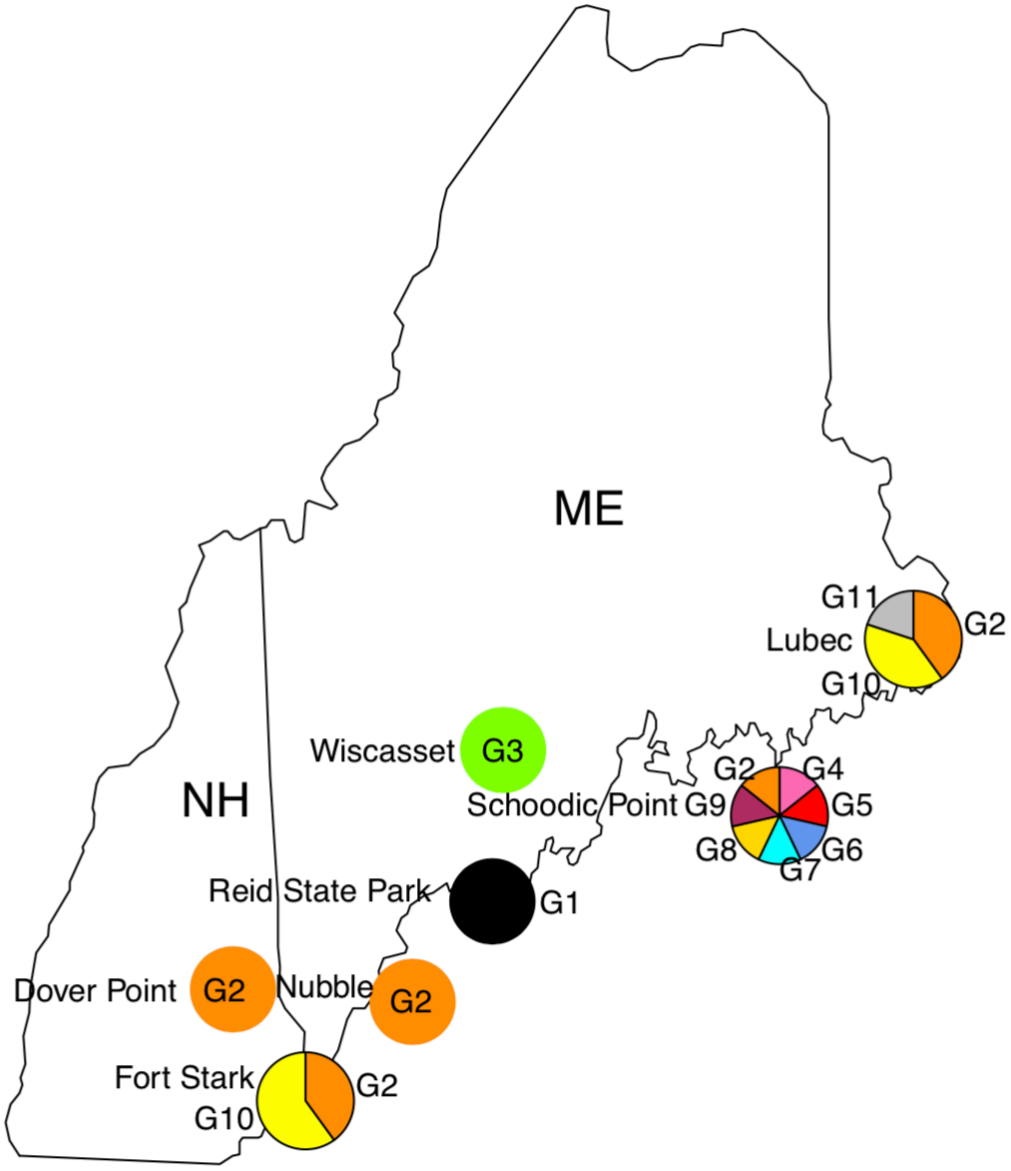


Figure 4 Sampling sites and pie graph showing genotype frequency in each population

Each color represents a different genotype. The number represents the name of the genotype. NH, New Hampshire; ME, Maine. Each population has a sample size of 5 except for Schoodic Point population has 7 individuals.

CHAPTER II

Trehalose and (iso)floridoside production under desiccation stress in the red alga *Porphyra umbilicalis* Kützing and the genes involved in their synthesis

Abstract

The marine red alga *Porphyra umbilicalis* Kützing has high tolerance towards various abiotic stresses. In this study, the amount of floridoside, isofloridoside and trehalose were measured in response to desiccation and rehydration treatments using gas chromatography mass spectrometry (GC-MS); these conditions are similar to the tidal cycles that the alga experiences in its natural habitats. The GC-MS analysis showed that the concentration of floridoside and isofloridoside did not change in response to desiccation as expected for compatible solutes. Genes involved in the synthesis of (iso)floridoside and trehalose were identified from the recently completed *Porphyra* genome, including four putative trehalose-6-phosphate synthase (TPS) genes, two putative trehalose-6-phosphate phosphatase (TPP) genes, and one putative trehalose synthase/amylase (TreS) gene. Based on the phylogenetic, conserved domain, and gene expression analyses, it is suggested that the *Pum4785* and *Pum5014* genes are related to floridoside and isofloridoside synthesis, respectively, and that the *Pum4637* gene is probably involved in trehalose synthesis. Our study verifies for the first time the occurrence of nanomolar concentrations of trehalose in *P. umbilicalis* and identifies additional genes possibly encoding trehalose phosphate synthases.

Introduction

Porphyra umbilicalis Kützing, a macrophytic red alga in the Bangiaceae family, grows in

a physically stressful habitat-the middle to high intertidal zones. The occurence of this alga is confirmed from the Northeast Atlantic to the Northwest Atlantic, although its global distribution requires further review (Brodie et al. 2008). The alga is exposed to daily and seasonally variable temperatures, high levels of irradiance (including UV), and severe osmotic stress and desiccation (Brawley et al. 2017). Some eukaryotic algae, eubacteria, actinomycetes, fungi and yeast maintain osmotic balance by ion or compatible solute accumulation (Wegmann 1986, Brown 1990). During desiccation, *P. umbilicalis* should employ complex mechanisms at morphological, physiological and molecular levels to maintain homeostasis, to sustain cellular integrity, to repair oxidative stress damage and to survive the extreme loss of water.

"Compatible solutes" or low-molecular-weight organic molecules (osmolytes) can accumulate to high levels in order to balance the external osmotic pressure but which do not inhibit cellular function at high intracellular concentration (Brown 1976). The accumulation of compatible solutes functions as thermostabilizing chemical chaperones that help to increase the stability of proteins (Roberts 2005), and as antioxidants (Amor et al. 2005). The most common types of compatible solutes include polyols, sugars, amino acids and their respective derivatives, betaines, ectoines, and occasionally some peptides (Galinski 1995). Specific taxonomic groups do not necessarily use same compatible solutes (Erdmann and Hagemann 2001). There seems to be some functional specialization of the different classes of compatible solutes, which may contribute to their preferential occurrence in organisms that settle in different ecological niches and are tolerant to adverse environment stresses. Polyols (Reed 1990), trehalose (Karsten et al. 2005), digeneaside (Karsten et al. 2005) and heterosides (Kremer and Vogl 1975, Bondu et al. 2009) have been observed in red algae and have been hypothesized to function as compatible solutes for these organisms.

The heterosides are composed of a sugar linked to a polyol; examples include glucosylglycerol (GG) and galactosylglycerols (floridoside and isofloridoside). Glucosylglycerol is mainly found in cyanobacteria, while galactosylglycerols, floridoside and isofloridoside, are widespread in red algae except for the *Ceramiales* (Kirst 1980). Floridoside and isofloridoside are identified in several *Porphyra* species, including *P. umbilicalis* (Wiencke and Läuchli 1981, Meng et al. 1987, Meng and Srivastava 1993, Karsten et al. 1993, Karsten 1999). Wiencke and Läuchli (1981) used various concentrations of artificial seawater medium to impose osmotic stress on *P. umbilicalis*, and found that the total amount of floridoside and isofloridoside increased at high hyperosmotic stress, but then decreased with more extreme hyperosmotic pressure. These authors measured the total galactosylglycerols concentration (floridoside+ isofloridoside) under salt stress, but did not measure how floridoside levels changed during desiccation and rehydration.

The disaccharide trehalose can also play a major role in osmotic adjustment by stabilizing membranes and replacing water (Jain and Roy 2009). Protein stabilizing properties of trehalose are greater than many other compatible solutes because of trehalose's ability to alter the water environment surrounding a protein and stabilize the protein in its native conformation (Kaushik and Bhat 2003, Magazù et al. 2005). High levels of trehalose have been found in some desiccation-tolerant plants, such as the moss *Selaginella lepidophylla* and the resurrection plant *Myrothammus flabellifolius* (Goddijn and Van 1999, Pampurova et al. 2014). Trehalose was found to be not sufficient for desiccation tolerance as a compatible solute in other organisms, such as yeast (Ratnakumar and Tunnacliffe 2006) and two strains of green algae *Klebsormidium* (Kaplan et al. 2012). Surprisingly, trehalose was not detected in various *Porphyra* species

(Craigie 1974, Karsten et al. 2007) despite the notation of a putative trehalose-6-phosphsphate synthase gene in *Pyropia yezoensis* (Deng et al. 2004) and *P. umbilicalis* (Brawley et al. 2017).

Among five known trehalose biosynthetic pathways, the pathway that involves two enzymatic steps catalyzed by trehalose-6-phosphosphate synthase (TPS) and trehalose-6-phosphate phosphatase (TPP) is the most widely distributed and has been found in eubacteria, archaea, fungi, insects, and plants (Avonce et al. 2006). The transfer of glucose from UDP-glucose to glucose 6-phosphate is catalyzed by TPS and forms trehalose 6-phosphate (T6P) and UDP, while TPP dephosphorylates T6P to trehalose and inorganic phosphate (Elbein et al. 2003). Other pathways are only found in eubacteria and archaea (De Smet et al. 2000, Avonce et al. 2006). According to Pade et al. (2015) the biochemical pathways for the two-step synthesis of trehalose and (iso)floridoside are very similar. In the floridoside synthesis pathway, floridoside phosphate synthase (FPS) first catalyzes the formation of floridoside phosphate from glycerol-3-phosphate (G3P) and UDP-galactose. Then floridoside phosphate is dephosphorylated to floridoside by (iso)floridoside phosphate phosphatase. It is possible that the putative trehalose-6-phosphosphate synthase gene found in *P. yezoensis* (Deng et al. 2004) and *P. umbilicalis* may not be involved in the synthesis of trehalose but rather in the synthesis of floridoside, as proposed for the unicellular red alga *Galdieria sulphuraria* (Galdieri) Merola (Pade et al. 2015). The *Porphyra umbilicalis* genome has been previously screened for candidate genes involved in trehalose and (iso)floridoside biosynthesis, but only two of the possible TPS/FPS genes have been identified (Brawley et al. 2017).

In this study, the accumulation of floridoside, isofloridoside and trehalose under desiccation and rehydration conditions were analyzed. Phylogenetic and conserved domain analyses of the putative *TPS* and *FPS* genes in *P. umbilicalis* were conducted in order to

distinguish the functions of genes that were involved in trehalose, floridoside, and isofloridoside synthesis. Quantitative real-time PCR (qPCR) was used to determine the expression profile of the four *TPS/FPS* genes at different times before, during and after desiccation of *P. umbilicalis.* This study represents the first successful attempt to measure trehalose in *P. umbilicalis.* It is also the first examination of how the expression profiles of the *TPS/FPS* genes change relative to the levels of compatible solutes (trehalose and (iso)floridoside) in *P. umbilicalis* under desiccation stress, which is the stress that the alga encounters twice daily in its mixed intertidal habitat. Furthermore, this study identifies probable genes in *P. umbilicalis* that are involved in trehalose and (iso)floridoside synthesis.

Material and Methods

Sample preparation

Porphyra umbilicalis samples were collected from an open coastal population at Fort Stark, New Castle, NH (43°03'30.3"N 70°42'43.0"W) in July 28, 2017 and brought to the lab for species confirmation by comparison to rbcL-rbcS intergenic spacer using methods described in Teasdale et al. (2002). Tissues that were confirmed to be *P. umbilicalis* were cultured under lab conditions at 15°C according to Redmond et al. (2014) for more than 6 hours before exposure to different treatments. The six treatments, each with four replications, were: hydrated, air-dried for half an hour, air-dried for one hour, air-dried for two hours, air-dried for three hours, and rehydrated with autoclaved seawater from Fort Stark, NH for a half hour after air-drying for three hours. Each "individual" (each defined as thalli connected to a single holdfast) was used as one replication for a treatment. Wet tissues were blotted with a soft paper towel to remove any free surface moisture. They were then weighed immediately to determine the initial fresh weight

in order to calculate trehalose and (iso)floridoside contents after various treatments. The initial fresh weight for each sample was around 0.5g.

Analysis of low-molecular-mass organic solutes

After samples were subjected to their various treatments, they were transferred into 5 ml 80% ethanol. Internal standards, sorbitol and inositol, were added into ethanol solution for subsequent gas chromatography mass spectrometry (GC-MS) analysis. The samples were then treated by ultrasound at 70°C in 80% ethanol for two hours in order to extract low-molecular-mass organic solutes. The extracts were centrifuged at 3360 RCF for 10 min to remove particulate matter and the supernatants were collected to dry in a Savant Speedvac Concentrator at -4°C. The concentrated material was then re-suspended in 4ml deionized water (HPLC grade), and centrifuged at 3360 RCF for 10 min. A total of 200 µl of the supernatants were collected and then dried for further derivatization. The derivatizations of samples were performed by Tri-Sil HTP (HDMS:TMCS:Pyridine) Reagent (ThermoFisher Scientific, USA) according to the manufacturer's protocol. After dying the samples in nitrogen gas, they were dissolved in 2ml chloroform and injected into GC-MS for further analysis.

Gas chromatography mass spectrometry analysis was performed using a Trace GC Ultra system (FINNIGAN) comprising a Tri Plus auto-sampler. For GC-MS detection, helium gas was used as a carrier gas at a constant flow rate of 1.0 ml min^{-1}, and an injection volume of 1 µl was employed (splitless injection). The injector and ion source temperature were maintained at 260°C and 200°C, respectively. The oven temperature was programmed for 100°C with an increase of 30°C min^{-1} to 160°C, isothermal for 2 min, and then increased to 280°C with 10°C min^{-1}, ending with a 9-min isothermal step at 280°C. Mass spectra were taken in a full scan mode for fragments from 50 to 650 amu. The mass detector used was a PolarisQ (FINNIGAN), and

Xcalibur™ software (ThermoFisher Scientific, USA) was applied to handle mass spectra and chromatogram results.

Phylogenetic analysis and conserved domain analysis

Representative protein sequences involved in trehalose synthesis in eubacteria, plants, fungi and animals (Avonce et al. 2006) and (iso)floridoside synthesis in the red alga *Galdieria sulphuraria* (Pade et al. 2015) were downloaded from the National Center for Biotechnology information (NCBI) database. The protein sequences were used to identify putative homologous genes in comparison to all versions of the *Porphyra umbilicalis* genome (Brawley et al. 2017), using TBLASTN with an E-value threshold of 10^{-10}. Nucleotide sequences predicted to encode TPS/TPP and FPS proteins were used as queries for BLASTX searches against protein sequences from the *P. umbilicalis* genome. The final version of *P. umbilicalis* genome can be accessed in NCBI (P_umbilicalis_V1) and earlier draft versions of the *P. umbilicalis* genome can be accessed at http://porphyra.rutgers.edu/bindex.php.

A phylogenetic tree was built using functionally characterized protein sequences for TPS enzymes from a wide range of species (Avonce et al. 2006), FPS enzymes from *Galdieria sulphuraria* (Pade et al. 2015) and glucosylglycerol-phosphate synthase (GgpS) from cyanobacteria *Synechocystis sp.* PPC 6803. The complete list of sequences used is provided in Supplemental Table 2. A total of 63 amino acid sequences were aligned using MUSCLE (Edgar 2004) and were edited with Jalview (Waterhouse et al. 2009). The best-fitting evolutionary model for the alignment was evaluated in ProtTest (Darriba et al. 2011). Maximum-likelihood trees were constructed in RAxML (Stamatakis 2014) using WAG model with 1000 bootstraps. Graphical representation and editing of the phylogenetic tree were performed with Dendroscope 3 (Huson and Scornavacca 2012).

The protein sequences of the four putative TPS/FPS enzymes and two putative TPP enzymes were analyzed in Pfam to identify protein domain families. The protein sequences were subsequently aligned with the trehalose-6-phosphate synthase (OtsA) and trehalose-6-phosphate phosphatase (OtsB) protein sequences of *Escherichia coli* (Gibson et al. 2002, Rao et al. 2006) using MUSCLE (Edgar 2004), in order to determine whether there are conserved amino acids important for the binding of substrates and for catalysis.

RNA isolation, cDNA synthesis and quantitative RT-PCR

Tissues were subjected to different treatments and then ground in liquid nitrogen to a fine powder. Total RNAs were extracted according to Eriksen (2014). To remove DNA, the extracts were treated with RNase-free ezDNase™ Enzyme (ThermoFisher Scientific, USA). The quality and quantity of the RNA were visually inspected on a 1% agarose gel and quantified a with DeNovix DS-11. The amount of total RNA was adjusted to 1 ng per sample. DNA-free RNA was reverse-transcribed into cDNA by SuperScript™ IV Reverse Transcriptase (ThermoFisher Scientific, USA) according to the manufacturer's protocol. One of the actin genes was previously demonstrated to be a good endogenous standard for relative RT-PCR under stress conditions in the red algae *Pyropia* (Li et al. 2014, Kong et al. 2015), so this actin was used in this study as an endogenous standard for *Porphyra umbilicalis* under desiccation stress. Gene-specific primers for actin and the *TPS*/*FPS* genes are shown in Table 6. The expression levels were determined with QuantStudio™ 3 Real-Time system (ThermoFisher Scientific, USA) and PowerUP™ SYBR® Green Master Mix (ThermoFisher Scientific, USA). Quantitative PCR was performed in a total reaction volume of 10 μL using PCR cycling of 50°C for 2 min and 95 °C for 2 min; followed by 40 cycles of 95 °C for 1 second and 60 °C 30 seconds. The melt curve was generated at 95°C for 15 seconds with a ramp rate of 1.6°C/second, followed by

60°C for 1 min with ramp rate of 16°C/second, and then ended with 95°C for 15 seconds with ramp rate of 0.15°C/second. Six biological replications were measured for each time point and quantitative PCR were performed in triplicate for each gene at each time point. Relative quantifications were performed following the ΔΔCt method (Livak and Schmittgen 2001). Cycle threshold values generated from the reactions were averaged. The expression levels for different genes were normalized by subtraction of their average cycle threshold values from the mean of the control housekeeping gene. The relative levels of expression changes at different time-point were normalized to the level of expression at the fresh hydrated time point.

Results

Floridoside, isofloridoside and trehalose accumulation under desiccation and rehydration conditions

The changes of floridoside, isofloridoside and trehalose content during desiccation and rehydration treatments are shown in Figure 5. Before desiccation treatment, the levels of floridoside and isofloridoside in *P. umbilicalis* were similar (40-50 μmol/g F.W.), with isofloridoside content being slightly higher than that of floridoside. However, the floridoside and isofloridoside contents were more than five thousand times higher than trehalose content (5-8 nmol/g F.W.) before desiccation. For floridoside content, there was a slight decrease from hydrated control to half hour dehydration treatment (P=0.086), and from 2-hour to 3-hour dehydration treatments (P=0.079). A sudden increase in floridoside occurred from 3-hour dehydration to rehydration treatment (P=0.098). Isofloridoside content did not change significantly for all treatment groups (P> 0.1; Figure 6b). Trehalose content continued to increase during drying, and then its level dropped slightly after long-term desiccation treatment and subsequent rehydration.

Identification of proteins involved in trehalose and (iso)floridoside synthesis in *P. umbilicalis*

The TBLASTN search of several sequential assemblies of the *P. umbilicalis genome* detected four proteins that were homologous to TPS, FPS or GgpS proteins from a wide range of species with significant E-values ($<10^{-10}$), one protein that had high similarity to trehalose synthase/amylase (TreS) in eubacteria with E-value of $2e^{-16}$, and two proteins that had significant similarity to TPP proteins. Information for each putative biosynthetic is listed in Table 7 and the predicted protein sequences for each putative gene are provided in the Supplemental Figure 1. Protein Pum4785 and Pum5014 correspond to protein ID Bu14_0021s0026 and BU14_0082s0055 from GenBank. Gene *Pum8501* was present in the final version of the *Porphyra* genome; however, its protein prediction was missing from the final genome model (Brawley et al. 2017).

The protein family search using Pfam showed that all four putative TPS proteins contained a glucosyltransferase family 20 (GT20) and phosphatase domain(s), and the two putative TPP proteins possessed phosphate phosphatase (PPase) family domains with putative trehalose-phosphatase functions (Figure 6). However, among the four putative TPS proteins, the alignment of the four predicted TPS proteins of *P. umbilicalis* with the trehalose-6-phosphate synthase OtsA of *E. coli* revealed that only protein *Pum4637* contained the conserved motif important for binding of substrates and for catalysis (Figure 6). The TPS and TPP proteins were also aligned with the trehalose-6-phosphate phosphatase OtsB protein sequences of *E. coli*. Although according to Pfam, all TPS and TPP protein sequences contained complete or partial trehalose PPase family, only Pum5014, Pum4738 and BU14_0615S0010 contained the highly conserved residues of the trehalose PPase family active site in their three respective TPP

conserved regions (Figure 6). The protein alignment of the four *P. umbilicalis* TPS proteins with OtsA of *E. coli* (UniProtKB/Swiss-Prot: P31677.3) are provided in Supplemental Figure 2 and the protein alignment of the four *P. umbilicalis* TPS proteins and two TPP proteins with OtsB of *E. coli* (UniProtKB/Swiss-Prot: P31678.2) are provided in Supplemental Figures 3 and 4.

A phylogenetic tree of 63 TPS proteins from a wide range of species, plus FPS proteins from *Galdieria sulphuraria* and GgpS proteins from cyanobacteria *Synechocystis sp.* PPC 6803 were constructed (Figure 7). According to the phylogenetic tree, there were six, four and three putative TPS/FPS proteins in *Pyropia haitanensis, Chondrus crispus,* and *Cyanidioschyzon merolae,* respectively. The phylogenetic tree showed that the red alga *Pyropia yezoensis* and the brown algae *Undaria pinnatifida* and *Sargassum henslowianum* only had one putative TPS/FPS protein.

The phylogenetic tree (Figure 7) showed two distinct clades. The first clade comprised enzymes for glucosylglycerol synthesis from cyanobacteria, whereas the other clade harbored enzymes for trehalose and floridoside synthesis from prokaryotes and eukaryotes. The second clade further divided into two subclades and a *Mycobacterium tuberculosis* outlier. The first subclade contained TPS proteins from nematodes and bacteria, and the second subclade contained all four putative TPS proteins from *Porphyra umbilicalis.* The four putative TPS proteins in *P. umbilicalis* clustered into four different groups. Pum4785 clustered with three other red algal TPS/FPS proteins, including *Galdieria sulphuraria*_Gasu26940 (bootstrap 100); found adjacently, with very low bootstraps, were TPS family proteins from Amoebozoa and Fungi. Pum5014 clustered with red and brown algae (bootstrap 100) adjacent to a cluster containing *G. sulphuraria*_10960 and the unicellular red alga *Cyanidioschyzon merolae* (bootstrap 100). Enzymes Gasu_10960 and Gasu_26940 have been biochemically characterized

and are responsible for isofloridoside phosphate synthase/phosphatase and floridoside phosphate synthase/phosphatase in *Galdieria sulphuraria.* Pum8501 and Pum4637 each clustered separately with red algal TPS proteins (bootstrap 99). The putative TPS/FPS proteins in *Chondrus crispus* and *Pyropia haitanensis* were also separated into the same four clusters as the *P. umbilicalis* proteins.

The expressions of the four *TPS/FPS* genes were analyzed in *P. umbilicalis* under different levels of desiccation and rehydration by qRT-PCR (Figure 8). Similar expression patterns were found for *Pum4785* and *Pum8501*, and their expression levels did not change significantly during gradual desiccation up to two hours; however, the expression levels of *Pum4785* and *Pum8501* dropped significantly after long-term desiccation and rehydration. The most significant increase in expression was found in *Pum4637* that reached a peak of about 3.5 fold at 2h drying in comparison with the fresh control. Its expression level decreased during long-term desiccation (3Hr) and after rehydration. Expression levels of *Pum5014* did not correlate with either desiccation or rehydration.

Discussion

Porphyra umbilicalis grows in the middle to high intertidal zone, thus is exposed to desiccation stress twice a day with changing tides. The alga is also simultaneously affected by many other abiotic stresses. The mechanism(s) of high tolerance to water loss in *P. umbilicalis* is still unclear. Desiccation and salt stress are different types of water deprivation regarding the physiological process of ion uptake and ion ratios (Munns 2002, Kumari et al. 2014). Under hypersaline conditions, seaweed cells can still be in full contact with liquid water of decreased water potential, while desiccation leads to more intense cellular dehydration (Holzinger and

Karsten 2013). Hence, it is not easy to distinguish between the effects of osmotic pressure and toxicity due to increased concentrations of Na^+ and Cl^- ions (Hirasawa et al. 2006). In this study, the levels of (iso)floridoside and trehalose in *P. umbilicalis* were measured under desiccation and rehydration conditions.

(Iso)floridoside and trehalose contents

The contents of floridoside and isofloridoside were both quantifiable before *P. umibilicalis* was osmotically stressed, which agrees with the conclusions from previous studies that these heterosides function as the alga's major carbon reservoir (Li et al. 2001). The content of floridoside and isofloridoside are around 40~50 μmol/g of fresh weight (fully hydrated). This study's measurements of combined floridoside and isofloridoside are in the same order of magnitude as those measured by Wiencke and Läuchli (1981), who showed that the combined floridoside and isofloridoside content are around 100 μmol/g fresh wt^{-1}. However, the level of floridoside in *P. umbilicalis* in this study (40 μmol/g fresh wt^{-1}) was almost ten times higher than that reported for *Pyropia haitanensis* (1140 +70 μg/g fresh wt^{-1} ≈ 4.5 μmol/g fresh wt^{-1}, Chen et al. 2014). The level of isofloridoside in *P. umbilicalis* (45 μmol/g fresh wt^{-1}) was similar to that in *P. haitanensis* (9260 +1190 μg/g fresh wt^{-1} ≈ 39 μmol/g fresh wt^{-1}, Chen et al. 2014). The amounts of floridoside in the latter study may be due to differences in species or treatments, as there was no description of how *P. haitanensis* was treated before the measurement of floridoside.

The current study showed that the floridoside content in *Porphyra umbilicalis* decreased slightly after short-term dehydration treatment and then dropped again after long-term desiccation, while floridoside levels increased significantly after rehydration treatment. The observed decline of floridoside content after dehydration was consistent with a previous study of

Pyropia perforata (another high intertidal red alga) which showed that the highest floridoside content was observed at noon when *P. perforata* was submerged for many hours (as *Porphyra perforata* Meng and Srivastava 1993) and dropped gradually after exposure to air. The drop in the level of floridoside could be possibly explained by circatidal clocks. These clocks, which modulate animal behavior in tune to changing tides, were discovered in animal residents of coastal or estuarine habitats, such as crustacea, annelids, molluscs, fish and some insects. (Wilcockson and Zhang, 2008). Altough circatidal clocks have not been previously identified in seaweeds, it is possible that the alga anticipates the dropping tide and prepares for the accumulation of floridoside before the onstart of water loss. This hypothesis is consistent with the decrease of floridoside observed in response to dehydration in *P. umbilicalis.* However, further studies of possible circatidal clocks in algae are needed to test this hypothesis.

In *P. umbilicalis*, the combined floridoside and isofloridoside levels almost doubled under salinity stress (Wiencke and Läuchli 1981). Similar patterns of increasing floridoside content under salinity stress have also been reported for *Galdieria sulphuraria* (Martinez-Garcia and Van der Maarel 2016), *Pyropia columbina* (Montagne) W.A. Nelson (as *Porphyra columbina* Karsten et al. 1993) and *Porphyra purpurea* (Roth) C.Agardh (Reed et al. 1980). The different amounts of floridoside and isofloridoside produced in *P. umbilicalis* during desiccation and salt stress suggests that *P. umbilicalis* may have different strategies for coping with desiccation versus salt stress. Lippert and Galinski (1992) suggest that the degree of protection against heating, freezing and drying by dissimilar compatible solutes is different, perhaps because (iso)floridoside has a less protective effect towards desiccation than salt stress. Such differences have been demonstrated for other compatible solutes like glycine betaine (Sakamoto and Murata 2002). Reed et al. (1980) and Martinez-Garcia and Van der Maarel (2016) showed

that increased $CaCl_2$ did not change floridoside content in *P. purpurea* and in *G. sulphuraria*. However, the results from Reed et al. (1980) and Martinez-Garcia and Van der Maarel (2016) contrasted with respect to floridoside content response to KCl. Based the two previous studies findings of consistent levels of floridoside under desiccation and $CaCl_2$ stress, and the increase of floridoside levels under NaCl stress, I postulate that floridoside content maybe related to specific ion(s) or to ionic ratios, possibly NaCl-induced K^+ efflux, as desiccation stress did not induce massive K^+ efflux from the *Arabidopsis* (Shabala et al. 2006). Floridoside is hypothesized to be the main soluble photosynthesis product in red alga (Bondu et al. 2009) and to act as a dynamic carbon pool (Li et al. 2001) that can be converted to polysaccharides and other end products. In this study, floridoside behaved like a typical reserve compound that was accumulated when exogenous resources and conditions were good (upon rehydration), and was consumed when water availability was low, unlike a typical stress metabolite (Baud et al. 2008).

The ability for the photosynthetic apparatus to recover from water loss after rehydration, not the ability to avoid desiccation, is the key factor for desiccation-tolerant plants (Kumari et al. 2014). The floridoside level in *P. umibilicalis* increased significantly after rehydration, which suggested that floridoside may either play an important role in the recovery process from desiccation or may be a by-product of the rehydration process. Hagemann (2016) stated that the function of high amounts of compatible solutes is to make the cytoplasm hyperosmotic towards the external medium ensuring water uptake and positive turgor. Increased floridoside after rehydration may help red algal cells take up water and recover from desiccation.

Isofloridoside is considered as one of the predominant heterosides involved in algal osmotic acclimation. Its content remained constant in all treatments in this study, which agrees with previous research on *P. purpurea* that was exposed to salt stress (Reed et al. 1980), but

differs with the increased isofloridoside levels in *Pyropia columbina* under salt stress (Karsten et al. 1993). The different changes in isofloridoside content, in response to salt stress in various *Porphyra* species (sensu lato), suggest that the mechanisms that control isofloridoside levels may vary between these red algae species.

This study is the first time in which trehalose has been quantified in *Porphyra.* Karsten and cowrokers (2007) were not able to detect trehalose in *Porphyra* species using Nuclear Magnetic Resonance (NMR), although trehalose was known to be wide spread in other red algae genera. However, I was successful in using the more sensitive Gas Chromatography Mass Spectrometry (GC-MS) to detect trehalose in *Porphyra umbilicalis.* In a large range of organisms such as bacteria, fungi, nematodes and invertebrates, trehalose accumulates to mM levels, serving as a reserve carbohydrate and stress protectant; in contrast, in most higher plants, only minute amounts can be detected (Wiemken 1990, Goddijn et al. 1997, Crowe et al. 1998). It has been shown that disaccharides preserve the structural integrity of membranes in the dried state by substituting for the structural water hydrogen bonded to the phospholipid headgroups (Welsh 2000). The small amount of trehalose found in *P. umbilicalis* (5-20 nmol/g fresh weight) is similar to that in green algae *Klebsormidium* (Streptophyta) and Cyanobacteria (Kaplan et al. 2012, Reed et al. 1984). The trehalose content remained too low both before and after desiccation treatment to function as a compatible solute. However, the amount of trehalose almost doubled after drying; this suggests that trehalose or its precursor (trehalose-6-phosphate) might function as a signaling molecule under dehydration stress; a signaling role for trehalose has been demonstrated in yeast and certain plants (Elbein et al. 2003, Argüelles 2000). The increase of trehalose corresponded to a decrease of floridoside, which is similar to the relationship between glucosylglycerol (GG) and trehalose in the cyanobacteria. Mikkat et al.

(1996) showed that synthesis of GG was repressed by trehalose and that trehalose uptake caused a decrease in the content of previously synthesized GG in *Synechocystis* sp.

Identifying putative functions of four TPS like-genes in *P. umbilicalis*

Avonce et al. (2006) proposed that eukaryotes only have a TPS/TPP pathway while several eubacterial species have multiple pathways. Using a TBLASTN search of the *Porphyra umbilicalis* genome, a total of four putative *TPS*s, two putative *TPP*s and one putative *Tre*S genes were identified in the genome of *P. umbilicalis* (Brawley et al. 2017). The putative trehalose synthase/amylase protein (TreS, BU14_0165s0020) found in the *P. umbilicalis* genome has previously only been found in eubacteria and archaea (Avonce et al. 2006), suggesting that this trehalose synthase/amylase gene may have resulted from horizontal gene transfer from eubacteria or archaea. Since the putative trehalose synthase/amylase protein (BU14_0165s0020) was found encoded on the scaffold 165 in the *P. umbilicalis* genome version P_umbilicalis_V1 and this gene has mRNA support, it is unlikely that this putative trehalose synthase/amylase gene resulted from microbial contamination of the *P. umbilicalis* genome.

A total of four putative TPS/FPS proteins were identified in *Porphyra umbilicalis,* which was the same number identified in *Chondrus crispus* (Collén et al. 2013). The four putative TPS/FPS found in *P. umbilicalis* and *C. crispus* separated into four clusters (Figure 7). Although there were six putative TPS/FPS proteins found in *P. haitanensis,* my phylogenetic analsysis showed these were distributed in the same four clusters as the four clades found in *P. umbilicalis* and *C. crispus*. The two additional TPS/FPS proteins in *P. haitanensis* may be due to recent gene duplication events. The four putative TPS proteins in *P. umbilicalis* were annotated based on their sequence similarity to TPS protein sequences in eubacteria, archaea, fungi, metazoan and plants (Avonce et al. 2006). These four putative TPS proteins in *P. umbilicalis* all possess

glucosyltransferase (GT20) and phosphatase domains (Figure 6). However, some of these proteins may be involved in (iso)floridoside metabolism in red algae because of the similarity of biochemical synthetic pathways and the physical structural similarities between trehalose and (iso)floridoside (Hagemann and Pade 2015, Brawley et al. 2017).

The ^{14}C labelling experiments conducted by Craigie et al (1968) suggested floridoside and isofloridoside are non-equivalent as metabolites in the Rhodophyta despite their structural similarities. Pade et al. (2015) demonstrated experimentally that the *Gasu_10960* gene coded for isofloridoside phosphate synthase/phosphatase enzyme and the *Gasu_26940* gene coded for floridoside phosphate synthase/phosphatase enzyme in *Galdieria sulphuraria*. Among the four putative TPS/FPS proteins, Pum5014, which clustered with with *Gasu_10960,* is probably involved in isofloridoside synthesis, while Pum4785, which clustered with Gasu_26940, is related to floridoside synthesis (see Figure 7). According to the results from the gene expression analysis (Figure 8), the mostly stable expression of the *Pum5014* gene was consistent with the unchanged isofloridoside content under desiccation and rehydration status, except for the minor down-regulation after short-term desiccation. The expression of the *Pum4785* gene had a minor increase right after short-term desiccation, while expression dropped after long-term desiccation and did not increase after rehydration. The decline of floridoside under long-term desiccation corresponded to the reduced expression of the *Pum4785* gene. However, the minor drop of floridoside content after short-term desiccation and the significant increase after rehydration treatment cannot be explained by the expression of *Pum4785* alone. An inconsistency between transcriptional expression and floridoside content has also been found in *G. sulphuraria* under high salt condition. Pade et al. (2015) suggested that biochemical activation of the synthetic

enzyme(s) probably regulates stress-proportional floridoside accumulation rather than transcriptional regulation.

The other two putative TPS proteins (Figure 7), Pum4637 and Pum8501, are candidates to have TPS functions involved in trehalose synthesis. Unicellular organisms and vertebrates typically possess single copy of gene encoding trehalose biosynthesis enzymes, while terrestrial plant genomes contain a remarkably larger number of homologs putatively involved in trehalose metabolism (Vandesteene et al. 2010). However, usually only one protein among the numbers of putative homologs possess TPS function. In *Arabidopsis* (to date), only *TPS1* among the 11 putative *TPS* genes has been shown to have TPS activity (Blázquez et al. 1998). In *Arabidopsis,* the TPS1 protein contains a TPS domain at its amino-terminus and a fused putative TPP domain at its carboxyterminus (Eastmond and Graham 2003); however, the TPP domain lacks consensus sequences that are conserved in phosphatases and has no demonstrated function (Zentella et al. 1999). Gibson et al. (2002) successfully identified the amino acid residues functioning in substrate binding and catalysis in *Arabidopsis* TPS when compared it with the reported 3D structure of the *E. coli* TPS enzyme (PDB: 1GZ5). In *P. umbilicalis,* all residues that are involved in the binding of glucose 6-phosphate (Arg9, Trp40, Tyr76, Trp85 and Arg300) and in the binding of UDP-glucose (Gly22, Asp130, His154, Arg262, Asp361 and Glu369) (Gibson et al. 2002) are conserved only in Pum4637 among the four putative *P. umbilicalis* TPS proteins. Avonce et al. (2006) showed that all the residues mentioned above are conserved in proteins displaying TPS activity, or in proteins from organisms that are known to produce trehalose. Proteins without conserved residues from *Arabidopsis* and yeast did not complement a *ΔTPS1* mutant of *Saccharomyces cerevisiae* (Bell et al. 1998, Vogel et al. 2001), and do not possess TPS activity. The expression of *Pum4637* was up-regulated in response to desiccation, which also

corresponded to the increase of trehalose content under desiccation; thus I propose that Pum4637 is the most promising candidate for TPS enzyme catalyzing the trehalose synthesis. The up-regulation of *Pum4637* was consistent with TPS up-regulation in a wide range of species, such as bacteria *Bradyrhizobium japonicum* (Kirchner 1896) Jordan 1982 *(*Cytryn et al. 2007) and higher plants (Nepomuceno et al. 2002, Jiang et al. 2010, Junior et al. 2013, Xu et al. 2017) under different stresses. The up-regulation of the *Pum4637* gene reached a peak of about 3.5-fold at 2h drying. However, this did not correspond to the changes of trehalose that peaked at half-hour drying and stayed almost stable during the desiccation process and rehydration. These observations may be explained by the regulation of the intracellular concentration of trehalose through hydrolysis by trehalase.

Trehalose-6-phosphate phosphatase, which belongs to the HAD (L-2-haloacid dehalogenase) superfamily of magnesium-dependent phosphatases/phosphotransferases, has three highly conserved motifs (Avonce et al. 2006, Farelli et al. 2014). TPP domains in TPS proteins usually lack consensus sequences that are conserved in other phosphatases and are unknown for domain function (Zentella et al. 1999). These observations are partially consistent with what was found in the present study. The two putative TPS-like proteins (Pum8501 and Pum4637; Figure 6) in *Porphyra umbilicalis* lacked conserved TPP domains, while TPP domains were conserved in the two putative FPS like proteins (Pum5014 and Pum 4785; Figure 6). In higher plants, such as *Arabidopsis* (Vogel et al. 1998), TPS domains are usually independent from TPP proteins. Although Pfam identified TPP signatures in *P. umbilicalis*; the sequences of BU14_0615S0014, Pum8501 and Pum4637 proteins lack some of the conserved motifs, suggesting that it is unlikely that these proteins have TPP activity. The lack of conserved TPP domains in Pum8501 and Pum4637 protein sequences is similar to those of the AtTPS1-4

subfamily in *Arabidopsis*, where only the TPS or TPS-similar domains are present and TPP domains are absent (Leyman et al. 2001). The protein BU14_0615S0014 sequence was conserved in all 3 key amino acid regions except for missing the key amino acids near the N-terminal end (DXDX sequence). In *Arabidopisis,* the first aspartate residue forms a phosphorylated intermediate with the substrate, while the second aspartate residue plays an important role in catalysis (Farelli et al. 2014). They showed that replacement of the catalytic DXDX resulted in the loss of all detectable activity of TPP in *Arabidopsis*; this leaves BU14_0615S0010 as the most probable candidate with TPP activity to convert trehalose-6-phosphate to trehalose.

In conclusion, a total of seven genes that are involved in (iso)floridoside and trehalose synthesis have been found in the *P. umbilicalis* genome. Based on phylogenetic analysis, *Pum4785* and *Pum5014* genes are probably involved in floridoside and isofloridoside synthesis. Gene *Pum4637* is probably related to trehalose-6-phosphate synthase as it has all the conserved amino acid residues important for the binding of glucose 6-phosphate and the binding of UDP-glucose for TPS. The enzyme function for *Pum8501* gene is still unclear. The trehalose synthase/amylase gene (*BU14_0165s0020*) in the *P. umbilicalis* genome may have resulted from horizontal gene transfer from eubacteria or archaea. Among the two predicted TPP proteins present in *P. umbilicalis,* only BU14_0615S0010 is likely to function as a TPP while the other TPP-like protein BU14_0615S0014 did not possess all the conserved motifs important for TPP activity.

The present study documents the first time that trehalose has been documented to occur in *P. umbilicalis.* However, the trehalose concentration was too low to be considered as a compatible solute in this taxon. Floridoside and isofloridoside concentrations were measured in

P. umbilicalis undergoing desiccation and rehydration. The content of isofloridoside did not change dramatically during the desiccation and rehydration processes, while the level of floridoside decreased during desiccation and then increased after rehydration, which suggests that neither floridoside nor isofloridoside act as compatible solutes in *P. umbilicalis* under desiccation stress.

Limitation and Future Directions

In this study, I identified genes involved in trehalose, floridoside and isofloridoside synthesis in *Porphyra umbilicalis* and predicted their putative functions by the phylogenetic and domain analyses. In the future, biochemical analyses are needed to confirm the functions of these genes. For example, the putative TPS and TPP genes can be transferred to mutant varieties of yeast to see see can rescue the wildtype phenotype of yeast.

Tables

Table 6 Primers used for the expression analysis of four putative TPS/FPS genes

Name	Sequence (from 5' to 3')	fragment length (bp)
Pum4785_qPCR_fw	CCGTTACCCTCCCTTCGTC	96
Pum4785_qPCR_rev	TACCAAGCGAGCCGTAATCA	
Pum5014_qPCR_fw	GTGAACGACGGGGAGTTTG	106
Pum5014_qPCR_rev	TGACCAAGCTCTCGATGGAA	
Pum8501_qPCR_fw	CCATGTGGTGCTCGTACAGA	139
Pum8501_qPCR_rev	GTAGTGGATGGGCATGTCGT	
Pum4637_qPCR_fw	CAATGCGGAACAAGGTGCTG	62
Pum4637_qPCR_rev	GTACGTGTGGAAGCCGATCA	
Actin_qPCR_fw	GGGTACAGCTTCACGACGT	135
Actin_qPCR_rev	TTCGTACTCCTTCTCCAGCG	

Table 7 Gene and protein information for the putative TPS/FPS, TreS and TPP enzymes in *P. umbilicalis*

	Gene name	Protein name	Genome assembly Version	Scaffold
TPS gene family[#]	*Pum4637*	Pum4637	Assembly V_0.77	None*
	Pum8501	Pum8501	Final version	483
	Pum4785	Pum4785 (Bu14_0021s0026[!])	Final version	21
	Pum5014	Pum5014 (BU14_0082s0055)	Final version	82
	TreS	TreS (BU14_0165s0020)	Final version	165
	TPP1	TPP1 (BU14_0615S0010)	Final version	615
	TPP2	TPP2 (BU14_0615S0014)	Final version	615

* There is no scaffold information for an earlier version of the *P. umbilicalis* genome.

[!]The nomenclature inside the parentheses correspond to the protein IDs listed in GenBank.

[#]The genes *Pum4637*, *Pum8501*, *Pum4785, Pum5014* are all putative TPS family genes.

Figures

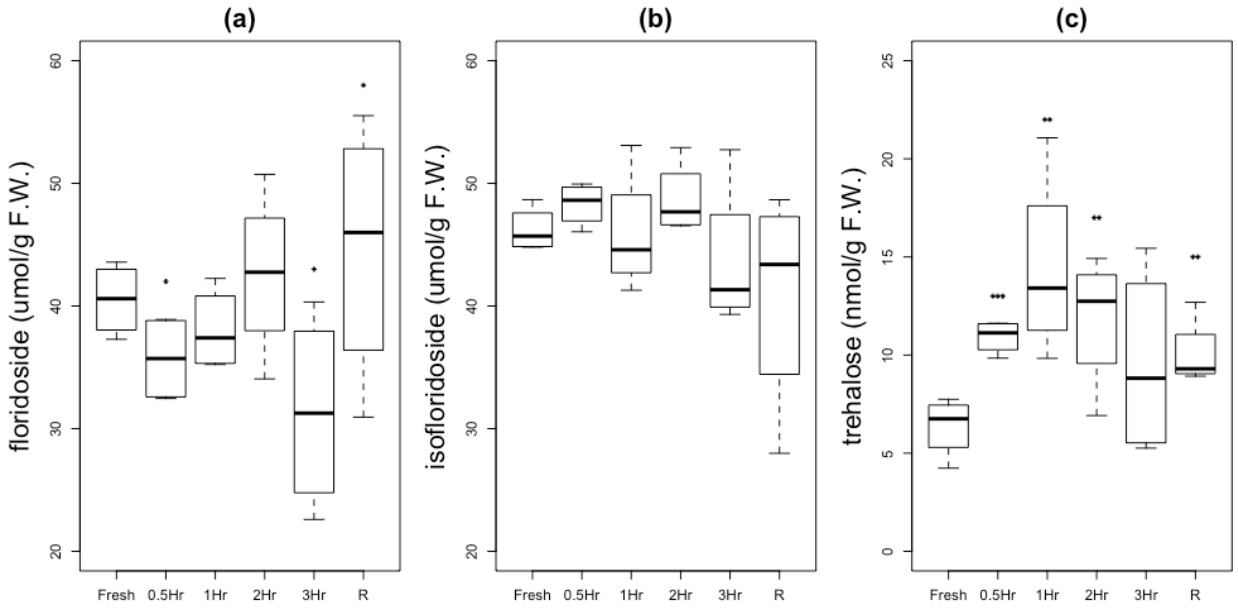


Figure 5 Floridoside, isofloridoside and trehalose contents under drying or rehydration conditions

The six time points are as follows: (1) fresh hydrated; (2) half-hour drying effect; (3) one-hour drying effect; (4) two-hour drying effect; (5) three-hour desiccation state; (6) half hour rehydration after three hours of desiccation. (*P<0.1 **P<0.05 ***<0.01)

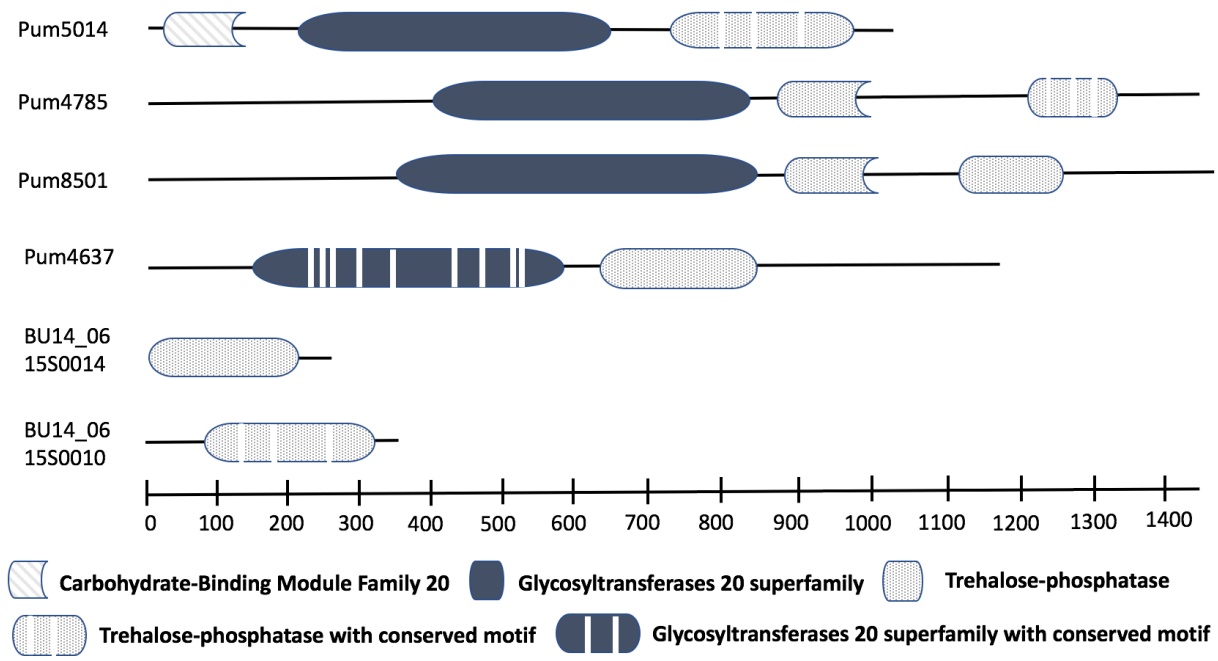


Figure 6 Results from the conserved domain search (CD-Search) for the TPP/FPS and TPP genes in *P. umbilicalis*

The domains with conserved catalytic amino acids required for the TPS or TPP activity contained white stripes.

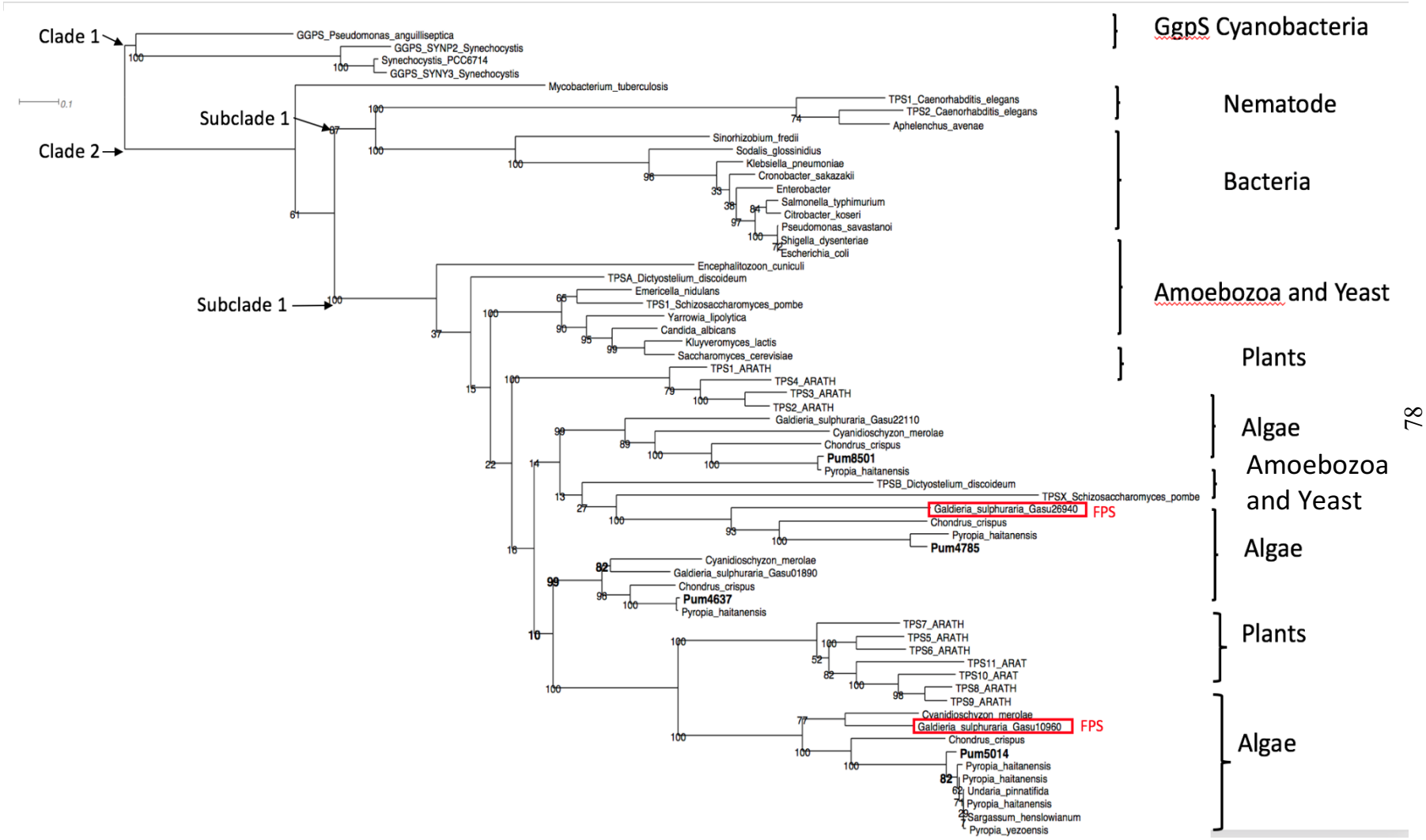


Figure 7 Unrooted phylogenetic tree of amino acid sequences from putative trehalose 6 phosphate synthases and putative (iso)floridoside phosphate synthases using maximum likelihood algorithm

The scale is shown in the upper left-hand corner and bootstrap values are shown on the nodes. The proteins of interest from *Porphyra umbilicalis* are in bold. The cutoff for statistically significant bootstraps is 70.

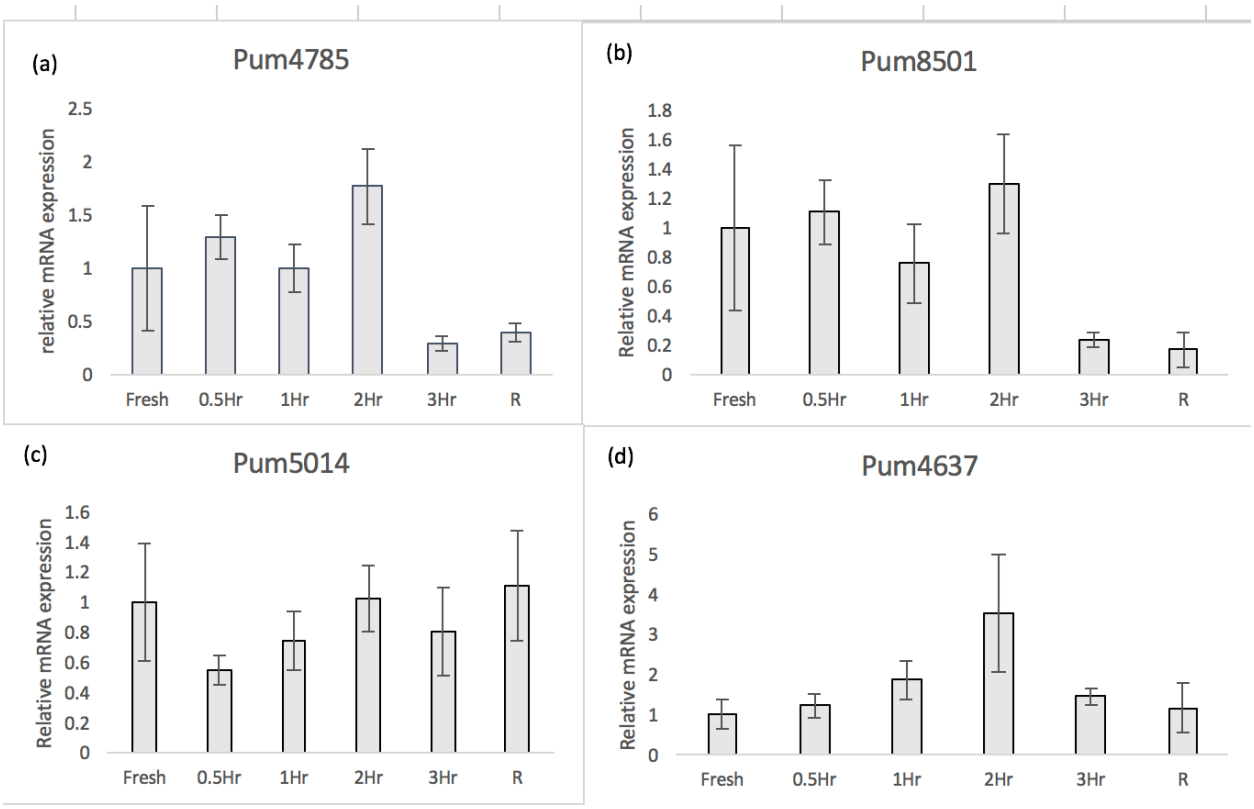


Figure 8 Relative expression of the putative trehalose-6-phosphate synthase genes in *P. umbilicalis* under drying or rehydration: (a) Pum4785; (b) Pum8501; (c) Pum5014; (d) Pum4637

The six time points are as follows: (1) fresh hydrated; (2) half-hour drying; (3) one-hour drying; (4) two-hour drying; (5) three-hour desiccation state; (6) half hour rehydration after three hours of desiccation. The constitutively expressed actin gene served as a control gene and the expression level of each gene was normalized by reference to the constitutively expressed actin gene; relative level of expression changes were normalized to that of the fresh hydrated time point. Error bars represent standard deviation.

Chapter III

Identification of desiccation stress related genes in the marine red algae *Porphyra umbilicalis* Kützing revealed by transcriptomics

Abstract

The red alga *Porphyra umbilicalis* is an ecologically and economically important macroalga in the Northern Atlantic. The alga regularly experiences desiccation and rehydration cycles with the tidal cycles. A detailed analysis for the responses to water loss and water gain is important to understand the mechanisms of desiccation tolerance in *P. umbilicalis*. In this study, a comparative transcriptomic analysis was performed using *P. umbilicalis* under fresh, dehydrated, desiccated and rehydrated conditions. During the water loss process, there were no differentially expressed genes (DEGs) in common to dehydration and desiccation, suggesting that distinct molecular responses may occur during dehydration and desiccation. In comparison to the water loss process, there were more DEGs found with rehydration. The functions of the DEGs under water gain processes suggested that rehydration-induced responses play an important role in the mechanisms of desiccation tolerance in *P. umbilicalis.* The functions of DEGs under stresses are discussed in detail for the roles of osmolytes, photosynthesis and signaling.

Introduction

Algae typically live in aquatic ecosystems, but some species also live partly or permanently under aeroterrestrial conditions. Prolonged desiccation and the subsequent rehydration have negative effects on living cells, such as DNA strand breakage, protein

denaturation, and perturbation of lipid membrane by dramatic phase changes (Bieger-Dose et al. 1992, Prestrelski et al. 1993). For nearly three decades, considerable attention has been paid to the salinity tolerance in marine algae (Kirst 1990). The extensive studies of the green alga *Dunaliella,* which inhabits high saline environments, have provided substantial insight into the response of halophytes to osmotic stress and their adaptive mechanisms (Adams et al. 1998, Bental et al. 1990, Einspahr et al. 1988, Ha and Thompson 1992, Niu et al. 1993). Few studies have focused on the desiccation stress mechanisms in marine algae because of the assumption that tolerance for salt and desiccation were due to closely related mechanisms (Zhu 2002); and because it was easier to manipulate salt stress in laboratory settings. However, desiccation stress is pervasive in intertidal marine algae and the study of desiccation stress provides direct answers to the stress-responsive mechanisms for water loss in algal natural habitats.

The red algae *Porphyra, sensu lato*, are economically important crops. Some *Pyropia* species that are widely used for nori were classified as *Porphyra*. Historically, *Porphyra umbilicalis* Kützing was collected for food as laver, and now represents a developing mariculture crop (Blouin et al. 2010). As a food source, *Porphyra* have high level of proteins, free amino acids and a high ratio of ω3: ω6 fatty acids (Mouritsen 2013). Furthermore, *P. umbilicalis* can be used as a biofilter for integrated multi-trophic aquaculture (IMTA), and as a partial replacement for fishmeal (Walker et al. 2009; Kremer et al. 2004). *Porphyra umbilicalis* grows in the intertidal zone, where it is exposed to diverse abiotic stresses, such as daily and seasonally variable temperatures, high levels of irradiance (including UV), and severe desiccation stress (Brawley et al. 2017). Southward (1958) stated that the desiccation tolerance was the most important factor in determining the zonation of intertidal organisms and whether an alga can survive in the intertidal zone. The red alga *P. umbilicalis* experiences alternation between aquatic

and aero-terrestrial environments with the tidal cycles so that it should be equipped with complex response mechanisms at morphological, physiological and molecular levels to survive the severe loss of water. Hence, it can serve as a red algal model system to study the mechanisms of desiccation in algae (Royer 2017).

Desiccation tolerance is well studied in cyanobacteria, lichens, and mosses (Holzinger and Karsten 2013). Reactive oxygen species (ROS) defense, repression of membrane phase transition and formation of cellular glass are considered as the three major desiccation tolerance mechanisms in land plants (Liu 2009). However, the lack of protective structures (e.g. the waxy cuticle of terrestrial plants) and the lack of the ability to actively regulate the transpiration rate (e.g. by plant stomata) can easily lead to severe desiccation in algae under water-limited conditions (Holzinger et al. 2014).

Some studies have attempted to unravel the cellular and metabolic mechanisms associated with the desiccation stress tolerance in *Porphyra* (Sampath-Wiley et al. 2008, Lipkin et al. 1993, Kim et al. 2008, Liu 2009). Sampath-Wiley et al. (2008) found that desiccation was the greater facilitator of photo inhibitory damage and ROS generation in *P. umbilicalis* during photosynthesis II. The abilities to recover photosynthesis (Lipkin et al. 1993) and to recover nutrient uptake after desiccation (Kim et al. 2008) play important roles in *P. umbilicalis* and other high intertidal *Porphyra* species under desiccation stress. The cytoplasm of *P. umbilicalis* forms a stable glass and the molecular mobility is lower in *P. umbilicalis* after drying (Liu 2009) than in a less desiccation tolerant species *Pyropia yezoensis*. A dehydrin-like protein with a molecular weight of 17 kDa, which is extremely hydrophilic under dehydrative conditions, was suggested to play a key role in *P. umbilicalis* under desiccation stress (Liu 2009). Contradictory observations were found for oxidative stresses and antioxidant activities in *Porphyra* and related

taxa. In *Pyropia columbina*, increased reactive oxygen species (ROS) levels and increased antioxidant enzymes activity were observed during desiccation, while both ROS levels and antioxidant enzymes activity decreased to the basal levels after rehydration (Contreras-Porcia et al. 2010). However, Sampath-Wiley (2008) found increased antioxidant metabolism during summer emersion but not during winter emersion when compared with submersed algae and suggested that irradiance levels and temperatures, instead of desiccation, were the cause of increased antioxidant metabolism in *P. umbilicalis*. A similar result was found by Liu (2009), who showed ROS defense was not the key for the high desiccation tolerance in *P. umbilicalis* when compared with less desiccation tolerant species.

Beyond the previous physiological and biochemical studies of desiccation in *Porphyra sensu latto*, little is known regarding the stress signaling pathways and downstream genes involved in the high desiccation tolerances in the high intertidal alga *P. umbilicalis.* With the rapid development of next generation sequencing (NGS) technologies in recent years, RNA-seq provides a cost-effective and high-throughput approach with a wide variety of applications to assess the whole transcriptome, including measuring expression levels. Genome-wide expression profiles of *Pyropia haitanensis* have been studied in response to dehydration/rehydration stress using deep sequencing technology (Wang et al. 2015). However, this earlier study was limited to two replicates per treatment, which is generally considered too few to achieve accurate statistical power to discover differential expression genes (Schurch et al. 2016). The desiccation response genes in a closely related red alga *Pyropia tenera* were identified using transcriptome sequencing generated by GS_FLX 454 platform (Im et al. 2017). However, again, there was no replication of treatments and thus no statistical analysis to support the DEGs. It is suggested that at least six

biological replicates should be used for RNA-seq experiments to identify DEGs (Schurch et al. 2016).

Holzinger and Karsten (2013) suggested that dehydration and desiccation tolerances must be clearly distinguished. In plants, dehydration was defined as a steady loss of water, while desiccation was the final result of dehydration, while the water status was equilibrated with the air (desiccated is equal to extremely dehydrated, Zhang and Bartels 2018). It has also been shown by transcriptome analysis (Ma et al. 2015) and proteome analysis (Ingle et al. 2007) in plants that the mechanisms underlying dehydration and desiccation may be different.

In the present study, the desiccation tolerant red alga *Porphyra umbilicalis* was exposed to dehydration, desiccation and rehydration in order to obtain information about the genes involved in the acquisition of dehydration and desiccation tolerance, and recovery after desiccation of the alga. The molecular genetic mechanisms in *Porphyra* is applicable to greater understanding of intertidal algae desiccation.

Material and methods

Porphyra umbilicalis growth and treatment conditions

Porphyra umbilicalis samples were collected on December 5, 2016 from the high intertidal zone from six exposed rocks during low tide in an open coastal population at Fort Stark, NH (43°03'30.3"N 70°42'43.0"W). The six rocks were at similar tidal level so that samples were adapted to similar levels of desiccation in the natural habitats. The individuals and corresponding rocks information are listed in Table 8. Samples were brought to the lab for species confirmation by comparison to rbcL-rbcS intergenic spacer using methods described in Teasdale et al. (2002). Tissues that were confirmed to be *P. umbilicalis* were cultured under lab

conditions at 15°C according to Redmond et al. (2014) for more than 6 hours before being exposed to different treatments. The four treatments were: the hydration group, the half hour air-dry group, the three-hour air-dry group, and rehydration for half hour after three-hour air-dry group. All treatments were maintained at the same temperature and light intensity. Each treatment had six biological replications/individuals. An "individual" was defined as an thallus connected to a single holdfast; an individual was used as one replication for a treatment.

RNA preparation and sequencing

Total RNA was isolated from each individual according to Eriksen (2014). To remove DNA, the extracts were treated with RNase-free ezDNase™ Enzyme (ThermoFisher Scientific, USA). The quality and quantity of the RNA were checked by electrophoresis on 1% agarose gel and by DeNovix DS-11 microspectrophometry. The integrity and purity of the RNA samples were assessed using an Agilent 2100 Bioanalyzer System (Agilent, Santa Clara, CA). All samples had an RNA Integrity Number (RIN) of 90 or higher.

The cDNA libraries were prepared with Illumina Truseq ® RNA LT Sample Prep Kit (Illumina, San Diego, USA) with poly A selection according to the manufacturer's instructions. Completed libraries were then subsequently sequenced on Illumina HiSeq 2000 with 2 × 150-bp paired-end reads at the Hubbard Center for Genome Studies at the University of New Hampshire. The raw reads are deposited in the sequence read archive (SRA, https://www.ncbi.nlm.nih.gov/sra/PRJNA496319).

RNA Seq analysis

Adapters and low-quality reads were trimmed with a PHRED score of 2 by BBDuk (Joint genome institute) from raw sequencing data, followed by Rcorrector (Song and Florea 2015) for error correction. Quality control was performed using FastQC

(http://www.bioinformatics.babraham.ac.uk/projects/fastqc). Reference genome sequences and gene annotation files of *Porphyra umbilicalis* were downloaded from NCBI. The processed reads were then mapped to the reference *P. umbilicalis* genome with default configuration using STAR (Dobin et al. 2013) with second mapping pass. HTSeq (Anders et al. 2015) was used to extract count results under four conditions, each with six biological replications.

Batch effects are prevalent in RNA-seq experiments, which can contribute unwanted variations to the data. Such effects are widely recognized as a major source of bias and variability in high-throughput experiment (Leek et al. 2012, Peixoto et al. 2015). A principal component analysis (PCA) plot was performed as a visual tool to determine whether batch effects existed in this data set. The Surrogate Variable Analysis (SVA) package (Leek et al. 2012, Peixoto et al. 2015), which functions to estimate and then remove unwanted surrogate variables, was used to identify and estimate potential unknown batch effects. The Differential Expression Sequence software, DESeq, was utilized to perform differential expression analysis after applying the normalization method for the batch effect identified by SVA package (Leek et al. 2012) and a PCA plot was used to determine if batch effects were corrected after the normalization method. Differentially expressed genes (DEGs) were identified, with the threshold $|\log2 (\text{Fold change})| > 1$ and adjust p-value < 0.05 used as criteria for significant differences of gene expression. All the scripts used can be found at https://github.com/yov2/Porphyra_RNAseq_desiccation.git.

All DEGs sequences were searched against the NCBI NR database using the BLASTX program with an E-value cutoff of $1.0E^{-10}$. Annotation was conducted using Blast2go (Götz et al. 2008). Specific gene functions and biological pathways were annotated using the Kyoto

Encyclopedia of Genes and Genomes (KEGG, http://www.kegg.jp/) database in Blast2go. A histogram of the GO annotation analysis was generated by the software WEGO (Ye et al. 2018).

Results

Global transcriptional changes under desiccation stress

To investigate the water stress responses, transcriptome analysis was performed on *Porphyra umbilicalis* in response to dehydration (short-term water loss), desiccation (long-term water loss), and rehydration. High-quality mRNAs were purified from twenty-four *P. umbilicalis* samples representing six biological replicates in each of four treatments [hydrated control condition (F), dehydration (H), desiccation (D) and rehydration (R)]. A total of 1,094,184,998 paired-end raw reads were obtained from Illumina HiSeq 2000 sequencing platform, with average read length of 150bp and GC-content of 58%. The number of raw reads per sample ranged from 36.68 million to 57.18 million, with an average of 45.59 million reads per sample. After quality filtering and adapter trimming, a total of 1,087,391,122 reads (99% of the raw reads) passed the quality filters and were kept for further analysis. The number of clean reads in each sample ranged from 36.46 to 56.84 million. The error correction step did not change the number of reads in each sample library. Clean reads from the twenty-four samples were then mapped to the genome of *Porphyra umbilicalis* GCA_002049455.2 P_umbilicalis_v1. A total of 32.43 to 42.97 million clean reads in each sample were mapped to the genome, accounting for 75.6% to 89.45% of the total clean reads. The overview of the sequencing for each library, such as the number of raw reads, the number of reads kept after trimming, the number of bases corrected by Rcorrector, and the number of reads mapped to the genome, are provided in supplemental Table 3.

To evaluate if RNA-seq data was affected by unknown biological factors (e.g. batch effects) in addition to the one of interest, PCA was performed on normalized data. The PCA plot A in Figure 9 showed that the first (50.58%) and second (6.88%) components represented most of the variation pattern with a cumulative proportion of 57.46%. All individuals from rock 1, 3 and 4, plus two individuals from rock 2 and rock 5 were grouped at the left side of PCA1, while the other samples were clustered at the right side of PCA1. After removing batch effects estimated by package *sva*, the PCA plot B showed a pattern that corresponded to the treatments. Although the sample D6 is separated from the other samples in the PCA plot, I can not conclude that sample is a significant outliner because the first component of PCA after removing batch effects only explain 15.92% of the variation. I compared the results with and without sample D6. A total of 22 differentially expressed genes (DEGs) were lost and less than 10% of the DEGs were different before and after removing sample D6. In order to provide a more complete picture of the genetic mechansims of *P. umbilicalis* under desiccation stress, all DEGs from the 24 samples were used for further analysis.

The DEGs were identified as those genes with more than 2-fold change in transcriptional expression (adjust p-value < 0.05). Of 12,027 total transcripts expressed by *P. umbilicalis,* there were respectively 117, 37, and 96 DEGs during water loss condition (between fresh and dehydration treatments, between dehydration and desiccation treatments, and between fresh and desiccation treatments) (Figure 10). A total of 100, 22 and 72 DEGs were up-regulated during the process of water loss, from fresh to dehydrated, from dehydrated to desiccation, and from fresh to desiccation, respectively. There were fewer DEGs found to be down-regulated during water loss process. A total of 17, 15 and 24 DEGs were down-regulated from fresh to dehydrated, from dehydrated to desiccation, and from fresh to desiccation, respectively.

A larger number of DEGs were found in the rehydration treatment in comparison to the water loss treatments (dehydration and desiccation). The comparative analysis showed that there were 194, 148 and 291 DEGs during the water loss process, of which 465 were up-regulated and 168 were down-regulated during the water recovery process. More DEGs were up-regulated in the rehydration process (124, 114 and 227 DEGs, respectively) than those that were down-regulated (70, 34, and 64 DEGs, respectively, Figure 10).

Venn analyses were plotted separatly as water loss and rehydration processes (Figure 11). The Venn diagram of the water loss process indicated that no gene was in common among fresh/dehydration, fresh/desiccation, and dehydration/desiccation conditions during the process of the water loss. In the rehydration process, a total of 55 genes were regulated in common to rehydration/fresh, rehydration/dehydration and rehydration/desiccation conditions during the process of rehydration. Of the shared 55 DEGs, 54 were up-regulated genes and only one was down-regulated genes.

Functional annotation

Top BLAST hits of the transcripts sequenced were assigned GO terms in categories of cellular component, molecular function, and biological process. Among the 490 unique DEGs found at different levels of water stress, a total of 250 unique DEGs (51%) were annotated with GO terms in Blast2GO software, and the remaining genes (49%) showed no significant sequence homology with any known genes in the NR database. The putative functions of the DEGs with GO terms are provided in supplemental Tables 4 to 9. A histogram of the GO annotation analysis of the differential expression genes during water-loss process and water gain process in *Porphyra umbilicalis* (Figure 12) showed that different biological processes occuried during the recovery and the water loss stages, such as signaling, localization, grow, immune system

processes, detoxification and locomotion. There were more DEGs with the GO functions of biological regulation, response to stimulus or regulation of biological process identified during water-gain process than those during water-loss process.

Discussion

Desiccation stress is the most common environmental stress that affects intertidal algae; hence, an understanding of desiccation tolerance mechanisms for these organisms will provide new insights about physiological adaptations to recurrent environmental extremes. The red alga *Porphyra umbilicalis*, which lives in the intertidal zone and is exposed to regular desiccation stress, can tolerate more than 95% water loss and resume physiological activities after rehydration (Liu 2009). The present study of the transcriptomic changes in *P. umbilicalis* under desiccation stress can provide clues about the molecular mechanisms for resilience in intertidal algae.

Osmolytes

Organic osmolytes are low-molecular-weight organic solutes used by cells of water-stressed organisms and tissues to maintain cell volume (Yancey 2005). Many different small molecules, such as amino acids and derivatives, polyols and derivatives, sugars, methylamines, methylsulfonium compounds and urea, can function as osmoregulators, protecting subcellular structures and scavenging reactive oxygen species (Slama et al. 2015, Yancey 2005).

Red algae, such as *Pyropia* (Qian et al. 2015), *Galdieria* (Pade et al. 2015) and *Porphyra* (Wiencke and Läuchli 1981) can accumulate photosynthetic carbohydrates, floridoside and isofloridoside, in response to salt stress. However, genes involved in (iso)floridoside pathways are not fully understood in algae. In this study, a putative galactosyltransferase gene was found

to be up-regulated in response to dehydration and rehydration. Barbier et al. (2005) suggested that the galactosyltransferase gene was the most promising candidate gene potentially involved in floridoside biosynthesis in the red alga *Galdieria* as floridoside phosphate synthase (FPS). The increased expression of the putative orthologous galactosyltransferase gene found in this study probably increases floridoside content during dehydration and rehydration in *P. umbilicalis*. The increased expression of the putative orthologous galactosyltransferase gene is partially consistent results in chapter II, which measured the contents of floridoside under dehydration and rehydration in *P. umbilicalis* by Gas chromatography–Mass Spectrometry (GC-MS). In the GC-MS study, the level of floridoside slightly decreased during dehydration and desiccation, but its level increased during rehydration. During rehydration, the up-regulation of the putative galactosyltransferase gene found in this study was consistent with increased levels of floridoside measured in the GC-MS study. However, the up-regulation of the putative orthologous galactosyltransferase gene during dehydration observed in this study does not agree with the decline of floridoside during dehydration in the GC-MS study. Floridoside has been suggested as a direct precursor of polysaccharides in the cell well of *Porphyriduum* sp. (Li et al. 2002). The galactosyl residue in floridoside can be transferred to other molecules (such as polysaccharides), which could explain the inconsistency between increased gene expression of putative galactosyltransferase gene measured in this study and the slight reduction floridoside measured by the previous GC-MS study.

Photosynthesis

Some intertidal macroalgae such as *Pyropia yezoensis* and epiphytic macroalgae *Bostrychia calliptera* (Montagne) Montagne exhibit increased photosynthetic rates in response to dehydration when compared with a fully hydrated state (Gao and Aruga 1987, Peña et al. 1999).

However, this kind of response was not found in *Pyropia haitanensis* (Xu et al. 2016, Zou and Gao 2002). In the current study, the unigene that related to phycobilisomes was up-regulated during dehydration. In the prokaryotic cyanobacteria and eukaryotic red algae, such as *Porphyra purpurea,* light harvesting is carried out primarily by phycobilisomes (Grossman et al. 1993). The up-regulation of the phycobilisome related unigene might suggest increased photosynthetic rate in *Porphyra umbilicalis* during dehydration*,* which is consistent with the previous findings for *Pyropia yezoensis* (Gao and Aruga 1987) and *B. calliptera* (Peña et al. 1999).

During long-term desiccation, the photosynthesis may decrease in red algae, such as in *Pyropia haitanensis* (Xu et al. 2016, Zou and Gao 2002), *Pyropia columbina* (Contreras-Porcia et al. 2010), or be unaffected by desiccation in *Porphyra umbilicalis* (Sampath-Wiley et al. 2008). In the current study, a gene that codes for beta-carotene isomerase, which is involved in the carotene desaturation pathway to form prolycopene in cyanobacteria and plants (Yu et al. 2011), was also up-regulated in response to desiccation. The up-regulation of carotene isomerase during desiccation suggests the possible increase of carotenoid during desiccation. In photosynthesis, carotenoids play an important role in light harvesting and in the protection of cells from photodamage caused by reactive oxygen species (Yu et al. 2011). The up-regulation of the gene that code for beta-carotene isomerase during desiccation probably relates to the function of photoprotection instead of light harvesting because the photosynthetic rate is either decreased or stable during desiccation in red algae.

Two additional genes in carotenoid synthesis pathway, phytoene synthase, and carotenoid cleavage dioxygenase genes, which were up-regulated during rehydration in *P.umbilicalis*. Phytoene synthase is the first enzyme in the carotenoid biosynthetic pathway (Welsch et al. 2000). Iron-binding protein beta-carotene isomerases transform *all-trans*-beta carotene to *9-cis*-

beta carotene and the carotenoid cleavage dioxygenase enzyme functions to cleave 9-cis-beta carotene to the precursor carlactone (Alder et al. 2012, Schwartz et al. 2004). The up-regulated genes that are related to carotenoid synthesis may suggest that the photosynthetic rate increased during rehydration, which is consistent with observations for *Pyropia haitanensis* (Xu et al. 2016). In response to rehydration, *P. haitanensis* recovers the expression of proteins relating to photosynthesis and energy metabolism to restart its normal metabolic activities (Xu et al. 2016). Some carotenoids are precursors of phytohormones such as abscisic acid (Schwartz et al. 1997) and the strigolactone (Umehara et al. 2008), which is discussed below in the phytohormones section.

Signaling

Complex signaling networks in plants, such as hormones, Ca^{2+}, reactive oxygen species (ROS), sugars, etc (Choudhury et al. 2017, Kurusu et al. 2015, Miller et al. 2010) are involved in abiotic stress sensing and metabolic regulation. However, there are limited studies about signaling responses to abiotic stress in algae. Based on the available studies of algae and some model land plants, this discussion will focus on the possible roles of hormones, ROS, and sugar in sensing and signaling during water stress.

Sugars

Trehalose, a non-reducing disaccharide, may accumulate to significant concentrations in bacteria, fungi and certain "resurrection plants", but it is rare in vascular plants (Fernandez et al. 2010, Lunn et al. 2014). The level of trehalose increased in response to environmental stresses and transgenic plants overexpressing trehalose biosynthesis genes increased their stress tolerance, which suggested that trehalose plays an important role in stress response in

Arabidopsis (Iordachescu and Imai 2008). Trehalose or its precursor trehalose-6-phosphate may function as a signaling molecule under osmotic stress in yeast and plants (Argüelles 2000, Elbein et al. 2003). Trehalose has been found in *P. umbilicalis* in nanomolar concentrations by gas chromatography–mass spectrometry (Chapter II). Among five known trehalose biosynthetic pathways, the two-enzymatic pathways that involve trehalose-6-phosphophosphate synthase (TPS) and trehalose-6-phosphate phosphatase (TPP) are the most widely distributed (Avonce et al. 2006).

In this study, no putative genes involved in trehalose biosynthesis were differentially expressed (fold change > 2, $P < 0.05$) during desiccation stress in *P. umbilicalis*. However, the putative gene encoding trehalase, the enzyme that hydrolyzes trehalose to two glucose molecules, was up-regulated during rehydration. High trehalase activity has been suggested as part of the defense mechanisms in terrestrial plants and it is associated with an increase in drought stress tolerance (Van Houtte et al. 2013). The increased expression of the trehalase gene suggested that excessive trehalose was broken down during rehydration in *P. umbilicalis*. It is unclear where excessive trehalose comes from, as no overexpression of trehalose synthesis related genes were found during the water loss process. It is also possible that extra trehalose is synthesized by commensal microorganisms, also impacted by water loss; It has been shown that *P. umbilicalis* harbored microbial communities on the surface and inside its cell wall (Miranda et al. 2013). Trehalose is known to be synthesized by microbes and trehalose synthesis occurs in certain symbiotic interactions (Iturriaga et al. 2009). Another possibility is that *Porphyra* TPS and TPP enzymes are regulated by phosphorylation and dephosphorylation or other post-translational modifications; these are known to enable fast modulation of protein function in response to metabolic and environmental changes (Martins and Swain 2013, Grabsztunowicz et

al. 2017). Another possibility is that trehalose concentration regulation results circatidal clocks that anticipate the drop of tide ; these clocks have been found in crustacea, annelids, molluscs, fish and some insects (Wilcockson and Zhang, 2008). Circatidal clocks have never been reported in algae, however, it is possible that algae can anticiplate the tidal changes and thus prepare for the following water loss. More studies regarding circatidal clocks in algae are needed, to draw any conclusion about this last hypothesis.

Phytohormones and signaling pathways

Hormones released by plants (also known as phytohormones) are active in very low concentrations (10^{-6}to10^{-5}mol/L); they stimulate development and growth and play an important role in signal transduction upon stress responses (Holzinger and Pichrtová 2016). Some common phytohormones, such as abscisic acid (ABA), cytokinin and ethylene, and calcium-dependent serine/threonine-protein kinases, are crucial for modulation of plant responses to abiotic stress (Campo et al. 2014, Holzinger and Becker 2015, Ramanjulu and Bartels 2002, Van de Poel et al. 2016). Phytohormones have also been found in algae, such as the intertidal seaweed species *Pyropia orbicularis* M.E.Ramírez, L.Contreras Porcia & M.-L.Guillemin (Guajardo et al. 2016). Cytokinin signaling, ABA signaling and ethylene response pathways were identified in the genome of the green alga *Hormidiopsis crenulata* (Kützing) Heering (Holzinger and Becker 2015). Signaling pathways for auxin, ABA, cytokinin, salicylic acid, and Jasmonic acid (JA) were found in the draft genome sequence of *Klebsormidium flaccidum* (Kützing) P.C.Silva, K.R.Mattox & W.H.Blackwell (Hori et al. 2014).

Based on the discussion above, the up-regulation of a gene that codes for beta-carotene

isomerase during desiccation could also relate to phytohormone production. Two classes of phytohormones: abscisic acid and strigolactones, are carotenoid derivatives. Among the five characteristic phytohormones, ABA is the most commonly known "stress hormones" that responds to a variety of biotic and abiotic stress (Zhang 2014). Pryce (1972) first suggested that algae and liverworts were distinct from higher plants and did not contain ABA. Later on, ABA was found to be present in various groups of algae, including *Porphyra* (Hartung 2010, Hirsch et al. 1989, Tominaga et al. 1993). However, ABA contents in algal cells (7-34 nmol ABA kg^{-1} FW) are significantly lower than those of unstressed terrestrial plants (Hartung 2010). It is suggested that ABA might not play the role of signaling in response to water stress in algae based on the absence of core components (including group A protein phosphatases type 2C) required for the induction of ABA signaling (Takezawa et al. 2011) and the fact that algae have low sensitivity to exogenous ABA (Komatsu et al. 2013). Even after the discovery of important ABA signaling components (protein phosphatase 2C and SNF1-related protein kinases2 genes) in green alga *Klebsormidium crenulatum* (Holzinger and Becker 2015), it is still unclear if ABA functions in signaling in response to stresses in algae. In *P. umbilicalis,* the increased expression of the unigene that codes for beta-carotene isomerase occured during desiccation, whereas the level of ABA increases after short-term dehydration in higher plants (Xiong and Zhu 2003). If ABA functions as a signal molecule in *Porphyra umbilicalis,* the present study suggests that ABA signaling takes place during prolonged period of water loss. A more likely explanation is that ABA in *Porphyra umbilicalis* may relate to the activation of antioxidant response instead of functioning as a signaling molecule during water stress as suggested for another intertidal seaweed species, *Pyropia orbicularis* (Guajardo et al. 2016).

Another possible signaling pathway for water stress in *Porphyra umbilicalis* may be the mitogen-activated protein kinase (MAPK) signaling pathway. Mitogen-activated protein kinase (MAPK) cascades are important mediators of signal transduction in cells. They convert signals generated at the receptors/sensors to cellular responses (Widmann et al. 1999). The present study showed that a gene encoding MAPK was up-regulated during dehydration, desiccation and rehydration in *P. umbilicalis.* Increasing numbers of studies demonstrate that MAPKs are activated in plants under various stress conditions (Mizoguchi et al. 1997, Hahn and Harter 2009) and thus MAPKs are speculated to control a subset of common defense responses (Zhang and Klessig 2001). The presence and activation of MAPK-like proteins have also been found in six representative species of intertidal macroalgae in response to desiccation or irradiance or both (Parages et al. 2014). It has been shown that the activation of MAPK gene can dramatically increase ethylene production (Hahn and Harter 2009). Ethylene responds to abiotic and biotic stresses and also regulates plant growth and development. In *Pyropia yezoensis,* ethylene was detected and shown to induce enhanced tolerance to oxidative stress (Uji et al. 2016). The *Porphyra umbilicalis* genome also contains genes encoding several proteins involved in ethylene biosynthesis (Brawley et al. 2017). In *P. umbilicalis,* it is possible that ethylene is one of the phytohormones produced in response to water stress.

Reactive oxygen species (ROS)

The accumulation of ROS is a general stress response that can be triggered by osmotic and ionic imbalance in water-stressed cells (Zhang and Bartels 2018). Reactive oxygen stress can be extremely harmful. When the production of ROS exceeds the buffering capacity of the antioxidant pools, ROS may increase membrane permeability, cause lipid peroxidation and enzyme damage, and reduce photosynthetic rates (Shiu and Lee 2005). The antioxidant response,

which includes the activation of antioxidant enzymes and the use of antioxidant compounds, is an important part of desiccation tolerance (Foyer et al. 1997, Noctor and Foyer 1998, Smirnoff 2000). Several studies have demonstrated that abiotic factors cause oxidative stress in algae, such as in *Fucus* spp. (Collén and Davison 1999), in *Mastocarpus stellatus* (Stackhouse) Guiry (Collén and Davison 1999), in *Ulva compressa* Linnaeus (Ratkevicius et al. 2003, Contreras-Porcia et al. 2010), in *Chondrus crispus* Stackhouse (Collén and Davison 1999), and in *Pyropia columbina* (Contreras-Porcia et al. 2010). However, there are some contradictory results regarding the importance of ROS defense in *Porphyra.* Liu (2009) suggested that ROS defense is not the key to the desiccation tolerance in *Porphyra umbilicalis* as the desiccation sensitive *P. yezoensis* had higher ROS scavenging enzyme activity and antioxidant contents than *P. umbilicalis* during desiccation. Similar results were found by Sampath-Wiley et al. (2008). They suggested that irradiance levels and temperatures, instead of desiccation, were the cause of increased antioxidant metabolism in *P. umbilicalis;* increased antioxidant metabolism was also observed during summer but not winter emersions. However, increased ROS levels and increased antioxidant enzymes activity were observed during desiccation, while ROS levels decreased to reduced levels after rehydration in *P. columbina* (Contreras-Porcia et al. 2010). Neither Liu (2009) nor Sampath-Wiley et al. (2008) examined these metabolites during the rehydration process. The transcriptome data in the current study showed no upregulation of genes related to antioxidant enzymes during dehydration and desiccation processes, however, genes with similarity to glutathione-S-transferase (GSTs) and catalases (CAT) were up-regulated during rehydration. Glutathione-S-transferase conjugates GSH to hydrophobic molecules (Rezaei et al. 2013) and CAT dismutates two molecules of H_2O_2 to water and O_2; both function to remove excessive ROS. Interestingly, the induction of ROS scavenger genes was only found

during the rehydration process, suggesting that the overproduction of ROS happens during rehydration in *P. umbilicalis*. The differential expression analysis results agree with oxidative stress levels measured in *P. columbina*: ROS was induced by desiccation and then quickly returned to the basal levels after rehydration (Contreras-Porcia et al. 2010).

Reactive oxygen species (ROS) also plays an integral part as signaling molecules that modifies gene expression and protein activity in response to biotic and/or abiotic stimuli in plants when ROS are within non-toxic levels (Mittler 2002, Baxter et al. 2014). Respiratory bursts oxidase-like proteins play a key role in the network of ROS production and its scavenging pathway (Torres and Dangl 2005). Superoxide, produced via respiratory burst oxidase-like proteins, dismutates to H_2O_2 spontaneously or catalyzed by superoxide dismutase (SOD). Membrane-permeable H_2O_2 is an important signaling molecule involved in stress response (Sagi et al. 2004, Xia et al. 2009). In *Porphyra umbilicalis,* the gene encoding for respiratory burst oxidase-like protein was up-regulated during dehydration and desiccation, which could increase the level of H_2O_2 during the water loss process. The increased H_2O_2 can function as signaling molecule in response to water loss. Two peaks of H_2O_2 were observed in *Pyropia columbina* at 1 and 3 h of desiccation (Contreras-Porcia et al. 2010), which is consistent with the present finding of the up-regulation of gene encodes for respiratory burst oxidase-like protein during dehydration and desiccation.

Transporters

The transfer of solutes is essential for species during desiccation stress. Solute transport in and out of the vacuoles occurs during desiccation and rehydration in desiccation-tolerant angiosperms (Ingram and Bartels 1996). Several transport-associated unigenes were exclusively detected during rehydration in the red alga *Pyropia haitanensis*, such as ABC transporters, ion

transporters, phosphate transporters, and aquaporins (Wang et al. 2015). The current study also identified several transport-associated unigenes in *Porphyra umbilicalis* upregulated during rehydration, which encode ABC transporters, ion transporters, amino acid transporters and folate/biopterin transporter family (Supplementary Table 3-8). The transport-associated unigenes may be involved in the transport machinery of *P. umbilicalis* under rehydration conditions.

ABC transporters play important roles in plant stress tolerance mechanisms and detoxification processes (Kretzschmar et al. 2011). In *Porphyra umbilicalis,* the most common putative transporters are ATP-binding cassette (ABC) superfamily (Chan et al. 2012b). In *Pyropia haitanensis*, genes encoding ABC transporters were downregulated during rehydration. In this study, only one of the genes that encoded ABC transporter was downregulated from fresh to desiccation treatments, while other genes encoding ABC transporters were upregulated during water loss and gain processes. The upregulation of genes for ABC transporter agrees with their detoxification function. In *P. umbilicalis*, ABC transporters were suggested to function in multidrug resistance, as used in bile salt pumps and the transport of lipids (Chan et al. 2012b).

Batch effects

Batch effects were found by principal component analysis plot (Figure 9). The results of PCA plot A revealed that the separation found in PCA at the first component does not correspond to the treatments. Instead, all individuals from rock 1, 3 and 4, plus two individuals from rock 2 and rock 5 were clustered together (group1), while the other samples were clustered (group2). The PCA plot suggested that the samples that clustered together may related to the rocks that samples were collected from. As all rocks are on the similar tidal levels, whether algae grow towards the wave or on the back of the wave could affect the sun exposure and hence possibly lead to different desiccation tolerance in *Porphyra.* The red alga *Porphyra umbilicalis* is

only found to reproduce asexually in the NW Atlantic. The twenty-four individuals used in this study could belong to two major different genotypes that are results of meiosis from the NE Atlantic. It is also possible that algae with different haplotypes may have different desiccation tolerance as group1 possessed one genotype while group2 possessed another genotype. However, more individuals from Fort Stark should be sampled and enhanced sequencing is necessary to support this hypothesis of intraspecific genotypic differences in *P. umbilicalis*.

Dehydration, desiccation and rehydration processes

Among the 250 different expressed genes (DEGs) that were regulated during the water loss process, only 42 DEGs were shared between fresh/dehydration and fresh/desiccation and 10 DEGs were shared between fresh/desiccated and dehydrated/desiccated conditions. A lower number of shared DEGs for dehydration and desiccation stress was also found in plants, such as *Myrothamnus flabellifolia* (Ma et al. 2015)*, Xerophyta viscosa* (Ingle et al. 2007) and *Arabidopsis thaliana* (Buchanan-Wollaston et al. 2005). The complete lack of shared DEGs between from fresh to dehydrated and from dehydrated to desiccated suggested that distinct responses may exist for dehydration and desiccation tolerances in *Porphyra umbilicalis*.

In the present study, several oxidoreductases related genes, such as a polyamine oxidase-like gene, respiratory burst oxidase-like gene, endoplasmic reticulum oxidoreduction 2-like gene, were up-regulated in *Porphyra umbilicalis* during dehydration (supplement Table 3-8). Also, genes involved in the signaling pathway (mitogen-activated protein kinase and serine/threonine protein kinase), ABC transporter system, compatible solutes synthesis pathway (galactosyltransferase) and photosynthesis (phycobilisome) were up-regulated during dehydration in *P. umbilicalis.* However very few biological processes were up-regulated during long term desiccation from dehydration to desiccation. Holzinger and Karsten (2013) suggested

that dehydration tolerance mechanisms involved maintaining homeostasis and actively regulating the water status while desiccation tolerance consisted of strategies to survive the complete loss of water. However, this study suggested that more biological processes were involved during dehydration than simply maintaining homeostasis and regulating the water status in *P. umbilicalis*.

Based on GO analysis, there were several unique biological processes that only occur during water gain, such as signaling, localization, growth, immune system process, detoxification and locomotion. For locomotion, *Porphyra* cells are not known to move. The GO annotation function of locomotion refers to as establishment of localization, such as nitrogen compound transport, nucleobase-containing compound transport, carbohydrate derivative transport, transmembrane transport. For GO term for immune system processes, it possibly relates to the response of alga associated pathogens or endophytes under the water gain process as an innate immune system. The red algal inate immune system has be partially characterized as by Bouarab et al. (2004) who showed that oxylipins play a role in the immunity of a marine red alga *Chondrus crispus*.

A higher number of DEGs were identified among the rehydration process than the water loss process (Figure 10), suggesting rehydration induces more gene expression changes in *Porphyra umbilicalis*. The same pattern of a higher number of induced gene expression changes happened in rehydrated status, like that reported in the highly desiccation-tolerant bryophyte *Tortula ruralis* (Oliver et al. 2004), where the primary response to a desiccation event at the gene expression level occurs during rehydration. Desiccation-tolerant intertidal algae lose water faster than less desiccation-tolerant intertidal algae (Dorgelo 1976). The red alga *P. umbilicalis* starts to loss its internal water content once the tide drops. Besides acquiring desiccation tolerance in

response to a dehydration event, rehydration-induced responses as discussed above also play an important role in the mechanisms of desiccation tolerance in *P. umbilicalis* as suggested by large number of the unigenes differently expressed as well as their discussed functions. This form of desiccation tolerance is considered to the most primitive mechanism of desiccation tolerance (Oliver et al. 2000), which is consistent with the evolutionary history of *P. umbilicalis* as a member of an ancient group of red algae (Brawley et al. 2017).

Limitation and Future direction

The twenty-four individuals used for this study possibly belong to two clones, which might contribute to the observed batch effect in RNA Seq data. In order to understand the desiccation tolerance mechanisms in *P. umbilicalis,* more individuals from different populations with possible different hyplotypes genotyped and investigated by transcriptome analysis.

Differential gene expression studies of abiotic stress in marine alage are relatively recent. Genome information for these species is still limit. In this study, the putative functions of DEG was ascertained primarily by comparing the DEG sequences' similarity to the known function genes in land plants. Further studies of algae genomes response to stress will be required to clarify the physiological functions of DEG.

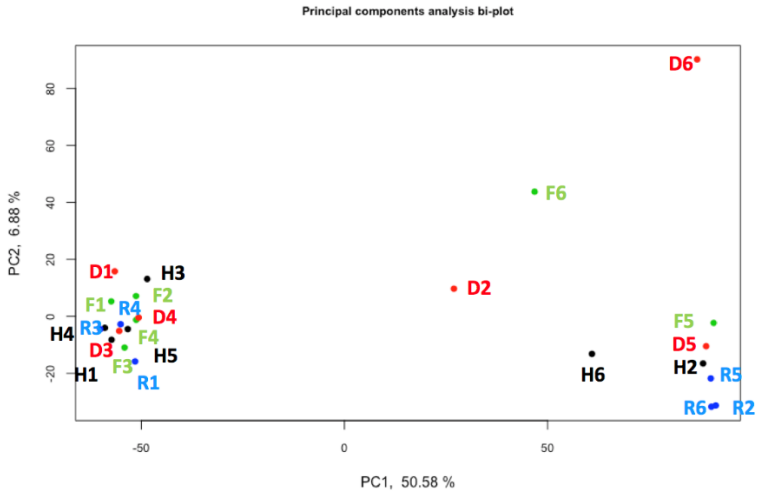
Tables

Table 8 Individual and corresponding sample location information (hydrated control condition (F), dehydration (H), desiccation (D) and rehydration (R))

Rock	individuals
A	F1, H1, D1, R1
B	F2, H2, D2, R2,
C	F3, H3, D3, R3
D	F4, H4, D4, R4
E	F5, H5, D5, R5,
F	F6, H6, D6, R6

Figures

A.



B

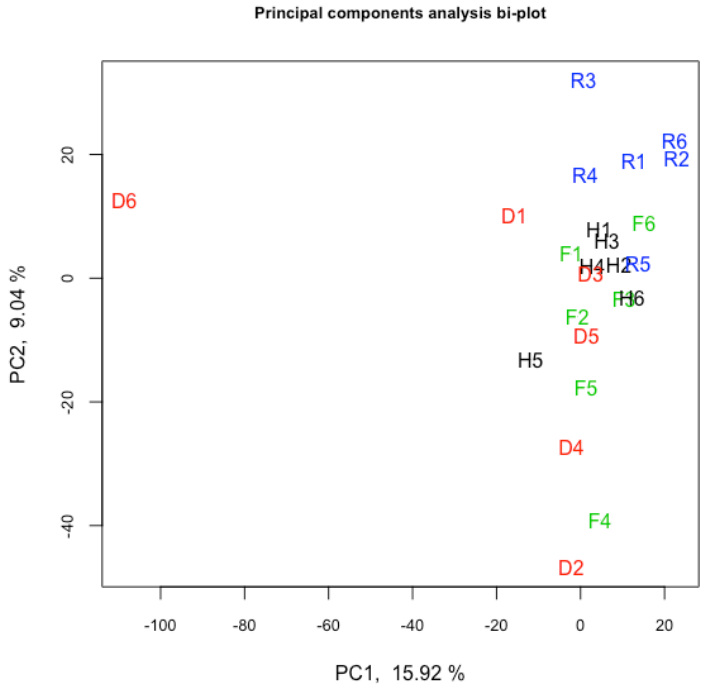


Figure 9 PCA plot of log2-normalized reads of twenty-four RNA-seq samples

In the PCA plot, green, black, red and blue color represent samples under fresh (F), dehydration (H), desiccation (D) and rehydration (R) treatments. Each number responses to each individual. Plot A and B represent PCA before and after batch effect normalization.

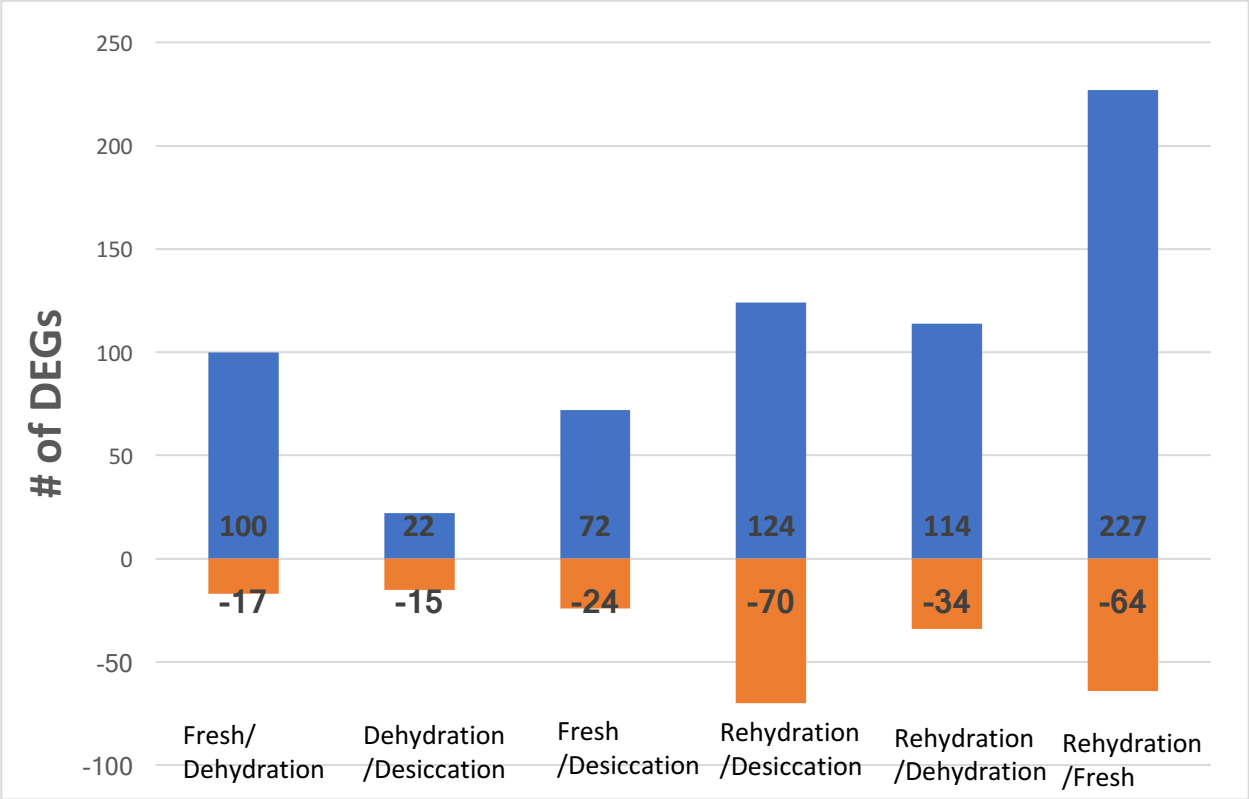
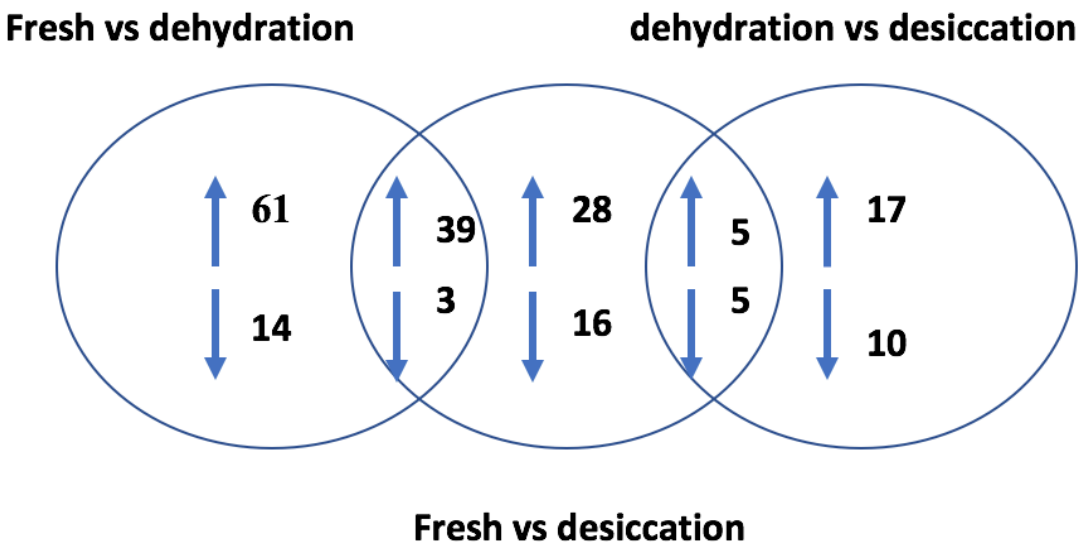


Figure 10 The number of up- and down- regulated genes under dehydration/desiccation stress conditions

The blue bar represents the up-regulated DEGs, while the orange bar represents the down-regulated DEGs

(A)



(B)

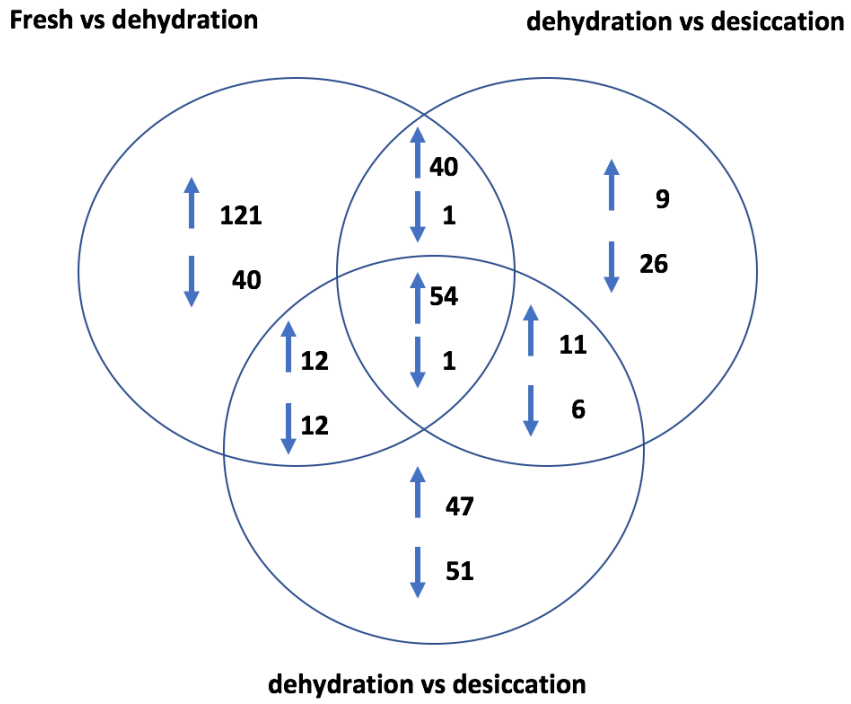


Figure 11 Venn diagram for DEGs in response to (A) water loss and (B) rehydration process

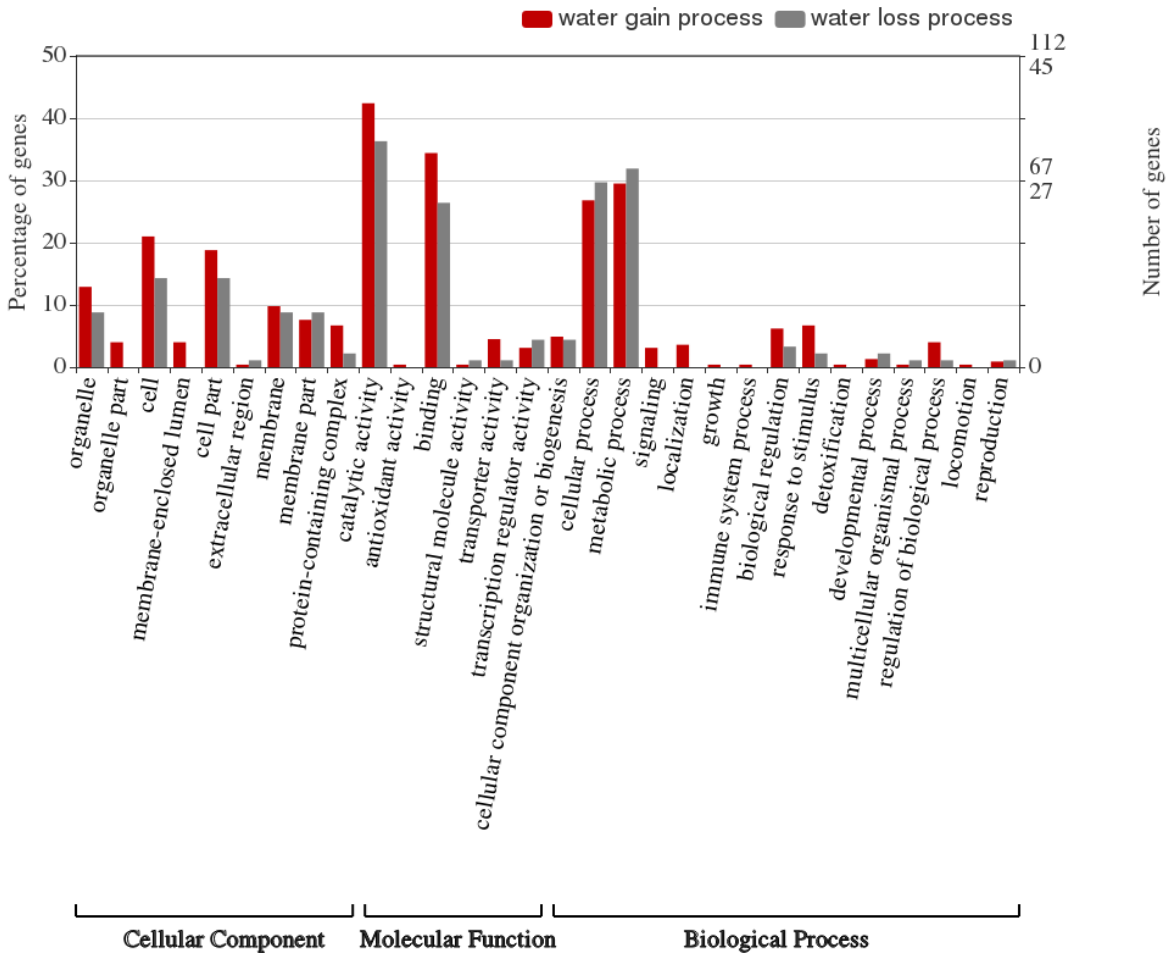


Figure 12 Histogram of the GO annotation analysis of the differential expression genes during water-loss process (grey) and water gain process (red) in *Porphyra umbilicalis* was generated by the Web Gene Ontology Annotation Plot (WEGO) software

Gene were grouped into three GO categories: Cellular component, molecular function, and biological process. The left Y-axis indicates the percentage of genes in a category, the right Y-axis indicates the number of genes in a category.

REFERENCES

Adams P, Nelson DE, Yamada S et al (1998) Growth and development of *Mesembryanthemum crystallinum* (Aizoaceae). New Phytologist 138(2):171-190. doi:10.1046/j.1469-8137.1998.00111.x.

Alder J, Jamil M, Marzorati M et al (2012) The path from beta-Carotene to Carlactone, a Strigolactone-Like Plant Hormone. Science 335(6074):1348-1351.

Amor NB, Ben Hamed K, Debez A et al (2005) Physiological and antioxidant responses of the perennial halophyte *Crithmum maritimum* to salinity. Plant Science 168(4):889-899. doi:10.1016/j.plantsci.2004.11.002.

Anders S, Pyl PT, Huber W (2015) HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics (Oxford, England) 31(2):166-169. doi:10.1093/bioinformatics/btu638.

Argüelles JC (2000) Physiological roles of trehalose in bacteria and yeasts: a comparative analysis. Archives of Microbiology 174(4):217-224. doi:10.1007/s002030000192.

Asamizu E, Nakajima M, Kitade Y et al (2003) Comparison of RNA expression profiles between the two generations of *Porphyra yezoensis* (Rhodophyta), based on expressed sequence tag frequency analysis. Journal of Phycology 39(5):923-930. doi:10.1046/j.1529-8817.2003.03003.x.

Avonce N, Mendoza-Vargas A, Morett E et al (2006) Insights on the evolution of trehalose biosynthesis. BMC Evolutionary Biology 6(1):109. doi:10.1186/1471-2148-6-109.

Baker CJ, Orlandi EW (1995) Active oxygen in plant pathogenesis. Annual Review of Phytopathology 33(1):299-321. doi:10.1146/annurev.phyto.33.1.299.

Baker HG (1965) Characteristics and modes of origins of weeds. In: Baker HG SG (ed) The Genetics of Colonizing Species Academic Press, New York, NY, USA, 147-172.

Barbier G, Oesterhelt C, Larson MD et al (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 Physiology 137(2):460-474. doi:10.1104/pp.104.051169.

Baud S, Dubreucq B, Miquel M et al (2008) Storage reserve accumulation in *Arabidopsis*: metabolic and developmental control of seed filling. The Arabidopsis Book 6:e0113. doi:10.1199/tab.0113.

Baxter A, Mittler R, Suzuki N (2014) ROS as key players in plant stress signalling. Journal of Experimental Botany 65(5):1229-1240. doi:10.1093/jxb/ert375.

Beach KS, Smith CM (1997) Ecophysiology of a tropical rhodophyte III: Recovery from emersion stresses in *Ahnfeltiopsis concinna* (J. Ag.) Silva et DeCew. Journal of Experimental Marine Biology and Ecology 211(2):151-167. doi:10.1016/S0022-0981(96)02721-9.

Bell W, Sun W, Hohmann S et al (1998) Composition and functional analysis of the *Saccaharomyces cerevisiae* trehalose synthase complex. J Biol Chem 273:33311–33319.

Bengtson S, Sallstedt T, Belivanova V et al (2017) Three-dimensional preservation of cellular and subcellular structures suggests 1.6 billion-year-old crown-group red algae. PLoS Biology 15(3):e2000735. doi:10.1371/journal.pbio.2000735.

Bental M, Pick U, Avron M et al (1990) The role of intracellular orthophosphate in triggering osmoregulation in the alga *Dunaliella salina*. European Journal of Biochemistry 188(1):117-122. doi:10.1111/j.1432-1033.1990.tb15378.x.

Bhattacharya D, Price DC, Chan CX et al (2013) Genome of the red alga *Porphyridium purpureum*. Nature Communications 4(1):1941. doi:10.1038/ncomms2931.

Bieger-Dose A, Dose K, Meffert R et al (1992) Extreme dryness and DNA-protein cross-links. Advances in Space Research 12(4):265-270. doi:10.1016/0273-1177(92)90181-V.

Bird CJ, McLachlan JL (1992) Seaweed Flora of the Maritimes: 1. Rhodophyta - The red Algae. Biopress Ltd, England.

Bisson MA, Kirst GO (1995) Osmotic acclimation and turgor pressure regulation in algae. Naturwissenschaften 82(10):461-471. doi:10.1007/BF01131597.

Blázquez MA, Santos E, Flores CL et al (1998) Isolation and molecular characterization of the Arabidopsis TPS1 gene, encoding trehalose-6-phosphate synthase. The Plant Journal: For Cell and Molecular Biology 13(5):685-689. doi:10.1046/j.1365-313X.1998.00063.x.

Bleecker AB, Kende H (2000) Ethylene: A Gaseous Signal Molecule in Plants. Annual Review of Cell and Developmental Biology 16(1):1-18. doi:10.1146/annurev.cellbio.16.1.1.

Blouin N, Brawley S (2012) An AFLP-based test of clonality in widespread, putatively asexual populations of *Porphyra umbilicalis* (Rhodophyta) in the Northwest Atlantic with an in silico analysis for bacterial contamination. Mar Biol 159(12):2723-2729. doi:10.1007/s00227-012-2029-z.

Blouin NA, Brodie JA, Grossman AC et al (2010) *Porphyra*: a marine crop shaped by stress. Trends in Plant Science 16(1):29-37. doi:10.1016/j.tplants.2010.10.004.

Blouin N, Xiugeng F, Peng J et al (2007) Seeding nets with neutral spores of the red alga *Porphyra umbilicalis* (L.) Kützing for use in integrated multi-trophic aquaculture (IMTA). Aquaculture 270(1):77-91. doi:10.1016/j.aquaculture.2007.03.002.

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics (Oxford, England) 30(15):2114-2120. doi:10.1093/bioinformatics/btu170.

Bondu S, Cerantola S, Kervarec N et al (2009) Impact of the salt stress on the photosynthetic carbon flux and 13C-label distribution within floridoside and digeneaside in *Solieria chordalis*. Phytochemistry 70(2):173-184. doi:10.1016/j.phytochem.2008.12.021.

Bouarab K, Adas F, Gaquerel E et al (2004) The innate immunity of a marine red alga involves oxylipins from both the eicosanoid and octadecanoid pathways. Plant Physiology 135(3) 1838-1848.

Brawley SH, Blouin NA, Ficko-Blean E et al (2017) Insights into the red algae and eukaryotic evolution from the genome of *Porphyra umbilicalis* (Bangiophyceae, Rhodophyta). Proceedings of the National Academy of Sciences of the United States of America 114(31):E6361. doi:10.1073/pnas.1703088114.

Brodie JA, Irvine LM (2003) Seaweeds of the British Isles Vol. 1 Rhodophyta Part 3B Bangiophycidae. The Natural History Museum, London, London.

Brodie J, Mols Mortensen A, Ramirez M et al (2008) Making the links: towards a global taxonomy for the red algal genus *Porphyra* (Bangiales, Rhodophyta). J Appl Phycol 20(5):939-949. doi:10.1007/s10811-008-9315-7.

Brown AD (1976) Microbial water stress. Bacteriol Rev 40(4):803–846.

Brown AD (1990) Microbial Water Stress Physiology, Principles and Perspectives. John Wiley & Sons, New York.

Buchanan-Wollaston V, Page T, Harrison E et al (2005) Comparative transcriptome analysis reveals significant differences in gene expression and signalling pathways between developmental and dark/starvation-induced senescence in *Arabidopsis*. The Plant Journal 42(4):567-585. doi:10.1111/j.1365-313X.2005.02399.x.

Burke C, Thomas T, Lewis M et al (2011) Composition, uniqueness and variability of the epiphytic bacterial community of the green alga *Ulva australis*. The ISME Journal 5(4):590-600. doi:10.1038/ismej.2010.164.

Burki F, Kaplan M, Tikhonenkov DV et al (2016) Untangling the early diversification of eukaryotes: a phylogenomic study of the evolutionary origins of Centrohelida, Haptophyta and Cryptista. Proceedings. Biological sciences 283(1823):20152802. doi:10.1098/rspb.2015.2802.

Campo S, Baldrich P, Messeguer J et al (2014) Overexpression of a Calcium-Dependent Protein Kinase Confers Salt and Drought Tolerance in Rice by Preventing Membrane Lipid Peroxidation. Plant Physiology 165(2):688-704. doi:10.1104/pp.113.230268.

Cardon LR, Bell JI (2001) Association study designs for complex diseases. Nature Reviews Genetics 2(2):91-99. doi:10.1038/35052543.

Chan CX, Blouin NA, Zhuang Y et al (2012a) Porphyra (Bangiophyceae) Transcriptomes Provide Insights Into Red Algal Development And Metabolism. Journal of Phycology 48(6):1328-1342. doi:10.1111/j.1529-8817.2012.01229.x.

Chan CX, Zäuner S, Wheeler G et al (2012b) Analysis of Porphyra Membrane Transporters Demonstrates Gene Transfer among Photosynthetic Eukaryotes and Numerous Sodium-Coupled Transport Systems. Plant Physiology 158(4):2001-2012. doi:10.1104/pp.112.193896.

Chan C, Ho C, Phang S (2006) Trends in seaweed research. Trends in Plant Science 11(4):165-166. doi:10.1016/j.tplants.2006.02.003.

Chen J, Song D, Luo Q et al (2014) Determination of Floridoside and Isofloridoside in Red Algae by High-Performance Liquid Chromatography–Tandem Mass Spectrometry. Analytical Letters 47(14):2307-2316. doi:10.1080/00032719.2014.905950.

Chepelev I, Wei G, Tang Q et al (2009) Detection of single nucleotide variations in expressed exons of the human genome using RNASeq. Nucleic Acids Res. 37(16):106.

Choudhury FK, Rivero RM, Blumwald E et al (2017) Reactive oxygen species, abiotic stress and stress combination. The Plant Journal 90(5):856-867. doi:10.1111/tpj.13299.

Cirulli ET, Singh A, Shianna KV et al (2010) Screening the human exome: a comparison of whole genome and whole transcriptome sequencing. Genome Biology 11(5):57. doi:10.1186/gb-2010-11-5-r57.

Cloonan N, Forrest ARR, Kolle G et al (2008) Stem cell transcriptome profiling via massive-scale mRNA sequencing. Nature Methods 5:613-619.

Cole JJ (1982) Interactions Between Bacteria and Algae in Aquatic Ecosystems. Annual Review of Ecology and Systematics 13(1):291-314. doi:10.1146/annurev.es.13.110182.001451.

Cole KM, Sheath RG (1995) Biology of the Red Seaweeds. Cambridge University Press, New York.

Cole KM (2010) Biology of the red algae. Cambridge Univ. Press, Cambridge [u.a.].

Collén J, Davison IR (1999) Reactive oxygen production and damage in intetidal *Fucus spp*. (Phaeophyceae). Journal of Phycology 35:54-61.

Collén J, Porcel B, Carré W et al (2013) Genome structure and metabolic features in the red seaweed *Chondrus crispus* shed light on evolution of the Archaeplastida. Proceedings of the National Academy of Sciences of the United States of America 110(13):5247-5252. doi:10.1073/pnas.1221259110.

Conesa A, Götz S, García-Gómez JM et al (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21(18):3674-3676. doi:10.1093/bioinformatics/bti610.

Conesa A, Madrigal P, Tarazona S et al (2016) A survey of best practices for RNA-seq data analysis. Genome Biology 17(10):13. doi:10.1186/s13059-016-0881-8.

Contreras-Porcia L, Thomas D, Flores V et al (2010) Tolerance to oxidative stress induced by desiccation in *Porphyra columbina* (Bangiales, Rhodophyta). Journal of Experimental Botany 62(6):1815-1829. doi:10.1093/jxb/erq364.

Cowan AK, Rose PD, Horne LG (1992) *Dunaliella salina*: A model System for Studying the Response of Plant Cells to Stress. Journal of Experimental Botany 43(257):1535-1547.

Craigie JS (1974) Storage products. Algal physiology and biochemistry 10: 206-235.

Craigie JS, McLachlan J, Tocher RD (1968) Some neutral constituents of the Rhodophyceae with special reference to the occurrence of the floridosides. Canadian Journal of Botany 46(5):605-611. doi:10.1139/b68-088.

Croft MT, Lawrence AD, Raux-Deery E et al (2005) Algae acquire vitamin B12 through a symbiotic relationship with bacteria. Nature 438(7064):90-93. doi:10.1038/nature04056.

Crowe JH, Carpenter JF, Crowe LM (1998) The role of vitrification in anhydrobiosis. Annual Review of Physiology 60(1):73-103. doi:10.1146/annurev.physiol.60.1.73.

Crowe JH, Crowe LM, Chapman D (1984) Preservation of Membranes in Anhydrobiotic Organisms: The Role of Trehalose. Science 223(4637):701-703. doi:10.1126/science.223.4637.701.

Crowley JJ, Zhabotynsky V, Sun W et al (2015) Analyses of allele-specific gene expression in highly divergent mouse crosses identifies pervasive allelic imbalance. Nature Genetics 47(4):353. doi:10.1038/ng.3222.

Cytryn EJ, Sangurdekar DP, Streeter JG et al (2007) Transcriptional and Physiological Responses of *Bradyrhizobium japonicum* to Desiccation-Induced Stress. Journal of Bacteriology 189(19):6751-6762. doi:10.1128/JB.00533-07.

Daly MA, Mathieson AC (1977) The effects of Sand Movement on Intertidal Seaweeds and Selected Invertebrates at Bound Rock, New Hampshire, USA. Marine Biology, 43: 45-55.

Darriba D, Taboada GL, Doallo R et al (2011) ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics (Oxford, England) 27(8):1164-1165. doi:10.1093/bioinformatics/btr088.

Davison IR, Pearson GA (1996) Stress tolerance in intertidal seaweeds. Journal of Phycology 32(2):197-211.

Day JP, Neefus CD, Yarish C (2009) Development of A Modular Integrated Recirculating Aquaculture System Using Porphyra for Bioremediation of Marine Finfish Effluent. Journal of Phycology 45:31-32.

De Smet KA, Weston A, Brown IN et al (2000) Three pathways for trehalose biosynthesis in mycobacteria. Microbiology 146(1):199.

Deelen P, Zhernakova DV, de Haan M et al (2015) Calling genotypes from public RNA-sequencing data enables identification of genetic variants that affect gene-expression levels. Genome Medicine 7(1):30.

Deighton N, Muckenschnabel I, Goodman BA et al (1999) Lipid peroxidation and the oxidative burst associated with infection of Capsicum annuum by Botrytis cinerea. The Plant Journal 20(4):485-492. doi:10.1046/j.1365-313x.1999.00622.x.

Deng D, Zhao G, Xuan J et al (2004) Construction and characterization of a bacterial artificial chromosome library of marine macroalga *Porphyra yezoensis* (Rhodophyta). Plant Mol Biol Rep 22(4):375-386. doi:10.1007/BF02772680.

Dietz K, Turkan I, Krieger-Liszkay A (2016) Redox- and Reactive Oxygen Species-Dependent Signaling into and out of the Photosynthesizing Chloroplast. Plant physiology 171(3):1541-1550. doi:10.1104/pp.16.00375.

Dobin A, Davis CA, Schlesinger F et al (2013) STAR: ultrafast universal RNA-seq aligner. Bioinformatics (Oxford, England) 29(1):15-21. doi:10.1093/bioinformatics/bts635.

Domozych DS, Ciancia M, Fangel JU et al (2012) The Cell Walls of Green Algae: A Journey through Evolution and Diversity. Frontiers In Plant Xcience 3:82. doi:10.3389/fpls.2012.00082.

Dorgelo J (1976) Intertidal fucoid zonation and desiccation. Hydrobiological Bulletin 10(2):115-122. doi:10.1007/BF02282219.

Drenth A, Tas ICQ, Govers F (1994) DNA fingerprinting uncovers a new sexually reproducing population of Phytophthora infestans in the Netherlands. European Journal of Plant Pathology 100(2):97-107. doi:10.1007/BF01876244.

Drew KM (1949) Conchocelis-Phase in the Life-History of *Porphyra umbilicalis* (L.) Kütz. Nature 164(4174):748-749. doi:10.1038/164748a0.

Eastmond PJ, Graham IA (2003) Trehalose metabolism: a regulatory role for trehalose-6-phosphate?. Current Opinion in Plant Biology 6(3):231-235. doi:10.1016/S1369-5266(03)00037-2.

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research 32(5):1792-1797. doi:10.1093/nar/gkh340.

Einspahr KJ, Peeler TC, Thompson GAJ (1988) Rapid changes in polyphosphoinositide metabolism associated with the response of *Dunaliella salina* to hypoosmotic shock. J. Biol. Chem 263(12):5775-5779.

Elbein AD, Pan YT, Pastuszak I et al (2003) New insights on trehalose: a multifunctional molecule. Glycobiology 13(4):27. doi:10.1093/glycob/cwg047.

Enquist-Newman M, Faust AME, Bravo DD et al (2014) Efficient ethanol production from brown macroalgae sugars by a synthetic yeast platform. Nature 505(7482):239-243. doi:10.1038/nature12771.

Erdmann N, Hagemann M (2001) Salt Acclimation of Algae and Cyanobacteria: A Comparison. In: Rai LC, Gaur JP (eds) Algal Adaptation to Environmental Stresses Springer, Berlin, Heidelberg, p 323-361.

Eriksen RL, Green LA, Klein AS (2016) Genetic variation within and among asexual populations of *Porphyra umbilicalis* Kützing (Bangiales, Rhodophyta) in the Gulf of Maine, USA. Botanica Marina 59(1):1-12. doi:10.1515/bot-2015-0017.

Eriksen RL (2014) Population genetics and organism-environment interactions of *Porphyra umbilicalis* Kützing in the Gulf of Maine, University of New Hampshire.

Eswaran J, Horvath A, Godbole S et al (2013) RNA sequencing of cancer reveals novel splicing alterations. Scientific Reports 3:1689. doi:10.1038/srep01689.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of Molecular Variance Inferred From Metric Distances Among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data. Genetics 131(2):479-491.

Fang J, Zhang J, Xiao T et al (2016) Integrated multi-trophic aquaculture (IMTA) in Sanggou Bay, China. Aquaculture Environment Interactions 8:201-205. doi:10.3354/aei00179.

Farelli JD, Galvin BD, Li Z et al (2014) Structure of the trehalose-6-phosphate phosphatase from Brugia malayi reveals key design principles for anthelmintic drugs. PLoS Pathogens 10(7):e1004245. doi:10.1371/journal.ppat.1004245.

Fei XG, Lu S, Bao Y et al (1998) Seaweed cultivation in China. World Aquaculture 29:22-24.

Fernandez O, Béthencourt L, Quero A et al (2010) Trehalose and plant stress responses: friend or foe?. Trends in Plant Science 15(7):409-417. doi:10.1016/j.tplants.2010.04.004.

Fernandez-Marin B, Holzinger A, Garcia-Plazaola JI (2016) Photosynthetic strategies of desiccation-tolerant organisms. In: Pessarakli M (ed) Handbook of Photosynthesis Boca Raton, FL: CRC Press, 719–737.

Ficko-Blean E, Hervé C, Michel G (2015) Sweet and sour sugars from the sea: the biosynthesis and remodeling of sulfated cell wall polysaccharides from marine macroalgae. Perspectives in Phycology 2(1):51-64. doi:10.1127/pip/2015/0028.

Flint-Garcia SA, Thornsberry JM, Buckler ES (2003) Structure of linkage disequilibria in plants. Annual Review of Plant Biology 54:357-374.

Flores-Molina MR, Thomas D, Lovazzano C et al (2014) Desiccation stress in intertidal seaweeds: Effects on morphology, antioxidant responses and photosynthetic performance. Aquatic Botany 113:90-99. doi:10.1016/j.aquabot.2013.11.004.

Folke C, Kautsky N, Troell M (1994) The cost of eutrophication from salmon farming: implications for policy. Journal of Environmental Management 40:173-182.

Food and Agriculture Organization of the United Nations (2010) Global Forest Resources Assessment 2010 (Rome)

Foyer CH, Lopez-Delgado H, Dat JF et al (1997) Hydrogen peroxide- and glutathione-associated mechanisms of acclimatory stress tolerance and signalling. Physiologia Plantarum 100(2):241-254. doi:10.1034/j.1399-3054.1997.1000205.x.

Fredericq S, Hommersand MH, Freshwater DW (1996) The molecular systematics of some agar- and carrageenan-containing marine red algae based on rbcL sequence analysis. Hydrobiologia 326-327(1):125-135. doi:10.1007/BF00047797.

Fuller CL (2013) Examining Morphological and Physiological Changes in *Zygnema irregulare* During A Desiccation and Recovery Period, CaliforniaState University.

Galinski EA (1995) Osmoadaptation in bacteria. Advances in Microbial Physiology 37:272.

Gantt E, Berg GM, Bhattacharya D et al (2010) *Porphyra*: Complex Life Histories in a Harsh Environment. *P. umbilicalis*, an Intertidal Red Alga for Genomic Analysis. In: Seckbach J, Chapman D (eds) Red Algae in the Genomic Age. Cellular Origin, Life in Extreme Habitats and Astrobiology Springer, New York, p 129-148.

Gao K, Aruga Y (1987) Preliminary studies on the photosynthesis and respiration of *Porphyra yezoensis* under emersed conditions. Journal of the Tokyo University of Fisheries 47(1):51-65.

García-Jiménez P, Robaina RR (2015) On reproduction in red algae: further research needed at the molecular level. Frontiers in Plant Science 6:93. doi:10.3389/fpls.2015.00093.

Gibson RP, Lloyd RM, Charnock SJ et al (2002) Characterization of *Escherichia coli* OtsA, a trehalose-6-phosphate synthase from glycosyltransferase family 20. Acta Crystallographica Section D 58(2):349-351. doi:10.1107/S090744490102056X.

Gill SS, Tuteja N (2010) Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. Plant Physiology and Biochemistry 48(12):909-930. doi:10.1016/j.plaphy.2010.08.016.

Goddijn OJM, Van DK (1999) Trehalose metabolism in plants. Trends in Plant Science 4:315 – 319.

Goddijn O, Verwoerd TC, Voogd E et al (1997) Inhibition of Trehalase Activity Enhances Trehalose Accumulation in Transgenic Plants. Plant Physiology 113(1):181-190. doi:10.1104/pp.113.1.181.

Goecke F, Labes A, Wiese J et al (2010) Chemical interactions between marine macroalgae and bacteria. Marine Ecology Progress Series 409:267-299. doi:10.3354/meps08607.

Good BH, Rouzine IM, Balick DJ et al (2012) Distribution of fixed beneficial mutations and the rate of adaptation in asexual populations. Proceedings of the National Academy of Sciences of the United States of America 109(13):4950-4955. doi:10.1073/pnas.1119910109.

Götz S, García-Gómez JM, Terol J et al (2008) High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Research 36(10):3420-3435. doi:10.1093/nar/gkn176.

Goya R, Sun MGF, Morin RD et al (2010) SNVMix: predicting single nucleotide variants from next-generation sequencing of tumors. Bioinformatics (Oxford, England) 26(6):730-736. doi:10.1093/bioinformatics/btq040.

Grabherr MG, Brian J Haas, Yassour M et al (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature Biotechnology 29(7):644-652. doi:10.1038/nbt.1883.

Grabsztunowicz M, Koskela MM, Mulo P (2017) Post-translational Modifications in Regulation of Chloroplast Function: Recent Advances. Frontiers in plant science 8:240. doi:10.3389/fpls.2017.00240.

Graham LE, Graham JM, Wilcox LE et al (2016) Algae, Third Edition edn. LJLM Press, Madison (WI).

Grossman AR, Schaefer MR, Chiang GG et al (1993) The phycobilisome, a light-harvesting complex responsive to environmental conditions. Microbiological Reviews 57(3):725-749.

Guajardo E, Correa J, Contreras-Porcia L (2016) Role of abscisic acid (ABA) in activating antioxidant tolerance responses to desiccation stress in intertidal seaweed species. Planta 243(3):767-781. doi:10.1007/s00425-015-2438-6.

Guillemin M, Faugeron S, Destombe C et al (2008) Genetic Variation in Wild and Cultivated Populations of the Haploid– Diploid Red Alga *Gracilaria chilensis*: How Farming Practices Favor Asexual Reproduction and Heterozygosity. Evolution 62(6):1500-1519. doi:10.1111/j.1558-5646.2008.00373.x.

Gupta PK, Rustgi S, Mir RR (2008) Array-based high-throughput DNA markers for crop improvement. Heredity 101(1):5-18. doi:10.1038/hdy.2008.35.

Ha KS, Thompson J, G A (1992) Biphasic changes in the level and composition of *Dunaliella salina* plasma membrane diacylglycerols following hypoosmotic shock. Biochemistry 31(2):596-603. doi:10.1021/bi00117a042.

Haas AF, Nelson CE, Kelly LW et al (2011) Effects of Coral Reef Benthic Primary Producers on Dissolved Organic Carbon and Microbial Activity. PLoS One 6(11):e27973. doi:10.1371/journal.pone.0027973.

Hagemann M (2016) Coping with High and Variable Salinity: Molecular Aspects of Compatible Solute Accumulation. In: Borowitzka M, Beardall J, Raven J (eds) The Physiology of Microalgae. Developments in Applied Phycology Springer, Cham.

Hagemann M, Pade N (2015) Heterosides – compatible solutes occurring in prokaryotic and eukaryotic phototrophs. Plant Biology banner 17(5):927-934. doi:10.1111/plb.12350.

Hahn A, Harter K (2009) Mitogen-Activated Protein Kinase Cascades and Ethylene: Signaling, Biosynthesis, or Both?. Plant Physiology 149(3):1207-1210. doi:10.1104/pp.108.132241.

Harker M, Berkaloff C, Lemoine Y et al (1999) Effects of high light and desiccation on the operation of the xanthophyll cycle in two marine brown algae. European Journal of Phycology 34(1):35-42. doi:10.1017/S0967026299001924.

Hartung W (2010) The evolution of abscisic acid (ABA) and ABA function in lower plants, fungi and lichen. Functional Plant Biology 37(9):806. doi:10.1071/FP10058.

Hawes I (1990) Effects of freezing and thawing on a species of *Zygnema* (Chlorophyta) from the Antarctic. Phycologia 29(3):326-331. doi:10.2216/i0031-8884-29-3-326.1.

Heo Y, Wu X, Chen D et al (2014) BLESS: Bloom filter-based error correction solution for high-throughput sequencing reads. Bioinformatics 30(10):1354-1362. doi:10.1093/bioinformatics/btu030.

Herburger K, Holzinger A (2015) Localization and Quantification of Callose in the Streptophyte Green Algae *Zygnema* and *Klebsormidium*: Correlation with Desiccation Tolerance. Plant & cell Physiology 56(11):2259.

Hirasawa T, Nakakura Y, Yoshikawa K et al (2006) Comparative analysis of transcriptional responses to saline stress in the laboratory and brewing strains of *Saccharomyces cerevisiae* with DNA microarray. Appl Microbiol Biotechnol 70(3):346-357. doi:10.1007/s00253-005-0192-6.

Hirsch R, Hartung W, Gimmler H (1989) Abscisic Acid Content of Algae under Stress. Botanica Acta 102(4):326-334. doi:10.1111/j.1438-8677.1989.tb00113.x.

Hoffmann AJ, Camus P (1989) Sinking rates and viability of spores from benthic algae in central Chile. Journal of Experimental Marine Biology and Ecology 126(3):281-291. doi:10.1016/0022-0981(89)90193-7.

Holdt S, Kraan S (2011) Bioactive compounds in seaweed: functional food applications and legislation. J Appl Phycol 23(3):543-597. doi:10.1007/s10811-010-9632-5.

Holzinger A, Becker B (2015) Desiccation tolerance in the streptophyte green alga *Klebsormidium*: The role of phytohormones. Communicative & Integrative Biology 8(4):e1059978. doi:10.1080/19420889.2015.1059978.

Holzinger A, Kaplan F, Blaas K et al (2014) Transcriptomics of Desiccation Tolerance in the Streptophyte Green Alga *Klebsormidium* Reveal a Land Plant-Like Defense Reaction. PLoS One 9(10):e110630. doi:10.1371/journal.pone.0110630.

Holzinger A, Karsten U (2013) Desiccation stress and tolerance in green algae: consequences for ultrastructure, physiological and molecular mechanisms. Frontiers in Plant Science 4:327. doi:10.3389/fpls.2013.00327.

Holzinger A, Pichrtová M (2016) Abiotic Stress Tolerance of Charophyte Green Algae: New Challenges for Omics Techniques. Frontiers in Plant Science 7:678. doi:10.3389/fpls.2016.00678.

Hori K, Maruyama F, Fujisawa T et al (2014) *Klebsormidium flaccidum* genome reveals primary factors for plant terrestrial adaptation. Nature Communications 5(1):3978. doi:10.1038/ncomms4978.

Hurd CL, Dring MJ (1991) Desiccation and phosphate uptake by intertidal fucoid algae in relation to zonation. British Phycological Journal 26(4):327-333. doi:10.1080/00071619100650291.

Huson DH, Scornavacca C (2012) Dendroscope 3: An Interactive Tool for Rooted Phylogenetic Trees and Networks. Systematic Biology 61(6):1061-1067. doi:10.1093/sysbio/sys062.

Im S, Lee H, Jung HS et al (2017) Transcriptome-Based Identification of the Desiccation Response Genes in Marine Red Algae *Pyropia tenera* (Rhodophyta) and Enhancement of Abiotic Stress Tolerance by PtDRG2 in Chlamydomonas. Marine Biotechnology 19(3):232-245. doi:10.1007/s10126-017-9744-x.

Ingle RA, Schmidt UG, Farrant JM et al (2007) Proteomic analysis of leaf proteins during dehydration of the resurrection plant *Xerophyta viscosa*. Plant, Cell & Environment 30(4):435-446. doi:10.1111/j.1365-3040.2006.01631.x.

Ingolfsson A (1992) The Origin of the Rocky Shore Fauna of Iceland and the Canadian Maritimes. Journal of Biogeography 19(6):705-712. doi:10.2307/2845711.

Ingram J, Bartels D (1996) The Molecular Basis Of Dehydration Tolerance In Plants. Annual Review of Plant Physiology and Plant Molecular Biology 47(1):377-403. doi:10.1146/annurev.arplant.47.1.377.

Inupakutika MA, Sengupta S, Devireddy AR et al (2016) The evolution of reactive oxygen species metabolism. Journal of Experimental Botany 67(21):5933-5943. doi:10.1093/jxb/erw382.

Iordachescu M, Imai R (2008) Trehalose Biosynthesis in Response to Abiotic Stresses. Journal of Integrative Plant Biology 50(10):1223-1229. doi:10.1111/j.1744-7909.2008.00736.x.

Iturriaga G, Suárez R, Nova-Franco B (2009) Trehalose metabolism: from osmoprotection to signaling. International journal of molecular sciences 10(9):3793-3810. doi:10.3390/ijms10093793.

Jahns P, Holzwarth AR (2012) The role of the xanthophyll cycle and of lutein in photoprotection of photosystem II. Biochimica et Biophysica Acta 1817(1):182.

Jain NK, Roy I (2009) Effect of trehalose on protein structure. Protein science : a publication of the Protein Society 18(1):NA. doi:10.1002/pro.3.

Jiang W, Fu F, Zhang S et al (2010) Cloning and Characterization of Functional Trehalose-6-Phosphate Synthase Gene in Maize. Journal of Plant Biology 53(2):134-141. doi:10.1007/s12374-010-9098-7.

Johnson WS, Gigon A, Gulmon SL et al (1974) Comparative Photosynthetic Capacities of Intertidal Algae under Exposed and Submerged Conditions. Ecology 55(2):450-453. doi:10.2307/1935235.

Ju C, Van de Poel B, Cooper ED et al (2015) Conservation of ethylene as a plant hormone over 450 million years of evolution. Nature Plants 1(1):14004. doi:10.1038/nplants.2014.4.

Junior NN, Pereira Nicolau MS, Mantovanini LJ et al (2013) Expression Analysis of Two Genes Coding for Trehalose-6-Phosphate Synthase (TPS), in Sugarcane (*Saccharum spp.*) under Water Stress. American Journal of Plant Sciences 4(12):91-99. doi:10.4236/ajps.2013.412A3011.

Kaplan F, Lewis L, Wastian J et al (2012) Plasmolysis effects and osmotic potential of two phylogenetically distinct alpine strains of *Klebsormidium* (Streptophyta). Protoplasma 249(3):789-804. doi:10.1007/s00709-011-0324-z.

Kapraun DF, Lemus AJ (1987) Field and Culture Studies of P*orphyra spirali*s var. amplifolia Oliveira Filho et Coll (Bangiales, Rhodophyta) from Isla de Margarita, Venezuela. Botanica Marina 30(6):483. doi:10.1515/botm.1987.30.6.483.

Karsten U (1999) Seasonal Variation in Heteroside Concentrations of Field-Collected *Porphyra* Species (Rhodophyta) from Different Biogeographic Regions. The New Phytologist 143(3):561-571. doi:10.1046/j.1469-8137.1999.00477.x.

Karsten U, Barrow KD, King RJ (1993) Floridoside, L-Isofloridoside, and D-Isofloridoside in the Red Alga *Porphyra columbina*: Seasonal and Osmotic Effects. Plant Physiology 103(2):485-491.

Karsten U, Görs S, Eggert A et al (2007) Trehalose, digeneaside, and floridoside in the Florideophyceae (Rhodophyta) – a reevaluation of its chemotaxonomic value. Phycologia 46(2):143-150. doi:10.2216/06-29.1.

Karsten U, Michalik D, Michalik M et al (2005) A new unusual low molecular weight carbohydrate in the red algal genus *Hypoglossum* (Delesseriaceae, Ceramiales) and its possible function as an osmolyte. Planta 222(2):319-326. doi:10.1007/s00425-005-1527-3.

Kaushik JK, Bhat R (2003) Why is trehalose an exceptional protein stabilizer? An analysis of the thermal stability of proteins in the presence of the compatible osmolyte trehalose. The Journal of Biological Chemistry 278(29):26458-26465. doi:10.1074/jbc.M300815200.

Kim JK, Kraemer GP, Yarish C (2008) Physiological activity of *Porphyra* in relation to eulittoral zonation. Journal of Experimental Marine Biology and Ecology 365(2):75-85. doi:10.1016/j.jembe.2008.07.040.

Kim JK, Kraemer GP, Yarish C (2009) Research note: Comparison of growth and nitrate uptake by New England Porphyra species from different tidal elevations in relation to desiccation. Phycological Research 57(2):152. doi:10.1111/j.1440-1835.2009.00533.x.

Kim S, Kim D, Cho SW et al (2014) Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. Genome Research 24(6):1012. doi:10.1101/gr.171322.113.

Kirst GO (1980) Low molecular weight carbohydrates and ions in Rhodophyceae: quantitative measurement of floridoside and digeneaside. Phytochemistry 19:107-1110.

Kirst GO (1990) Salinity Tolerance of Eukaryotic Maine Algae. Annual Review of Plant Physiology and Plant Molecular Biology 41:21-53.

Klein AS, Mathieson AC, Neefus CD et al (2003) Identification of north-western Atlantic *Porphyra* (Bangiaceae, Bangiales) based on sequence variation in nuclear SSU and plastid rbcL genes. Phycologia 42(2):109.

Kojima M, Kamada-Nobusada T, Komatsu H et al (2009) Highly sensitive and high-throughput analysis of plant hormones using MS-probe modification and liquid chromatography-tandem mass spectrometry: an application for hormone profiling in *Oryza sativa*. Plant and Cell Physiology 50(7):1201-1214.

Komatsu K, Suzuki N, Kuwamura M et al (2013) Group A PP2Cs evolved in land plants as key regulators of intrinsic desiccation tolerance. Nature Communications 4(1):2219. doi:10.1038/ncomms3219.

Kong F, Cao M, Sun P et al (2015) Selection of reference genes for gene expression normalization in *Pyropia yezoensis* using quantitative real-time PCR. J Appl Phycol 27(2):1003-1010. doi:10.1007/s10811-014-0359-6.

Kremer BP, Vogl R (1975) Zur chemotaxonomischen Bedeutung des [14C] Markierungs-musters bei Rhodophyceen. Phytochemistry 14:1309–1314.

Kretzschmar T, Burla B, Lee Y et al (2011) Functions of ABC transporters in plants. Essays in Biochemistry 50(1):145-160. doi:10.1042/bse0500145.

Krom MD, Ellner S, Rijn Jv et al (1995) Nitrogen and phosphorus cycling and transformations in a prototype 'non-polluting' integrated mariculture system, Eilat, Israel. Marine Ecology Progress Series 118(1/3):25-36. doi:10.3354/meps118025.

Kumari A, Jewaria PK, Bergmann DC et al (2014) Arabidopsis reduces growth under osmotic stress by decreasing SPEECHLESS protein. Plant & Cell Physiology 55(12):2037-2046. doi:10.1093/pcp/pcu159.

Kurusu T, Kuchitsu K, Tada Y (2015) Plant signaling networks involving Ca2+ and Rboh/Nox-mediated ROS production under salinity stress. Frontiers in Plant Science 6:427. doi:10.3389/fpls.2015.00427.

Kwok P (2001) Methods for genotyping signle nucleotide polymorphisms. Annual Review of Genomics and Human Genetics 2(1):235-258. doi:10.1146/annurev.genom.2.1.235.

Lachnit T, Meske D, Wahl M et al (2011) Epibacterial community patterns on marine macroalgae are host-specific but temporally variable. Environmental Microbiology 13(3):655-665. doi:10.1111/j.1462-2920.2010.02371.x.

Lang GI, Botstein D, Desai MM (2011) Genetic variation and the fate of beneficial mutations in asexual populations. Genetics 188(3):647-661. doi:10.1534/genetics.111.128942.

Larkum AWD, Barrett J (1983) Light-harvesting Processes in Algae. In: Anonymous Advances in Botanical Research, vol 10. Elsevier Science & Technology, p 1-219.

Latifi A, Ruiz M, Zhang C (2009) Oxidative stress in cyanobacteria. FEMS microbiology Reviews 33(2):258-278. doi:10.1111/j.1574-6976.2008.00134.x.

Lee J, Yang EC, Graf L et al (2018) Analysis of the Draft Genome of the Red Seaweed *Gracilariopsis chorda* Provides Insights into Genome Size Evolution in Rhodophyta. Molecular Biology and Evolution 35(8):1869-1886. doi:10.1093/molbev/msy081.

Leek JT, Johnson WE, Parker HS et al (2012) The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics (Oxford, England) 28(6):882-883. doi:10.1093/bioinformatics/bts034.

Leyman B, Van Dijck P, Thevelein JM (2001) An unexpected plethora of trehalose biosynthesis genes in *Arabidopsis thaliana*. Trends in Plant Science 6(11):510-513. doi:10.1016/S1360-1385(01)02125-2.

Li B, Chen C, Xu Y et al (2014) Validation of housekeeping genes as internal controls for studying the gene expression in *Pyropia haitanensis* (Bangiales, Rhodophyta) by quantitative real-time PCR. Acta Oceanol Sin 33(9):152-159. doi:10.1007/s13131-014-0526-2.

Li H, Durbin R (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26(5):589-595. doi:10.1093/bioinformatics/btp698.

Li H, Handsaker B, Wysoker A et al (2009) The Sequence Alignment Map format and SAMtools. Bioinformatics 25(16):2078-2079. doi:10.1093/bioinformatics/btp352.

Li Q, Liu J, Zhang L et al (2014) De Novo Transcriptome Analysis of an Aerial Microalga *Trentepohlia jolithus*: Pathway Description and Gene Discovery for Carbon Fixation and Carotenoid Biosynthesis. PLoS One 9(9). doi:10.1371/journal.pone.0108488.

Li R, Li Y, Fang X et al (2009) SNP detection for massively parallel whole-genome resequencing. Genome Research 19(6):1124.

Li S, Lellouche J, Shabtai Y et al (2001) Fixed carbon partitioning in the red microalga *Porphyridium* (Rhodophyta). Journal of Phycology 37(2):289-297. doi:10.1046/j.1529-8817.2001.037002289.x.

Li S, Shabtai Y, Arad S (2002) Floridoside as a carbon precursor for the synthesis of cell-wall polysaccharide in the red microalga *Porphyridium* sp (Rhodophyta). Journal of Phycology 38(5):931-938. doi:10.1046/j.1529-8817.2002.01143.x.

Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22(13):1658-1659. doi:10.1093/bioinformatics/btl158.

Li Z, Wakao S, Fischer BB et al (2009) Sensing and Responding to Excess Light. Annual Review of Plant Biology 60(1):239-260. doi:10.1146/annurev.arplant.58.032806.103844.

Lipkin Y, Beer S, Eshel A (1993) The Ability of Porphyra linearis (Rhodophyta) to Tolerate Prolonged Periods of Desiccation. Botanica Marina 36(6):517. doi:10.1515/botm.1993.36.6.517.

Lippert K, Galinski E (1992) Enzyme stabilization be ectoine-type compatible solutes: protection against heating, freezing and drying. Applied Microbiology and Biotechnology 37(1). doi:10.1007/BF00174204.

Liu K, Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21(9):2128-2129. doi:10.1093/bioinformatics/bti282.

Liu Y (2009) Mechanism for differential desiccation tolerance in *Porphyra* species, Northeastern University

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods (San Diego, Calif.) 25(4):402.

Long S, Humphries S, Falkowski PG (1994) Photoinhibition of Photosynthesis in Nature. Annual Review of Plant Physiology and Plant Molecular Biology 45(1):633-661. doi:10.1146/annurev.arplant.45.1.633.

Longford S, Tujula N, Crocetti G et al (2007) Comparisons of diversity of bacterial communities associated with three sessile marine eukaryotes. Aquatic Microbial Ecology 48:217-229. doi:10.3354/ame048217.

Lopez-Maestre H, Brinza L, Marchet C et al (2016) SNP calling from RNA-seq data without a reference genome: identification, quantification, differential analysis and impact on the protein sequence. Nucleic Acids Research:gkw655. doi:10.1093/nar/gkw655.

Lüning K (1990) Seaweeds. Their environment, biogeography and ecophysiology.John Wiley & Son

Lunn JE, Delorge I, Figueroa CM et al (2014) Trehalose metabolism in plants. The Plant Journal 79(4):544-567. doi:10.1111/tpj.12509.

Lütz C, Seidlitz H, Meindl U (1997) Physiological and structural changes in the chloroplast of the green alga *Micrasterias denticulata* induced by UV-B simulation. Plant Ecology 128(1):55-64. doi:1009754722357.

Lynch M, Ackerman MS, Gout J et al (2016) Genetic drift, selection and the evolution of the mutation rate. Nature Reviews Genetics 17(11):704-714. doi:10.1038/nrg.2016.104.

Ma C, Wang H, Macnish AJ et al (2015) Transcriptomic analysis reveals numerous diverse protein kinases and transcription factors involved in desiccation tolerance in the resurrection plant *Myrothamnus flabellifolia*. Horticulture Research 2(1):15034. doi:10.1038/hortres.2015.34.

MacColl R (1998) Cyanobacterial Phycobilisomes. Journal of Structural Biology 124(2-3):311-334. doi:10.1006/jsbi.1998.4062.

Macmanes MD (2014) On the optimal trimming of high-throughput mRNA sequence data. Frontiers in Genetics 5:13. doi:10.3389/fgene.2014.00013.

Magazù S, Migliardo F, Mondelli C et al (2005) Correlation between bioprotective effectiveness and dynamic properties of trehalose–water, maltose–water and sucrose–water mixtures. Carbohydrate Research 340(18):2796-2801. doi:10.1016/j.carres.2005.09.026.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. Cancer Res. 27:209-220.

Markert JA, Champlin DM, Gutjahr-Gobell R et al (2010) Population genetic diversity and fitness in multiple environments. BMC Evolutionary Biology 10(1):205. doi:10.1186/1471-2148-10-205.

Martinez-Garcia M, Van der Maarel, M. J. E. C (2016) Floridoside production by the red microalga *Galdieria sulphuraria* under different conditions of growth and osmotic stress. AMB Express 6(1):1. doi:10.1186/s13568-016-0244-6.

Martins BMC, Swain PS (2013) Ultrasensitivity in phosphorylation-dephosphorylation cycles with little substrate. PLoS Computational Biology 9(8):e1003175. doi:10.1371/journal.pcbi.1003175.

Mathieson AC, Burns RL (1971) Ecological studies of economic red algae. I. Photosynthesis and respiration of *Chondrus crispus* Stackhouse and *Gigartina stellata* (Stackhouse) Batters. Journal of Experimental Marine Biology and Ecology 7(2):197-206. doi:10.1016/0022-0981(71)90031-1.

Matsuzaki M, Misumi O, Shin-I T et al (2004) Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. Nature 428(6983):653-657. doi:10.1038/nature02398.

McKenna A, Hanna M, Banks E et al (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research 20(9):1297-1303. doi:10.1101/gr.107524.110.

Meng J, Rosell K, Srivastava LM (1987) Chemical characterization of floridosides from *Porphyra perforata*. Carbohydrate Research 161(2):171-180. doi:10.1016/S0008-6215(00)90075-2.

Meng J, Srivastava LM (1993) Variations in floridoside content and floridoside phosphate synthase activity in *Porphyria perforate* (Rhodophyta). Journal of Phycology 29(1):82-84. doi:10.1111/j.1529-8817.1993.tb00283.x.

Mikkat S, Hagemann M, Schoor A (1996) Active transport of glucosylglycerol is involved in salt adaptation of the cyanobacterium *Synechocystis* sp. strain PCC 6803. Microbiology 142(7):1725-1732. doi:10.1099/13500872-142-7-1725.

Miller G, Suzuki N, Ciftci-Yilmaz S et al (2010) Reactive oxygen species homeostasis and signalling during drought and salinity stresses. Plant, cell & environment 33(4):453-467. doi:10.1111/j.1365-3040.2009.02041.x.

Milner JM, Albon SD, Illius AW et al (1999) Repeated Selection of Morphometric Traits in the Soay Sheep on St Kilda. Journal of Animal Ecology 68(3):472-488. doi:10.1046/j.1365-2656.1999.00299.x.

Miranda H (2011) Stress response in the cyanobacterium *Synechocystis* sp. PCC 6803. Umeå University, Sweden.

Miranda LN, Hutchison K, Grossman AR et al (2013) Diversity and abundance of the bacterial community of the red Macroalga *Porphyra umbilicalis*: did bacterial farmers produce macroalgae?. PloS one 8(3):e58269. doi:10.1371/journal.pone.0058269.

Mishra A, Mishra KB, Höermiller II et al (2011) Chlorophyll fluorescence emission as a reporter on cold tolerance in *Arabidopsis thaliana* accessions. Plant Signaling & Behavior 6(2):301-310. doi:10.4161/psb.6.2.15278.

Mittler R (2002) Oxidative stress, antioxidants and stress tolerance. Trends Plant Sci 7(9):405-410.

Mizoguchi T, Ichimura K, Shinozaki K (1997) Environmental stress response in plants: the role of mitogen-activated protein kinases. Trends in Biotechnology 15(1):15-19. doi:10.1016/S0167-7799(96)10074-3.

Mouritsen OG (2013) Seaweeds: edible, available & sustainable. The University of Chicago Press, Chicago.

Munns R (2002) Comparative physiology of salt and water stress. Plant, Cell & Environment 25(2):239-250. doi:10.1046/j.0016-8025.2001.00808.x.

Nei M (1972) Genetic distance between populations. American Naturalist 106:283-292.

Nakamura Y, Sasaki N, Kobayashi M et al (2013) The First Symbiont-Free Genome Sequence of Marine Red Alga, Susabi-nori (*Pyropia yezoensis*). PLoS One 8(3):e57122. doi:10.1371/journal.pone.0057122.

Nakitandwe J, Trognitz F, Trognitz B (2007) Reliable allele detection using SNP-based PCR primers containing Locked Nucleic Acid: application in genetic mapping. Plant Methods 3(1):2. doi:10.1186/1746-4811-3-2.

Nascimento I, Marques S, Cabanelas I et al (2013) Screening Microalgae Strains for Biodiesel Production: Lipid Productivity and Estimation of Fuel Quality Based on Fatty Acids Profiles as Selective Criteria. Bioenerg Res 6(1):1-13. doi:10.1007/s12155-012-9222-2.

Neale PJ (1987) Algal photoinhibition and photosynthesis in the aquatic environment. In: Kyle DJ, Osmond CB, Arntzen CJ (eds) Photoinhibition Elsevier, Amsterdam, p 35–65.

Nepomuceno AL, Oosterhuis D, Stewart JM et al (2002) Expression of heat shock protein and trehalose-6-phosphate synthase homologues induced during water deficit in cotton. Brazilian Journal of Plant Physiology 14(1):11-20.

Nikaido I, Asamizu E, Nakajima M et al (2000) Generation of 10,154 expressed sequence tags from a leafy gametophyte of a marine red alga, *Porphyra yezoensis*. DNA research: An International Journal for Rapid Publication of Reports on Genes and Genomes 7(3):223-227. doi:10.1093/dnares/7.3.223.

Niu X, Narasimhan ML, Salzman RA et al (1993) NaCl Regulation of Plasma Membrane H+-ATPase Gene Expression in a Glycophyte and a Halophyte. Plant Physiology 103(3):713-718. doi:10.1104/pp.103.3.713.

Niwa K, Kikuchi N, Iwabuchi M et al (2004) Morphological and AFLP variation of *Porphyra yezoensis* Ueda form, narawaensis Miura (Bangiales, Rhodophyta). Phycological Research 52(2):180-190. doi:10.1111/j.1440-1835.2004.tb00327.x.

Noctor G, Foyer CH (1998) Ascorbate and glutathione: Keeping active oxygen under control. Annual Review of Plant Physiology and Plant Molecular Biology 49(1):249-279. doi:10.1146/annurev.arplant.49.1.249.

Olie JJ, Potts M (1986) Purification and Biochemical Analysis of the Cytoplasmic Membrane from the Desiccation-Tolerant Cyanobacterium *Nostoc commune* UTEX 584. Applied and Environmental Microbiology 52(4):706-710.

Oliveira LSd, Gregoracci GB, Silva GGZ et al (2012) Transcriptomic analysis of the red seaweed *Laurencia dendroidea* (Florideophyceae, Rhodophyta) and its microbiome. BMC Genomics 13(1):487. doi:10.1186/1471-2164-13-487.

Oliver MJ, Dowd SE, Zaragoza J et al (2004) The rehydration transcriptome of the desiccation-tolerant bryophyte *Tortula ruralis*: transcript classification and analysis. BMC Genomics 5(1):89. doi:10.1186/1471-2164-5-89.

Oliver MJ, O'Mahony P, Wood AJ (1998) "To dryness and beyond" - Preparation for the dried state and rehydration in vegetative desiccation-tolerant plants. Plant Growth Regulation 24(3):193. doi:1005863015130.

Oliver MJ, Tuba Z, Mishler BD (2000) The Evolution of Vegetative Desiccation Tolerance in Land Plants. Plant Ecology 151(1):85-100. doi:1026550808557.

Öquist G, Fork DC (1982) Effects of desiccation on the excitation energy distribution from phycoerythrin to the two photosystems in the red alga *Porphyra perforata*. Physiologia Plantarum 56(1):56-62. doi:10.1111/j.1399-3054.1982.tb04899.x.

Costa-Silva JC, Domingues D, Lopes FM (2017) RNA-Seq differential expression analysis: An extended review and a software tool. PLoS One 12(12):e0190152. doi:10.1371/journal.pone.0190152.

Pabinger S, Dander A, Fischer M et al (2014) A survey of tools for variant analysis of next-generation genome sequencing data. Briefings in Bioinformatics 15(2):256-278. doi:10.1093/bib/bbs086.

Pade N, Linka N, Ruth W et al (2015) Floridoside and isofloridoside are synthesized by trehalose 6-phosphate synthase-like enzymes in the red alga *Galdieria sulphuraria*. New Phytologist 205(3):1227-1238. doi:10.1111/nph.13108.

Pampurova S, Verschooten K, Avonce N et al (2014) Functional screening of a cDNA library from the desiccation-tolerant plant *Selaginella lepidophylla* in yeast mutants identifies trehalose biosynthesis genes of plant and microbial origin. J Plant Res 127(6):803-813. doi:10.1007/s10265-014-0663-x.

Parages ML, Capasso JM, Niell FX et al (2014) Responses of cyclic phosphorylation of MAPK-like proteins in intertidal macroalgae after environmental stress. Journal of Plant Physiology 171(3-4):276-284. doi:10.1016/j.jplph.2013.08.005.

Paritosh K, Yadava SK, Gupta V et al (2013) RNA-seq based SNPs in some agronomically important oleiferous lines of *Brassica rapa* and their use for genome-wide linkage mapping and specific-region fine mapping. BMC Genomics 14(1):463. doi:10.1186/1471-2164-14-463.

Parra G, Bradnam K, Korf I (2007) CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. Bioinformatics 23(9):1061-1067. doi:10.1093/bioinformatics/btm071.

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. Bioinformatics (Oxford, England) 28(19):2537-2539. doi:10.1093/bioinformatics/bts460.

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6:288-295. doi:10.1111/j.1471-8286.2005.01155.x.

Peixoto L, Risso D, Poplawski SG et al (2015) How data analysis affects power, reproducibility and biological insight of RNA-seq studies in complex datasets. Nucleic Acids Research 43(16):7664-7674. doi:10.1093/nar/gkv736.

Peña EJ, Zingmark R, Nietch C (1999) Comparative photosynthesis of two species of intertidal epiphytic macroalgae on mangrove roots during submersion and emersion. Journal of Phycology 35(6):1206-1214. doi:10.1046/j.1529-8817.1999.3561206.x.

Persson F, Svensson R, Nylund GM et al (2011) Ecological role of a seaweed secondary metabolite for a colonizing bacterial community. Biofouling 27(6):579-588. doi:10.1080/08927014.2011.589001.

Pick U (1998) Dunaliella-A Model Extremophilic Alga. Israel Journal of Plant Sciences 46(2):131-139. doi:10.1080/07929978.1998.10676720.

Piskol R, Ramaswami G, Li JB (2013) Reliable identification of genomic variants from RNA-seq data. American Journal of Human Genetics 93(4):641-651. doi:10.1016/j.ajhg.2013.08.008.

Plettner I, Steinke M, Malin G (2005) Ethene (ethylene) production in the marine macroalga *Ulva* (*Enteromorpha*) *intestinalis* L. (Chlorophyta, Ulvophyceae): effect of light-stress and co-production with dimethyl sulphide. Plant, Cell & Environment 28(9):1136-1145. doi:10.1111/j.1365-3040.2005.01351.x.

Prestrelski SJ, Tedeschi N, Arakawa T et al (1993) Dehydration-induced conformational transitions in proteins and their inhibition by stabilizers. Biophysical Journal 65(2):661-671. doi:10.1016/S0006-3495(93)81120-2.

Price TD, Grant PR, Gibbs HL et al (1984) Recurrent patterns of natural selection in a population of Darwin's finches. Nature 309(5971):787-789. doi:10.1038/309787a0.

Pryce RJ (1972) The occurrence of lunularic and abscisic acids in plants. Phytochemistry 11(5):1759-1761. doi:10.1016/0031-9422(72)85033-7.

Qian F, Luo Q, Yang R et al (2015) The littoral red alga *Pyropia haitanensis* uses rapid accumulation of floridoside as the desiccation acclimation strategy. Journal of Applied Phycology 27(1):621-632. doi:10.1007/s10811-014-0336-0.

Quadir A, Harrison PJ, DeWreede RE (1979) The effects of emergence and submergence on the photosynthesis and respiration of marine macrophytes. Phycologia 18(1):83-88. doi:10.2216/i0031-8884-18-1-83.1.

Quinn EM, Cormican P, Kenny EM et al (2013) Development of strategies for SNP detection in RNA-seq data: application to lymphoblastoid cell lines and evaluation using 1000 Genomes data. PloS one 8(3):e58815. doi:10.1371/journal.pone.0058815.

Raja R, Hemaiswarya S, Kumar NA et al (2008) A Perspective on the Biotechnological Potential of Microalgae. Critical Reviews in Microbiology 34(2):77-88. doi:10.1080/10408410802086783.

Ram Y, Hadany L (2014) Stress-induced mutagenesis and complex adaptation. Proceedings. Biological Sciences 281(1792). doi:10.1098/rspb.2014.1025.

Ramanjulu S, Bartels DME (2002) Drought and desiccation-induced modulation of gene expression in plants. Plant, Cell and Environment 25(2):141-151. doi:10.1046/j.0016-8025.2001.00764.x.

Ramos AA, Polle J, Tran D et al (2011) The unicellular green alga *Dunaliella salina* Teod. as a model for abiotic stress tolerance: genetic advances and future perspectives. Algae 26(1):3-20. doi:10.4490/algae.2011.26.1.003.

Rao KN, Kumaran D, Seetharaman J et al (2006) Crystal structure of trehalose-6-phosphate phosphatase–related protein: Biochemical and biological implications. Protein Science 15(7):1735-1744. doi:10.1110/ps.062096606.

Rapaport F, Khanin R, Liang Y et al (2013) Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. Genome biology 14(9):R95. doi:10.1186/gb-2013-14-9-r95.

Ratkevicius N, Correa JA, Moenne A (2003) Copper accumulation, synthesis of ascorbate and activation of ascorbate peroxidase in *Enteromorpha compressa* (L.) Grev. (Chlorophyta) from heavy metal-enriched environments in northern Chile. Plant, Cell & Environment 26(10):1599-1608. doi:10.1046/j.1365-3040.2003.01073.x.

Ratnakumar S, Tunnacliffe A (2006) Intracellular trehalose is neither necessary nor sufficient for desiccation tolerance in yeast. FEMS Yeast Research 6(6):902-913. doi:10.1111/j.1567-1364.2006.00066.x.

Redlinger T, Gantt E (1982) A M(r) 95,000 polypeptide in *Porphyridium cruentum* phycobilisomes and thylakoids: Possible function in linkage of phycobilisomes to thylakoids and in energy transfer. Proceedings of the National Academy of Sciences of the United States of America 79(18):5542.

Redmond S, Green L, Yarish C et al (2014) New England Seaweed Culture Handbook. Seaweed Cultivation.

Reed RH (1990) Solute accumulation and osmotic acclimation. In: Cole KM, Sheath RG (eds) Biology of the Red Algae Cambridge University Press, Cambridge, 147-170.

Reed RH, Collins JC, Russell G (1980) The Effects of Salinity upon Galactosyl-Glycerol Content and Concentration of the Marine Red Alga *Porphyra purpurea* (Roth) C.Ag. Journal of Experimental Botany 31(6):1539-1554. doi:10.1093/jxb/31.6.1539.

Reed RH, Richardson DL, Warr SRC et al (1984) Carbohydrate Accumulation and Osmotic Stress in Cyanobacteria. Journal of General Microbiology 130(1):1. doi:10.1099/00221287-130-1-1.

Rezaei MK, Shobbar Z, Shahbazi M et al (2013) Glutathione S-transferase (GST) family in barley: Identification of members, enzyme activity, and gene expression pattern. Journal of Plant Physiology 170(14):1277-1284. doi:10.1016/j.jplph.2013.04.005.

Rippin M, Becker B, Holzinger A (2017) Enhanced Desiccation Tolerance in Mature Cultures of the Streptophytic Green Alga *Zygnema circumcarinatum* Revealed by Transcriptomics. Plant & Cell Physiology 58(12):2067-2084. doi:10.1093/pcp/pcx136.

Roach T, Krieger-Liszka A (2014) Regulation of Photosynthetic Electron Transport and Photoinhibition. Current Protein & Peptide Science 15(4):351-362.

Roberts MF (2005) Organic compatible solutes of halotolerant and halophilic microorganisms. Saline Systems 1:5. doi:10.1186/1746-1448-1-5.

Robinson N, Winberg P, Kirkendale L (2013) Genetic improvement of macroalgae: status to date and needs for the future. J Appl Phycol 25(3):703-716. doi:10.1007/s10811-012-9950-x.

Romiguier J, Gayral P, Ballenghien M et al (2014) Comparative population genomics in animals uncovers the determinants of genetic diversity. Nature 515(7526):261-263. doi:10.1038/nature13685.

Royer C (2017) Advancing Development of *Porphyra umbilicalis* as a Red Algal Model System and Aquaculture Crop, University of Maine.

Rustérucci C, Montillet JC, Agnel JP et al (1999) Involvement of lypoxigenase-dependent production of fatty acid hydroperoxides in the development of the hypersensitive cell death induced by criptogein on tobacco leaves. Plant Physiology 11:885-891.

Sagi M, Davydov O, Orazova S et al (2004) Plant Respiratory Burst Oxidase Homologs Impinge on Wound Responsiveness and Development in *Lycopersicon esculentum*. The Plant Cell 16(3):616-628. doi:10.1105/tpc.019398.

Sakamoto A, Murata N (2002) The role of glycine betaine in the protection of plants from stress: clues from transgenic plants. Plant, Cell & Environment 25(2):163-171. doi:10.1046/j.0016-8025.2001.00790.x.

Sampath-Wiley P, Neefus CD, Jahnke LS (2008) Seasonal effects of sun exposure and emersion on intertidal seaweed physiology: Fluctuations in antioxidant contents, photosynthetic pigments and photosynthetic efficiency in the red alga *Porphyra umbilicalis* Kützing (Rhodophyta, Bangiales). Journal of Experimental Marine Biology and Ecology 361(2):83-91. doi:10.1016/j.jembe.2008.05.001.

Schurch NJ, Schofield P, Gierliński M et al (2016) How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use?. RNA (New York, N.Y.) 22(6):839-851. doi:10.1261/rna.053959.115.

Schwartz SH, Qin X, Loewen MC (2004) The Biochemical Characterization of Two Carotenoid Cleavage Enzymes from *Arabidopsis* Indicates That a Carotenoid-derived Compound Inhibits Lateral Branching. Journal of Biological Chemistry 279(45):46940. doi:10.1074/jbc.M409004200.

Schwartz SH, Tan BC, Gage DA et al (1997) Specific Oxidative Cleavage of Carotenoids by VP14 of Maize. Science 276(5320):1872-1874. doi:10.1126/science.276.5320.1872.

Selander RK, Hudson RO (1976) Animal Population Structure Under Close Inbreeding: The Land Snail Rumina in Southern France. The American Naturalist 110(974):695-718. doi:10.1086/283098.

Seo J, Ju YS, Lee W et al (2012) The transcriptional landscape and mutational profile of lung adenocarcinoma. Genome Research 22(11):2109.

Shabala S, Demidchik V, Shabala L et al (2006) Extracellular Ca^{2+} ameliorates NaCl-induced K+ loss from *Arabidopsis* root and leaf cells by controlling plasma membrane K^{+} -permeable channels. Plant Physiol 141(4):1653-65.

Shearman JR, Sangsrakru D, Jomchai N et al (2015) SNP Identification from RNA Sequencing and Linkage Map Construction of Rubber Tree for Anchoring the Draft Genome. PLoS One 10(4). doi:10.1371/journal.pone.0121961.

Shiu C, Lee T (2005) Ultraviolet-B-induced oxidative stress and responses of the ascorbate–glutathione cycle in a marine macroalga *Ulva fasciata*. Journal of Experimental Botany 56(421):2851-2865. doi:10.1093/jxb/eri277.

Shoubaky GAE, Salem EA (2016) Effect of Abiotic Stress on Endogenous Phytohormones Profile in Some Seaweeds. International Journal of Pharmacognosy and Phytochemical Research 8(1):124-134.

Slama I, Abdelly C, Bouchereau A et al (2015) Diversity, distribution and roles of osmoprotective compounds accumulated in halophytes under abiotic stress. Annals of botany 115(3):433-447. doi:10.1093/aob/mcu239.

Smirnoff N (2000) Ascorbate biosynthesis and function in photoprotection. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 355(1402):1455-1464. doi:10.1098/rstb.2000.0706.

Smith BM, Melis A (1987) Photosystem Stoichiometry and Excitation Distribution in Chloroplasts from Surface and Minus 20 Meter Blades of Macrocystis pyrifera, the Giant Kelp. Plant Physiology 84(4):1325-1330. doi:10.1104/pp.84.4.1325.

Song L, Florea L (2015) Rcorrector: efficient and accurate error correction for Illumina RNA-seq reads. GigaScience 4(1):48. doi:10.1186/s13742-015-0089-y.

Sørensen I, Pettolino FA, Bacic A et al (2011) The charophycean green algae provide insights into the early origins of plant cell walls. The Plant Journal 68(2):201-211. doi:10.1111/j.1365-313X.2011.04686.x.

Southward AJ (1958) The zonation of plants and animals on rocky sea shores. Biological Reviews 33(2):137-177. doi:10.1111/j.1469-185X.1958.tb01444.x.

Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30(9):1312-1313. doi:10.1093/bioinformatics/btu033.

Staufenberger T, Thiel V, Wiese J et al (2008) Phylogenetic analysis of bacteria associated with *Laminaria saccharina*. FEMS Microbiology Ecology 64(1):65-77. doi:10.1111/j.1574-6941.2008.00445.x.

Sutherland JE, Lindstrom SC, Nelson WA et al (2011) A New Look At Ancient Order: Generic Revision Of The Bangiales (Rhodophyta). Journal of Phycology 47(5):1131-1151. doi:10.1111/j.1529-8817.2011.01052.x.

Takezawa D, Komatsu K, Sakata Y (2011) ABA in bryophytes: how a universal growth regulator in life became a plant hormone?. J Plant Res 124(4):437-453. doi:10.1007/s10265-011-0410-5.

Teasdale B, West A, Taylor H et al (2002) A simple restriction fragment length polymorphism (RFLP) assay to discriminate common *Porphyra* (Bangiophyceae, Rhodophyta) taxa from the Northwest Atlantic. Journal of Applied Phycology 14(4):293-298. doi:1021180315743.

Thoma I, Loeffler C, Sinha AK et al (2003) Cyclopentenone isoprostanes induced by reactive oxygen species trigger defense gene activation and phytoalexin accumulation in plants. The Plant Journal 34(3):363-375. doi:10.1046/j.1365-313X.2003.01730.x.

Timme RE, Delwiche CF (2010) Uncovering the evolutionary origin of plant molecular processes: comparison of *Coleochaete* (Coleochaetales) and *Spirogyra* (Zygnematales) transcriptomes. BMC Plant Biology 10(1):96. doi:10.1186/1471-2229-10-96.

Tominaga N, Takahata M, Tominaga H (1993) Effects of NaCl and KNO3 concentrations on the abscisic acid content of *Dunaliella* sp. (Chlorophyta). Hydrobiologia 267(1-3):163-168. doi:10.1007/BF00018798.

Torres MA, Dangl JL (2005) Functions of the respiratory burst oxidase in biotic interactions, abiotic stress and development. Current Opinion in Plant Biology 8(4):397-403. doi:10.1016/j.pbi.2005.05.014.

Toulza E, Blanc-Mathieu R, Gourbière S et al (2012) Environmental and Evolutionary Genomics of Microbial Algae: Power and Challenges of Metagenomics. In: Anonymous Advances in Botanical Research, vol 64. Elsevier Science & Technology, 383-427.

Tripathy BC, Oelmüller R (2012) Reactive oxygen species generation and signaling in plants. Plant Signaling & Behavior 7(12):1621-1633. doi:10.4161/psb.22455.

Tujula NA, Crocetti GR, Burke C et al (2010) Variability and abundance of the epiphytic bacterial community associated with a green marine Ulvacean alga. The ISME journal 4(2):301-311. doi:10.1038/ismej.2009.107.

Uji T, Matsuda R, Takechi K et al (2016) Ethylene regulation of sexual reproduction in the marine red alga *Pyropia yezoensis* (Rhodophyta). J Appl Phycol 28(6):3501-3509. doi:10.1007/s10811-016-0904-6.

Ulf Karsten, Kevin D. Barrow, Robert J. King (1993) Floridoside, L-Isofloridoside, and D-Isofloridoside in the Red Alga *Porphyra columbina*: Seasonal and Osmotic Effects. Plant Physiology 103(2):485-491.

Umehara M, Hanada A, Yoshida S et al (2008) Inhibition of shoot branching by new terpenoid plant hormones. Nature 455(7210):195-200. doi:10.1038/nature07272.

Van Belleghem SM, Roelofs D, Van Houdt J et al (2012) De novo transcriptome assembly and SNP discovery in the wing polymorphic salt marsh beetle *Pogonus chalceus* (Coleoptera, Carabidae). PloS one 7(8):e42605. doi:10.1371/journal.pone.0042605.

Van de Poel B, Cooper ED, Van Der Straeten D et al (2016) Transcriptome Profiling of the Green Alga *Spirogyra pratensis* (Charophyta) Suggests an Ancestral Role for Ethylene in Cell Wall Metabolism, Photosynthesis, and Abiotic Stress Responses. Plant Physiology 172(1):533-545. doi:10.1104/pp.16.00299.

Van Heerden, P. D. R, Tsimilli-Michael M, Krüger GHJ et al (2003) Dark chilling effects on soybean genotypes during vegetative development: parallel studies of CO_2 assimilation, chlorophyll a fluorescence kinetics O-J-I-P and nitrogen fixation. Physiologia Plantarum 117(4):476-491. doi:10.1034/j.1399-3054.2003.00056.x.

Van Houtte H, Vandesteene L, López-Galvis L et al (2013) Overexpression of the Trehalase Gene AtTRE1 Leads to Increased Drought Stress Tolerance in Arabidopsis and Is Involved in Abscisic Acid-Induced Stomatal Closure. Plant Physiology 161(3):1158-1171. doi:10.1104/pp.112.211391.

Vandesteene L, Ramon M, Le Roy K et al (2010) A single active trehalose-6-P synthase (TPS) and a family of putative regulatory TPS-like proteins in Arabidopsis. Molecular Plant 3(2):406-419. doi:10.1093/mp/ssp114.

Vertucci CW, Farrant JM (1995) Acquisition and loss of desiccation tolerance. In: Kigel J, Galili G (eds) Seed development and germination Marcel Dekker, New York, 237–271.

Vogel G, Aeschbacher RA, Müller J et al (1998) Trehalose-6-phosphate phosphatases from *Arabidopsis thaliana*: identification by functional complementation of the yeast tps2 mutant. The Plant Journal : for Cell and Molecular Biology 13(5):673-683. doi:10.1046/j.1365-313X.1998.00064.x.

Vogel G, Fiehn O, Jean-Richard-dit-Bressel L et al (2001) Trehalose metabolism in *Arabidopsis*: occurrence of trehalose and molecular cloning and characterization of trehalose-6-phosphate synthase homologues. Journal of Experimental Botany 52(362):1817-1826. doi:10.1093/jexbot/52.362.1817.

Wachi Y, Grant Burgess J, Iwamoto K et al (1995) Effect of ultraviolet-A (UV-A) light on growth, photosynthetic activity and production of biopterin glucoside by the marine UV-A resistant cyanobacterium *Oscillatoria sp.* BBA - General Subjects 1244(1):165-168. doi:10.1016/0304-4165(94)00219-N.

Wada N, Sakamoto T, Matsugo S (2013) Multiple roles of photosynthetic and sunscreen pigments in cyanobacteria focusing on the oxidative stress. Metabolites 3(2):463-483. doi:10.3390/metabo3020463.

Walker AB, Fournier HR, Neefus CD et al. (2009) Partial Replacement of Fish Meal with Laver *Porphyra* spp. in Diets for Atlantic Cod. North American Journal of Aquaculture 71(1):39-45.

Wang G, Zhao G, Feng Y et al (2010) Cloning and comparative studies of seaweed trehalose-6-phosphate synthase genes. Marine Drugs 8(7):2065-2079. doi:10.3390/md8072065.

Wang L, Mao Y, Kong F et al (2015) Genome-wide expression profiles of *Pyropia haitanensis* in response to osmotic stress by using deep sequencing technology. BMC Genomics 16(1):1012. doi:10.1186/s12864-015-2226-5.

Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nature Reviews Genetics 10(1):57-63. doi:10.1038/nrg2484.

Waterhouse AM, Procter JB, Martin DMA et al (2009) Jalview Version 2-a multiple sequence alignment editor and analysis workbench. Bioinformatics 25(9):1189-1191. doi:10.1093/bioinformatics/btp033.

Wegmann K (1986) Osmoregulation in eukaryotic algae. FEMS Microbiology Letters 39(1-2):37-43. doi:10.1111/j.1574-6968.1986.tb01840.x.

Welsch R, Beyer P, Hugueney P et al (2000) Regulation and activation of phytoene synthase, a key enzyme in carotenoid biosynthesis, during photomorphogenesis. Planta 211(6):846-854. doi:10.1007/s004250000352.

Welsh DT (2000) Ecological significance of compatible solute accumulation by micro-organisms: from single cells to global climate. FEMS Microbiology Reviews 24(3):263-290. doi:10.1016/S0168-6445(99)00038-8.

Wensel P (2018) Factors to Consider in Algal-Strain Selection for Commercial Manufacture of Biofuels and Co-Products. Biomedical Journal of Scientific & Technical Research 2(1). doi:10.26717/BJSTR.2018.02.000691.

Widmann C, Gibson S, Jarpe MB et al (1999) Mitogen-Activated Protein Kinase: Conservation of a Three-Kinase Module From Yeast to Human. Physiological Reviews 79(1):143-180. doi:10.1152/physrev.1999.79.1.143.

Wiemken A (1990) Trehalose in yeast, stress protectant rather than reserve carbohydrate. Antonie van Leeuwenhoek 58(3):209-217. doi:10.1007/BF00548935.

Wiencke C, Läuchli A (1981) Inorganic ions and floridoside as osmotic solutes in *Porphyra umbilicalis*. Z. Pflanzenphysio 103:247–258.

Wilcockson D, Zhang L (2008) Circatidal clocks. Current Biology, 18(17):753-755.

Wilhelm BT, Marguerat S, Watt S et al (2008) Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. Nature 453(7199):1239-1243.

Worrall JJ (2012) Structure and dynamics of fungal populations. Springer Science & Business Media, Netherlands.

Xia X, Wang Y, Zhou Y et al (2009) Reactive Oxygen Species Are Involved in Brassinosteroid-Induced Stress Tolerance in Cucumber. Plant Physiology 150(2):801-814. doi:10.1104/pp.109.138230.

Xiong L, Zhu J (2003) Regulation of Abscisic Acid Biosynthesis. Plant Physiology 133(1):29-36. doi:10.1104/pp.103.025395.

Xu K, Xu Y, Ji D et al (2016) Proteomic analysis of the economic seaweed *Pyropia haitanensis* in response to desiccation. Algal Research 19:198-206. doi:10.1016/j.algal.2016.08.010.

Xu Y, Wang Y, Mattson N et al (2017) Genome-wide analysis of the *Solanum tuberosum* (potato) trehalose-6-phosphate synthase (TPS) gene family: evolution and differential expression during development and stress. BMC Genomics 18:926. doi:10.1186/s12864-016-3381-z.

Yancey PH (2005) Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. Journal of Experimental Biology 208(15):2819-2830. doi:10.1242/jeb.01730.

Ye J, Zhang Y, Cui H et al (2018) WEGO 2.0: a web tool for analyzing and plotting GO annotations, 2018 update. Nucleic acids research 46(W1):W75. doi:10.1093/nar/gky400.

Yu Q, Ghisla S, Hirschberg J et al (2011) Plant Carotene Cis-Trans Isomerase CRTISO. Journal of Biological Chemistry 286(10):8666-8676. doi:10.1074/jbc.M110.208017.

Zentella R, Mascorro-Gallardo JO, Van Dijck P et al (1999) A Selaginella lepidophylla Trehalose-6-Phosphate Synthase Complements Growth and Stress-Tolerance Defects in a Yeast tps1 Mutant. Plant Physiology 119(4):1473-1482. doi:10.1104/pp.119.4.1473.

Zhang B, Yan X, Huang L (2011) Evaluation of an improved strain of *Porphyra yezoensis* Ueda (Bangiales, Rhodophyta) with high-temperature tolerance. J Appl Phycol 23(5):841-847. doi:10.1007/s10811-010-9587-6.

Zhang D (2014) Abscisic Acid: Metabolism, Transport and Signaling, Zhang, Da-Peng edn. Springer, New York, NY.

Zhang J, Wheeler DA, Yakub I et al (2005) SNPdetector: a software tool for sensitive and accurate SNP detection. PLoS Computational Biology 1(5):e53. doi:10.1371/journal.pcbi.0010053.

Zhang Q, Bartels D (2018) Molecular responses to dehydration and desiccation in desiccation-tolerant angiosperm plants. Journal of Experimental Botany 69(13):3211-3222. doi:10.1093/jxb/erx489.

Zhang S, Klessig DF (2001) MAPK cascades in plant defense signaling. Trends in Plant Science 6(11):520-527.

Zhu J (2002) Salt and drought stress signal transduction in plants. Annual Review of Plant Biology 53:247.

Zou D, Gao K (2002) Effects of desiccation and CO_2 concentrations on emersed photosynthesis in *Porphyra haitanensis* (Bangiales, Rhodophyta), a species farmed in China. European Journal of Phycology 37(4):587-592. doi:10.1017/S0967026202003876.

APPENDICIES

Supplement Table 1: The haplotype of the 37 individuals of *P. umbilicalis* from 7 populations along the Gulf of Maine

Ind	P2_1	P2_2	P2_3	P8	P13_1	P13_2	P13_3	P13_4	P13_5	P13_6	P13_7	P18	P22
F1	T	C	T	T	T	C	C	A	G	C	C	G	G
F3	T	C	T	T	T	C	C	A	G	C	C	G	G
F4	T	C	T	T	C	C	C	A	G	C	C	G	G
F5	T	C	T	T	C	C	C	A	G	C	C	G	G
F6	T	C	T	T	C	C	C	A	G	C	C	G	G
L1	T	C	T	T	T	T	T	G	A	T	C	G	G
L2	T	C	T	T	C	C	C	A	G	C	C	G	G
L4	T	C	T	T	T	C	C	A	G	C	C	G	G
L7	T	C	T	T	T	C	C	A	G	C	C	G	G
L8	T	C	T	T	C	C	C	A	G	C	C	G	G
W1	T	T	T	T	T	C	C	A	G	C	C	G	G
W2	T	T	T	T	T	C	C	A	G	C	C	G	G
W3	T	T	T	T	T	C	C	A	G	C	C	G	G
W6	T	T	T	T	T	C	C	A	G	C	C	G	G
W8	T	T	T	T	T	C	C	A	G	C	C	G	G
N1	T	C	T	T	T	C	C	A	G	C	C	G	G
N2	T	C	T	T	T	C	C	A	G	C	C	G	G
N3	T	C	T	T	T	C	C	A	G	C	C	G	G
N4	T	C	T	T	T	C	C	A	G	C	C	G	G
N5	T	C	T	T	T	C	C	A	G	C	C	G	G

S1	C	C	T	G	T	C	C	A	G	C	C	G	A
S2	T	C	C	T	T	C	C	A	G	C	C	G	G
S3	T	C	T	T	T	C	C	A	G	C	C	G	A
S4	T	C	T	T	T	C	C	A	G	C	T	G	G
S6	C	C	T	T	T	C	C	A	G	C	T	G	A
S7	T	C	T	T	T	C	C	A	G	C	C	G	G
S9	C	C	T	T	T	C	C	A	G	C	C	G	A
D1	T	C	T	T	T	C	C	A	G	C	C	G	G
D2	T	C	T	T	T	C	C	A	G	C	C	G	G
D3	T	C	T	T	T	C	C	A	G	C	C	G	G
D4	T	C	T	T	T	C	C	A	G	C	C	G	G
D5	T	C	T	T	T	C	C	A	G	C	C	G	G
R1	T	C	T	G	C	C	C	A	G	C	C	C	G
R2	T	C	T	G	C	C	C	A	G	C	C	C	G
R3	T	C	T	G	C	C	C	A	G	C	C	C	G
R4	T	C	T	G	C	C	C	A	G	C	C	C	G
R5	T	C	T	G	C	C	C	A	G	C	C	C	G

*The individual and their haplotype at 13 SNP loci

Supplement Table 2: List of the trehalose-6-phosphate synthase, (iso)floridoside phosphate synthase and glucosylglycerol-phosphate synthase proteins used in constructing the phylogenetic tree

Protein	organism	NCBI Reference Accession number	
GGPS1	*Synechocystis sp.* PPC 6803	WP_010872972	Bacteria
GGPS2	*Synechocystis sp*. PCC 6803	WP_012308437	Bacteria
GGPS	*Pseudomonas anguilliseptica*	Q93JY3	Bacteria

Gasu10960	*Galdieria sulphuraria*	EME31717	Algae
Gasu22110	*Galdieria sulphuraria*	EME30539	Algae
Gasu26940	*Galdieria sulphuraria*	EME29908	Algae
Gasu01890	*Galdieria sulphuraria*	EME32831	Algae
TPSA	*Caenorhabditis elegans*	Q7YZT6	Nematode
TPSB	*Caenorhabditis elegans*	O45380	Nematode
TPS	*Aphelenchus avenae*	Q5K2C4	Nematode
TPS	*Saccharomyces cerevisiae*	Q00764	Fungi
TPSA	*Schizosaccharomyces pombe*	P40387	Fungi
TPSB	*Schizosaccharomyces pombe*	O14081	Fungi
TPS	*Candida albicans*	Q92410	Fungi
TPS	*Emericella nidulans*	O59921	Fungi
TPS	*Kluyveromyces lactis*	Q07158	Fungi
TPSA	*Dictyostelium discoideum*	Q54K57	Fungi
TPSB	*Dictyostelium discoideum*	Q54NU9	Fungi
TPS	*Yarrowia lipolytica*	O74932	Fungi
TPS	*Encephalitozoon cuniculi*	Q8SSL2	Fungi
TPS	*Shigella dysenteriae*	Q32H58	Bacteria
TPS	*Cronobacter sakazakii*	A7MEE9	Bacteria
TPS	*Sinorhizobium fredii*	P55612	Bacteria
TPS	*Klebsiella pneumoniae*	A6TB47	Bacteria
TPS	*Sodalis glossinidius*	Q2NTK9	Bacteria
TPS	*Enterobacter sp.* (strain 638)	A4WBR1	Bacteria
TPS	*Pseudomonas savastanoi*	Q6JTB2	Bacteria
TPS	*Citrobacter koseri*	A8AFD4	Bacteria
OSTA	*Escherichia coli*	Q8XCE7	Bacteria
TPS	*Synechocystis sp.* PCC 6714	AIE73228	Bacteria

TPS	*Mycobacterium tuberculosis*	P9WN11	Bacteria
TPS	*Salmonella enterica*	P0A1Q1	Bacteria
TPS1	*Arabidopsis thaliana*	Q9SYM4	Plantae
TPS2	*Arabidopsis thaliana*	Q9FZ57	Plantae
TPS3	*Arabidopsis thaliana*	Q9SHG0	Plantae
TPS4	*Arabidopsis thaliana*	Q9T079	Plantae
TPS5	*Arabidopsis thaliana*	O23617	Plantae
TPS6	*Arabidopsis thaliana*	Q94AH8	Plantae
TPS7	*Arabidopsis thaliana*	Q9LMI0	Plantae
TPS8	*Arabidopsis thaliana*	Q0WUI9	Plantae
TPS9	*Arabidopsis thaliana*	Q9LRA7	Plantae
TPS10	*Arabidopsis thaliana*	O80738	Plantae
TPS11	*Arabidopsis thaliana*	Q9ZV48	Plantae
TPS1	*Pyropia haitanensis*	AGT98612	Algae
TPS2	*Pyropia haitanensis*	AJA06053	Algae
TPS3	*Pyropia haitanensis*	AJA06054	Algae
TPS4	*Pyropia haitanensis*	AHB86958	Algae
TPS5	*Pyropia haitanensis*	ABG75728	Algae
TPS6	*Pyropia haitanensis*	AJA04606	Algae
TPS	*Pyropia yezoensis*	AAW27916	Algae
TPS	*Undaria pinnatifida*	ADB19855	Algae
TPS	*Sargassum henslowianum*	ADB19856	Algae
TPS1	*Cyanidioschyzon merolae*	XP_005536433	Algae
TPS2	*Cyanidioschyzon merolae*	XP_005537876	Algae
TPS3	*Cyanidioschyzon merolae*	XP_005537475	Algae
TPS	*Chondrus crispus*	XP_005713653	Algae
TPS	*Chondrus crispus*	XP_005712424	Algae

TPS	*Chondrus crispus*	XP_005710571	Algae
TPS	*Chondrus crispus*	XP_005713306	Algae
TPS4637	*Porphyra umbilicalis*		Algae
TPS8501	*Porphyra umbilicalis*		Algae
TPS4785	*Porphyra umbilicalis*	OSX81421	Algae
TPS5014	*Porphyra umbilicalis*	OSX79290	Algae

>Pum5014 (BU14_0082s0055)
MSFPSSSVSNSRVTFVLRCATEFGQRVRVIGNDSRLGGWDPRRGVELTTSTGAWPSWRVSVDLPAGHKVEY
KYVVLQSNPSALAGDTLVVGSPDGPGMKAPTPIAGGGAEGSGSGSSAGILSPESVDGTMQWELFEENRTLT
TRSGQMIVNDGEFGKWNDKEEGRVKCFLTTDEGFGDSIESLVIVLYRLPIISKRDHTTGAWDFKWDDDALYL
TSTGLRKGLEQLKVAPLWVGILNSDEEVPMRERDGVADRLLEEFNCVPVFIPHDTLKQFYEGFCKGVLWPLF
HMVSTATDHTQHTTRFDDRLWRVYMNVNRMFRDKVVEVYDGDRQLIWVHDYHLMLLPQALRSRLSGVKI
GFFLHIPWPSSEVYRVLPWRNELLKGMLSATLLGFHLFDYARHFLSACVRLLNLEHEANRGSLGLEYDGRHV
MLRVSHIGVDPERFSEGLNESSLADRVAEFKQRFADCMVLGAVDDLDLIKGISLKLLGFQRYLDTAPNMRGK
VVLVQVAIPKAARVKESVRNEIRELVAAINDKHGDGSGRRPVWYLEESISFESRLALYSIMDALVLTPIRDGLNL
IPYEYIVSTSEGKGQLVLSEFTGCSRALSSAVRVNPWDIEELRGVLDMVVQKALSKAPEVELKRRADKSYVSAH
SSQQWAQSFLHDLKEASEPARAVVKVGPLAGLPGVLTYDEFTLLNRSSVLRAYKAAKRRLFLFDYDGTLTSITE
QSSQMAHAWARPNESVVANLDTLSKDPLNDVYIMSGRKTEVLEAGLNNSPAIGIAAEHGFYFRKKNSTEWN
KLLEDADLSWMELALRIMLMYTDRTDGSYVEQKKAGLVWHYLDADREFGSWQAKEMRDHLESLLSPFSVQ
VVSGYGWLQVRMSAMNKGVTVETILRDMPEAPDFVLCCGDDRTDEDMFAYLDTHLDPSVKQFTCTVGVK
PSHARYYLHSSNEVGALLETLVTGAYPRGGRSRHGGSMSLADMVADDEPSPPPPPQSNNGPRPAANSSG
QKRRGLASSLSGL

>Pum4785 (Bu14_0021s0026)
MFLVDVTHVPRLDRASLPPSRCRAGLPALFTSRSRTPLRVLIMSLVDALGDVLQPDGTAQSTLVPMTLTDME
RQERLIIGQMQDLRKKLDDARIRSGVKRSRRPKRRGAALSSDGISGEIAIHDLLSPGALAASAGPAAAAAAAA
ACARPPLVPGSASGRPSAAAATAAAAGAAAAVGGGSGGSGGGGGGCGGPAGLSAAMGAPPAGSFLSFAP
SGRADDSAFFLPAPSYDRPGRQRIVSVSLRLPSKADRQSARQQLFDSGLPPKGMFVLERTADVPFVWVGLLP
STEDGGGRSRGGDGHHGNGGRGNGRGGVGVGGGGGGGGGADSSGGSGGGVGGGGGGCYGGGWHR
GGGSGGGGFDPDDSVSHSFNPWSGEPSICGRRPRGGGGSSGRRLPRGSRGKDPAERSRHVAVTLPSSLEAP
FHAFCEETLWRLLHYDYGSLGNGGSGGVDTRDWDAYQAVNRRFAEAVTEVYEEGDLVWVHNYHLLLLPA
MLRRRLWYAKIGFFLYTPFPSSEIFRLLPQRSEVLRGVLGADLAGFHTYDYSKQFLASCARLLGLEGSPKGIQVE
PGGGHVCEIGIYPPGIDVAGLKAHVVSKAVRARVLELRDRFADRAVVVSVERLDDAFAGIPLTLLAFESLLHKY
PEYVTSAVLVLVATIPRHPRQLSSYRALASQINTSVGRINSTYGSIGSSPVHFINAELPQDELYALLSVGSVCVVS
SIRDGMSLVPYEWAICQHAGNRGPLILSEFAAAAHSFSTARHVNPWDVDDLRDKLAACLTMPAIDKRSRDE
AAYRFVTTHTAQLWGLNFLEDLEEIEPVRARLVATPHLDEAALSDSLRSSTKRKLFVLDYDGTLIPFHPLWQLA
APPPTVTDFVTNLVSQPDVEVIILSGRDRAKLSAWFPDSRVGLAAEHGHFLRLPNETEWQVLPRGYSANTSG
PPSTPAPVPPAAPPLPPPMGGGSSPAGAGAGEGVPDALVTPLSPLSTAPERLSMAPAEPPAAAGDCGGRED
GIDGGGPPAAAAATDRGDGGVDAFPLRFAASTRPSACDGAAHGASASSSTAGNEQRASILSTDGSVGSGSS
SGIGDDSGSVGGVDGGGGGGSRVVGGGRGGGVEGGSDRGGGCRVPNRSSSVGARPDSGNASPEDGICR
WGHGEPEAPSGGSGGGGFGGGGGAPPPPPSPLSSVAITNTPAPPGPLSPILGGTTTSTGHIAAWKSAVLPV
MRHFVERTPGAVMEEGEATVTWHYYDADVDFGRWQARDLQKHLESFLLQHLSVEVVSGERPGKWIKVRP
SGVDKAGAVQRALELSGARFDWAIVAGDDRSDEPMYELLRNERKLGELGFRGRALSVRVGLGIQTAADYVV
DSSTRLMTVIDSVLFEEEGDRGVPDERDFLSNPLDDGWGGAGFEGGARRARTIFLAGGPEPRLQLMQYCGE
CGGRPRRVGGDRCSARGKRGGKCR

>Pum8501
MPVTITNKILSPPRGWRGGPLLSAVHLRSPGRTRRFLVSPLYGPSTAFQGAFPVGHQGGDPTARAGPTDGPR
SASMAGILSRLLSLGHPVPPSAPPTEASSRDSSDPEDVPITSGTIGGVGGIGGAGGDGDGGDDDATTASSGG
RSRSSTVVSMERGGVSKVRMGASAERSSTNVVAPGGGGPGGGGGGSGGGTSATTTPAGKSGLLQTGPFSE
GELYDSGMGRGLTVPFPLDDMSGSVTGGSECGTGASGGDGSAVDTEVEADEAGAGLNEYQRELAEEEAAA
QMSSLASTLTEEVLLSRLQDLQMELKLLRDRRAAFLAAASGAPAPALSSDSLPSAATTGFAAHASAASRAAKR
VPPATRMHSGPLFVASFRLQLSVDITSETGAVKASLSAGGLGLVPAFRHLVSRTRIVWLGMPVFAAGQTPNG
ATQARIQARLRARKPSAVLSYAPIFPRRSDAVTHQAFCNNVLWALFHYLPLSFEGDRSFRPEMFESYKRVNEE
YALALLREFERSRRDENGVFWVHDFQLMLVPKMLRERMPHAKIGFFLHTPFPAGELYRTLPPRRELLEGMLG
ADLIGFHTYDYARHFLSACERILGLDIRPNCVDNHGVLVHVAIFPFGIDTKTFTSAMLRPTVIRQRDALKEELAG
KKVLLGIDRLDYIKGIPHKLLAFEHFLETYPEYVGHVVLVQIATPSSTTSEEYAGFRAEILEEVGRINGRFGTVDD
MPIHYREHAMSFDTLCALYSMADVAVITSLRDGMNLVSYEYIVCQAKNRGVLVLSEYTGAAQSLPGALLCNP
WSVEEVSHTLHVALTMSDAERELKHKKLYRYILMHSSSQWGLNFVSDLLQYSTARRMAVEKLMRLPVAHVR
WLFARRRRRLILLDYDGTLRTWESQPELAEPSPRLRALLKRLAADESNMVFIITGRQKDTMRRWFSGLGVGF
AVEHGYALQWPARVRSVFRQSAERVSVARPLDGPSPLAGVSPLVGPASLTAVPSPPIGPSPLASVLPPPGDVP
PPWGAPPPPPRRQERASPASSTAALPVPSRRPASSPLLMSSDGGSLEVPPFDTPALSLDLSWASGAAGGLAG
DATAAPGGEEEDWWEERPENDPATERALVAALAQARSLLKTFERHTPGSFVVDKESSVTWLYRDADTNFAF
TQAKEARQRLEEVLGGSPLEVLVGHKILYVRPRGVNKGATVRQIVRRMRHVDGGVGGPDLLFAIGDDKSDE
QMFEEVEGIREAAAAQRNTTSGALDEATHDEDRRREWEAEQHAARKSASATVAAAAATVATVAAAAAAG
EVPSLPAAPPPAPSPWAGAPVTVPGVSPRLRATSTHGSTRSLATLASDDVAAGDSSADEGGPDGGGPDGD
GLGGDAEAPAPPRRRRRRRRQPRWAPDEDGRAGASTAAGAPSTPPPSPPTPSAAAAASGTTPVAAGAAW
PTTPAAADGPGGGGSGGGGGSGGGSFGAPDKTTRLAATETDEDVAAAGSGSADIAVLTCTVGRKTTKARYC
FDDVADVIRALEELVDGGP

>Pum4637
MSAPTVQIPLTPSVPQPLSRGQAGNVSLPHLAAADVRSLGQDALRGTAGGEATSHLTEGEILARIQELHRELT
VVRESEEAAAGGAGGIGGASSETAAGADGRSGGSVLGENGLNGQLQPGRRLLVVSNRLPVTINKNPDGQW
DFKMSAGGLVSALAGVKNEFPFVWVGWSGSEVPDTDQDSLRRELRAQHGCVPVYLSDYDAHLYYNGFCN
DVLWPLFHYVPLPIVSSDGERKFDVKYWEAYSKANHRFAEAIMQVYEPGDLIWVQDYHLMLLPSLLRKRIRD
VTIGFFLHTPFPSSEVYRILPMRNKVLQGVLAADLIGFHTYDYARHFLSVCTRILGLEASPKGVDYKDHFAHVGI
FPIGIDPSAFIRALDLPSVQERASELQAKFAGKKVLLGVDRLDYIKGVPHKLMAFETLLARHPEWNERAVLVQI
AVPSRTEVEEYKKLSSQTHELVGRINGKFGSVDYSPIVFINQSVNFHDLVALYSVADVCVVSSIRDGMNLVSYE
YVMCQRERHGVLVLSEFAGSAQSLSGALRVNPWNIEELATALHDALSASPRERQLKQQKLYRYVTTHTAAF
WAQSFVSELRELANLEREARDARGANLPLPFLRVERELVSALQPRRHRLFLFEYEGTLCAPVALADLATPSSSL
RRYLARLSSDRANAVYIFSSRSKTVLDGWFGDLHVGLVAEHGCDFRHPGHPAWEPLIGLHDPAWRDEVVPIL
QYFSERTPGSHLELKEKVITWHFRDADPQFGSWQAKELQLLLAESCTNLPVEVVSGSKFLELRPVGVNRVAA
VQRIIAELPEPALDFVLCLGSDKADEEVFSYLSTYIKNEPEVSTLCCRVGKQQSSAADRYIQDVDSALRVLREIAP
TAKRTAGMSTGTSGQLSFFGLPTAGKPPNTQPTLGNFAAASRASSFLKNAGGVKNALQMAKKGSGGRLRDI
ATASLAASMSARAMNGKAPASSSASAAYLPTMVGGSGGIGSVNTASKSHAVLPSVTQPAASFSSPPQRSVL
THDPGVSSSPSIPTAMLPPSRSAVAGGRTAAAEGAAGPPSPTTHATPLPASDLAAVTGPSRPLPLATPSAVAS
RGASSLDMSALPSGGPGAPPYPSAPPPADGATTLDRPSAVSTPTSSSSGAAAAAAAGAMGHDRPTLKAAAG
VRFMDLEADAAEVSCSPP

>TPP1 (Bu14_0615S0014)
MRQVLSDLAHLFPTAIVSGRGRKKLQSLIGLAGHPGLFYGGSHGFDISGPTGVNDSLRRKEASEMLPALKRAS
AQLSASVAEFFGSQVEDNELAVSVHYRNLADKEALPELERRVDDVVEGSPELTKHYGKCVFEVRPRSDWHKG
KAVEYLLDALDLGGPDVLPVYIGDDVTDEDAFKTLHGRGLGIVVMSDEEVGEAAAADGERGAEGGAAGGEA
AGRGRRTAATMRVATTDEVRAFLRRFSEAGREGRCQGKAGASAARG

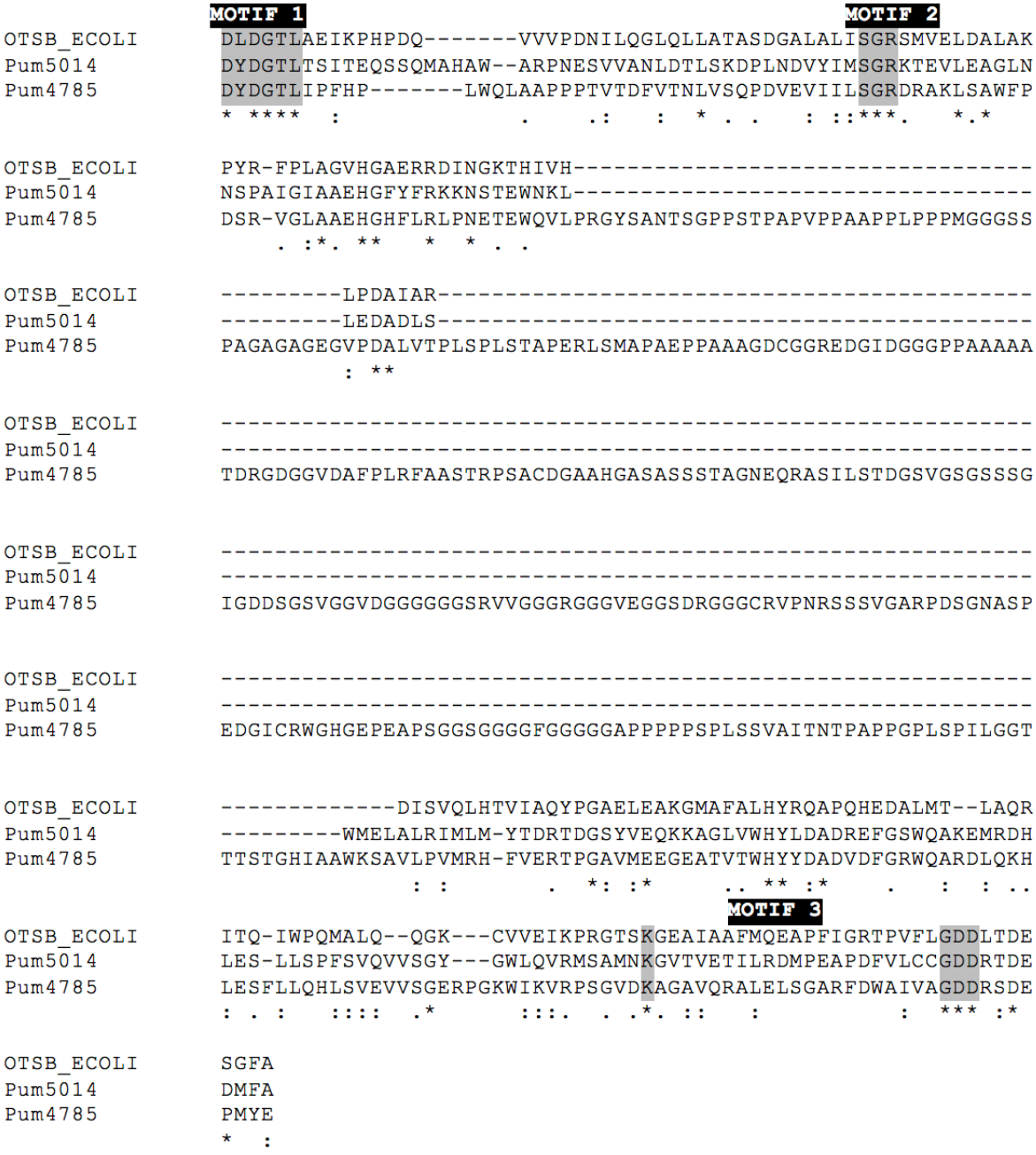
>TPP2 (Bu14_0615S0010)
MAPAPDAAGAPGAVAAAPSGAGEASNGHAAGGPGRDEAVWALGMEPSGRAPAGPVVKKNPMQLPPAL
DRVDTLQAALESGLRRPVFFLDYDGTLTPIVKDPDAALLSTEMRKVLTELATLFPTAIVSGRGRVKLQSLIGLHD
HQGLYYAGSHGFDISGPMGGDESLRRKEASETLPVLKRASSQLVATVADFSGSQVEDNELAVSVHYRNLVDK
EQLPALESCVDTVVNDLPGLSKHHGKCVFELRPETDWHKGKAVEYLLDALGLGGPDVLPIYIGDDVSDEDAFK
ALYGRGLGIIVMSDADVAEYRAKGHPDGRHTAATMRLCNTDDVRGFLGSFAAAGREGRCKGRSTPPPASVA
GRAVTVESGGT

>TreS (BU14_0165s0020)
MAKRTNPERWKDIFGAFFSELVAARDARPPALDARDAALPADWHTAPWYCVYPHLFGVPGDDAAPGATG
TYDDLIGALPYWADLGFRNLLLLPHYASSGGDGGYDVVDYSVDPALGGEAAWARFMAAAVAAGFRVATDA
VFNHTSTEHPWFQAALGGSAEHMDYYILRNGREKVGEYDRGGDILCRYRDPDGTVTERVCVFPDVDRTHGL
WVDIGGRTVTFYRSFYPFQVDLNLHNPAVLRALLRLLGAELNAGILGKRADAVAHWTKVAGGSDDGRPQTH
TLLALVKSFLRHVCGAAVVLPEVVRSAETAARYTGVPTAVCAVGCPSEGDLPLAFEMQAALREMTYFQTVAP
FWRRVFHTPRLPAGAAFLNLLEHHDETYMGFFAPDVRPWLREYITAHGGVVYKNGMSAGARVADCLDQD
PQRLATALFALYITPGVPLVYAGTELGAGNRYDVASAAMRARHVTFQRLGVFVAQSACYDPRDLQRGPIGA
ATLAAAGGGGGLPHAIVRCLNRLREAYRALRVPRVHPVDSGDIGVLCLVRDGRGGAGGGGRRPRAPRRRRR
ASRRAAWLVGGGGGTGGGGGGAAAVVAAAAAAAAAAAAPDASILCVANLTPFDKTVVMPVWQLVQRLH
LDEAGGGGTAAPPPPPPATVAAAVAAAVAAVDGGVHGGGGSVADGSLPAAVAAVAAASAAAVADAAAA
AAADDAPQSPLPPPPTPSDGGLTMVDLLTEAAVALERRGTAYVLALPPFSRYALVVRCSEGE

Supplement Figure 1: Protein sequences for the four putative tre(TPS)-like genes, two putative (TPP)-like genes, and one trehalose synthase/amylase (TreS) like gene in *Porphyra umbilicalis*

Supplement Figure 2: Sequence alignment of Pum4637 from *P. umbilicalis* and ostA protein from *E. coli*. Residues that form the conserved domains are highlight in black. The alignment was constructed using CLUSTAL-W with the *bottom* row providing the degree of homology: (*) identical, (:) conserved, (.) semiconserved.

Supplement Figure 3: Sequence alignment of TPP1 and TPP2 proteins from *P. umbilicalis* and ostB protein from *E. coli*. Residues that form the three characterizing motifs are highlight in grey. The alignment was constructed using CLUSTAL-W with the *bottom* row providing the degree of homology: (*) identical, (:) conserved, (.) semiconserved.



Supplement Figure 4: Sequence alignment of Pum5014 and Pum4785 proteins from *P. umbilicalis* and ostB protein from *E. coli*. Residues that form the three characterizing motifs are highlight in grey. The alignment was constructed using CLUSTAL-W with the *bottom* row providing the degree of homology: (*) identical, (:) conserved, (.) semiconserved.

Supplement Table 2: The overview of the sequencing for each library: the number of raw reads, the number of reads kept after trimming, the number of bases corrected by Rcorrector, and the number of reads mapped to the genome

Library	#of raw reads	# of reads of trimming	# of bases corrected	# of reads mapped
F1	44,148,108	43,915,304 (99.47%)	18,318,910 (0.31%)	37,280,010 (84.89%)
F2	45,338,280	45,092,582 (99.46%)	16,591,908 (0.29%)	35,976,119 (79.78%)
F4	36,686,194	36,468,932 (99.41%)	14,948,133 (0.31%)	32,582,165 (89.34%)
F5	42,682,278	42,449,512 (99.45%)	17,915,002 (0.32%)	37,971,735 (89.45%)
F8	43,731,432	43,492,352 (99.45%)	18,803,619 (0.31%)	35,879,720 (82.5%)
F9	57,181,144	56,841,188 (99.41%)	23,264,739 (0.31%)	42,970,433 (75.6%)
H1	49,103,882	48,784,568 (99.35%)	19,966,691 (0.31%)	42,565,028 (87.25%)
H2	49,326,800	49,010,286 (99.36%)	19,455,449 (0.3%)	41,781,107 (85.25%)
H4	44,892,262	44,565,386 (99.27%)	17,706,324 (0.3%)	38,367,882 (86.09%)
H5	48,855,876	48,429,034 (99.13%)	18,794,076 (0.3%)	42,195,373 (87.13%)
H8	43,221,270	42,976,622 (99.43%)	18,672,759 (0.32%)	37,391,108 (87.0%)
H9	46,798,412	46,543,134 (99.45%)	18,154,261 (0.29%)	39,898,709 (85.72%)
D1	44,420,298	44,148,240 (99.39%)	16,183,585 (0.28%)	35,187,906 (79.7%)
D2	44,353,000	44,081,236 (99.39%)	18,399,016 (0.32%)	34,592 155 (78.47%)
D3	46,517,310	46,160,900 (99.23%)	19,842,603 (0.32%)	38,741,751 (83.93%)

D4	47,213,486	46,897,990 (99.33%)	17,700,436 (0.29%)	38,998,946 (83.16%)
D5	43,863,970	43,614,702 (99.43%)	17,398,900 (0.3%)	36, 895,278 (84.59%)
D6	47,677,412	47,398,970 (99.42%)	21,819,591 (0.33%)	37, 191,500 (78.46%)
R1	46,074,840	45,809,204 (99.42%)	18,396,949 (0.3%)	40,388,709 (88.17%)
R2	44,045,426	43,781,790 (99.40%)	16,321,070 (0.29%)	37,564,336 (85.8%)
R4	43,987,146	43,722,990 (99.40%)	18,660,677 (0.3%)	38,191,397 (87.35%)
R5	48,191,884	47,873,106 (99.34%)	19,052,055 (0.3%)	41,543,111 (86.78%)
R8	45,888,518	45,628,176 (99.43%)	18,943,676 (0.31%)	38,517,956 (84.42%)
R9	39,985,770	39,704,918 (99.30%)	15,562,249 (0.3%)	32,432,487 (81.68%)

Supplement Table 3: DEGs between dehydration and fresh with gene name, Log_2fold change (Log2FC), putative protein name and putative GO functions

Gene	Log2FC	protein name	GO term	Go Function
gene8115	-6.8	hypothetical protein BU14_0357s0016	F:GO:0003674	F:molecular_function
gene4684	-6.1	Bile salt-activated lipase	F:GO:0003674	F:molecular_function
gene4487	-5.9	hypothetical protein BU14_0283s0005	F:GO:0008270	F:zinc ion binding
gene8272	-4.9	alkaline phosphatase family protein	F:GO:0003674; P:GO:0008150	F:molecular_function; P:biological_process
gene6615	-4.9	hypothetical protein BU14_0033s0078	F:GO:0004386; P:GO:0006259; P:GO:0006950; P:GO:0042592; F:GO:0043167; P:GO:0051276	F:helicase activity; P:DNA metabolic process; P:response to stress; P:homeostatic process; F:ion binding; P:chromosome organization

gene13370	-4.8	ycf68 protein	C:GO:0005737; P:GO:0006399; P:GO:0006412; P:GO:0006520; F:GO:0016874; F:GO:0043167	C:cytoplasm; P:tRNA metabolic process; P:translation; P:cellular amino acid metabolic process; F:ligase activity; F:ion binding
gene13323	-4.7	ycf68 protein	C:GO:0005737; P:GO:0006399; P:GO:0006412; P:GO:0006520; F:GO:0016874; F:GO:0043167	C:cytoplasm; P:tRNA metabolic process; P:translation; P:cellular amino acid metabolic process; F:ligase activity; F:ion binding
gene13373	-3.9	Uncharacterised protein	C:GO:0009536	C:plastid
gene6715	-1.6	germin-like protein subfamily 3 member 2	C:GO:0005576; F:GO:0043167	C:extracellular region; F:ion binding
gene6114	-1.5	hypothetical protein BU14_0239s0013	C:GO:0016020; C:GO:0016021	C:membrane; C:integral component of membrane
gene10540	-1.4	alpha carbonic anhydrase 1, chloroplastic	F:GO:0016829; F:GO:0043167	F:lyase activity; F:ion binding
gene9750	-1.0	histone H3	F:GO:0003677; C:GO:0005634; C:GO:0005694; P:GO:0007059; C:GO:0032991; P:GO:0051276; P:GO:0065003; P:GO:0140014	F:DNA binding; C:nucleus; C:chromosome; P:chromosome segregation; C:protein-containing complex; P:chromosome organization; P:protein-containing complex assembly; P:mitotic nuclear division
gene2158	1.0	hypothetical protein BU14_0077s0060	C:GO:0005575	C:cellular_component
gene10503	1.0	CRTAC1 family protein	C:GO:0016020; C:GO:0016021	C:membrane; C:integral component of membrane
gene11687	1.0	zinc finger protein OZF-like	F:GO:0003674	F:molecular_function
gene10209	1.0	peptidyl-prolyl cis-trans isomerase A-like	P:GO:0006457; P:GO:0006464; F:GO:0016853	P:protein folding; P:cellular protein modification process; F:isomerase activity

gene11909	1.1	alpha-1,2-fucosyltransferase	C:GO:0005575; P:GO:0005975; F:GO:0016757	C:cellular_component; P:carbohydrate metabolic process; F:transferase activity, transferring glycosyl groups
gene6565	1.1	predicted protein	F:GO:0003824; F:GO:0016740	F:catalytic activity; F:transferase activity
gene7450	1.1	uridine phosphorylase 1 isoform X2	F:GO:0016757; P:GO:0034641; P:GO:0044281	F:transferase activity, transferring glycosyl groups; P:cellular nitrogen compound metabolic process; P:small molecule metabolic process
gene4718	1.1	hypothetical protein BU14_0176s0030	C:GO:0005575	C:cellular_component
gene9525	1.1	hypothetical protein BU14_0492s0007	C:GO:0005575	C:cellular_component
gene12890	1.2	Serine/threonine protein kinase	P:GO:0006464; F:GO:0016301; F:GO:0043167	P:cellular protein modification process; F:kinase activity; F:ion binding
gene5519	1.2	hypothetical protein BU14_0215s0012	C:GO:0005575	C:cellular_component
gene10824	1.2	calcium/proton exchanger	C:GO:0005622; F:GO:0022857	C:intracellular; F:transmembrane transporter activity
gene9591	1.2	hypothetical protein BU14_0500s0001	P:GO:0006464; P:GO:0007049; F:GO:0019899	P:cellular protein modification process; P:cell cycle; F:enzyme binding
gene5068	1.2	protein kinase	C:GO:0016020; C:GO:0016021; F:GO:0016301; P:GO:0016310	C:membrane; C:integral component of membrane; F:kinase activity; P:phosphorylation
gene9310	1.2	hypothetical protein BU14_0466s0008	F:GO:0003700; P:GO:0009058; P:GO:0034641	F:DNA binding transcription factor activity; P:biosynthetic process; P:cellular

				nitrogen compound metabolic process
gene268	1.2	Di-copper centre-containing protein	C:GO:0005575; P:GO:0008150	C:cellular_component; P:biological_process
gene6602	1.3	hypothetical protein BU14_0263s0016	C:GO:0016020; C:GO:0016021	C:membrane; C:integral component of membrane
gene6894	1.3	hypothetical protein BU14_0283s0014	F:GO:0003674	F:molecular_function
gene3974	1.3	hypothetical protein BU14_0144s0021	P:GO:0006629	P:lipid metabolic process
gene277	1.3	Cyclin-P3-1	P:GO:0006464; P:GO:0007049; F:GO:0019899	P:cellular protein modification process; P:cell cycle; F:enzyme binding
gene11771	1.4	mitogen-activated protein kinase 17 isoform X2	P:GO:0006464; F:GO:0016301; F:GO:0043167	P:cellular protein modification process; F:kinase activity; F:ion binding
gene9486	1.4	CRTAC1 family protein	C:GO:0016020; C:GO:0016021	C:membrane; C:integral component of membrane
gene7927	1.4	hypothetical protein BU14_0341s0007	F:GO:0003674	F:molecular_function
gene11919	1.4	iron ABC transporter substrate-binding protein	P:GO:0006091; F:GO:0016491	P:generation of precursor metabolites and energy; F:oxidoreductase activity
gene3168	1.4	zinc finger protein OZF-like	F:GO:0003674	F:molecular_function
gene7984	1.5	respiratory burst oxidase-like protein	C:GO:0005575; F:GO:0016491	C:cellular_component; F:oxidoreductase activity
gene4876	1.5	polyamine oxidase-like	F:GO:0016491	F:oxidoreductase activity
gene3641	1.5	zinc finger protein 239-like	F:GO:0003674	F:molecular_function
gene4396	1.8	zinc finger protein 708-like	F:GO:0003674	F:molecular_function

gene4695	1.8	calcium/calmodulin-dependent protein kinase type IV-like	P:GO:0006464; F:GO:0016301; F:GO:0043167	P:cellular protein modification process; F:kinase activity; F:ion binding
gene2054	1.8	Monogalactosyldiacylglycerol synthase fragment, family GT28	F:GO:0016757	F:transferase activity, transferring glycosyl groups
gene6219	1.9	hypothetical protein BU14_0244s0008	F:GO:0003700; P:GO:0009058; P:GO:0034641	F:DNA binding transcription factor activity; P:biosynthetic process; P:cellular nitrogen compound metabolic process
gene13076	1.9	RNA polymerase sigma factor, RpoD/SigA family	F:GO:0003677; F:GO:0003700; P:GO:0006351; P:GO:0006352; P:GO:0006355; F:GO:0016987; P:GO:2000142	F:DNA binding; F:DNA binding transcription factor activity; P:transcription, DNA-templated; P:DNA-templated transcription, initiation; P:regulation of transcription, DNA-templated; F:sigma factor activity; P:regulation of DNA-templated transcription, initiation
gene9464	1.9	endoplasmic reticulum oxidoreductin-2-like	C:GO:0005783; P:GO:0006457; P:GO:0006464; F:GO:0016491; F:GO:0016853	C:endoplasmic reticulum; P:protein folding; P:cellular protein modification process; F:oxidoreductase activity; F:isomerase activity
gene9631	2.0	Endoribonuclease XendoU	F:GO:0004518	F:nuclease activity
gene9160	2.2	Transcription factor YY2	F:GO:0003674	F:molecular_function
gene4525	2.3	hypothetical protein BU14_0166s0032	F:GO:0008233	F:peptidase activity
gene12686	2.3	gelsolin-like protein 2	C:GO:0005622; P:GO:0007010; F:GO:0008092; C:GO:0043226	C:intracellular; P:cytoskeleton organization;

				F:cytoskeletal protein binding; C:organelle
gene12149	3.0	phycobilisome linker polypeptide	C:GO:0009579; P:GO:0015979; C:GO:0032991	C:thylakoid; P:photosynthesis; C:protein-containing complex
gene1902	3.6	hypothetical protein BU14_0072s0103, partial	F:GO:0003674	F:molecular_function
gene191	4.6	hypothetical protein BU14_0022s0012	C:GO:0016020; C:GO:0016021	C:membrane; C:integral component of membrane
gene6154	5.6	hypothetical protein BU14_0240s0022	C:GO:0005575	C:cellular_component
gene6589	7.0	hypothetical protein BU14_0263s0003	C:GO:0005575	C:cellular_component

Supplement Table 4: DEGs between desiccation and dehydration with gene name, Log_2fold change (Log2FC), putative protein name and putative GO functions

Gene	Log2FC	Protein name	GO term	Go Function
gene6589	-6.4	hypothetical protein BU14_0263s0003	C:GO:0005575	C:cellular_component
gene2909	-3.5	hypothetical protein BU14_0238s0011	C:GO:0005575	C:cellular_component
gene6082	-3.3	hypothetical protein BU14_0238s0011	C:GO:0005575	C:cellular_component
gene7632	-1.8	hypothetical protein BU14_0131s0014	F:GO:0003674; P:GO:0006259	F:molecular_function; P:DNA metabolic process
gene2463	-1.6	hypothetical protein BU14_0713s0004	C:GO:0005575	C:cellular_component
gene1902	-1.6	hypothetical protein BU14_0072s0103	F:GO:0003674	F:molecular_function
gene6795	-1.4	hypothetical protein BU14_0166s0020	F:GO:0003677; P:GO:0009058; F:GO:0016779; P:GO:0034641	F:DNA binding; P:biosynthetic process; F:nucleotidyltransferase activity; P:cellular

				nitrogen compound metabolic process
gene1859	-1.3	copper amine oxidase	F:GO:0016491; F:GO:0043167	F:oxidoreductase activity;F:ion binding
gene9160	-1.2	Transcription factor YY2	F:GO:0003674	F:molecular_function
gene1516	-1.2	alcohol dehydrogenase	F:GO:0016491	F:oxidoreductase activity
gene6894	-1.1	hypothetical protein BU14_0283s0014	F:GO:0003674	F:molecular_function
gene7415	-1.1	hypothetical protein BU14_0313s0013	C:GO:0005575	C:cellular_component
gene3169	-1.0	hypothetical protein BU14_0111s0045	C:GO:0005575	C:cellular_component
gene10680	-1.0	arabinogalactan endo-1-4-beta-galactosidase	F:GO:0016798	F:hydrolase activity acting on glycosyl bonds
gene99	1.7	hypothetical protein BU14_0014s0100	F:GO:0008233	F:peptidase activity
gene2799	2.1	MEDEA	P:GO:0000003; F:GO:0003677; C:GO:0005634; P:GO:0009058; P:GO:0034641; P:GO:0048856	P:reproduction; F:DNA binding; C:nucleus; P:biosynthetic process; P:cellular nitrogen compound metabolic process; P:anatomical structure development
gene7738	2.4	hypothetical protein BU14_0331s0006	C:GO:0016020; C:GO:0016021	C:membrane; C:integral component of membrane
gene6885	4.6	hypothetical protein BU14_0283s0005	P:GO:0006508; F:GO:0008234; F:GO:0008270	P:proteolysis; F:cysteine-type peptidase activity; F:zinc ion binding
gene13326	5.2	Uncharacterised protein	C:GO:0009536	C:plastid
gene13373	5.4	Uncharacterised protein	C:GO:0009536	C:plastid
gene8272	5.7	alkaline phosphatase family protein	F:GO:0003674; P:GO:0008150	F:molecular_function; P:biological_process

gene13323	6.3	ycf68 protein	C:GO:0005737; P:GO:0006399; P:GO:0006412; P:GO:0006520; F:GO:0016874; F:GO:0043167	C:cytoplasm; P:tRNA metabolic process; P:translation; P:cellular amino acid metabolic process; F:ligase activity; F:ion binding
gene13370	6.4	ycf68 protein	C:GO:0005737; P:GO:0006399; P:GO:0006412; P:GO:0006520; F:GO:0016874; F:GO:0043167	C:cytoplasm; P:tRNA metabolic process; P:translation; P:cellular amino acid metabolic process; F:ligase activity; F:ion binding
gene4487	7.4	hypothetical protein BU14_0283s0005	F:GO:0008270	F:zinc ion binding

Supplement Table 5: DEGs between desiccation and fresh with gene name, Log_2fold change (Log2FC), putative protein name and putative GO functions

Gene	Log2FC	Protein name	GO term	Go Function
gene8115	-7.2	hypothetical protein BU14_0357s0016	F:GO:0003674	F:molecular_function
gene4684	-5.1	Bile salt-activated lipase	F:GO:0003674	F:molecular_function
gene2463	-2.5	hypothetical protein BU14_0713s0004	C:GO:0005575	C:cellular_component
gene2909	-2.1	hypothetical protein BU14_0238s0011	C:GO:0005575	C:cellular_component
gene6715	-1.7	germin-like protein subfamily 3 member 2	C:GO:0005576; F:GO:0043167	C:extracellular region; F:ion binding
gene10785	-1.5	hypothetical protein BU14_0148s0028	C:GO:0005575	C:cellular_component
gene2932	-1.5	ubiquitin-conjugating enzyme E2 T	F:GO:0003674; C:GO:0005622; P:GO:0006464; P:GO:0006950	F:molecular_function; C:intracellular; P:cellular protein modification process; P:response to stress

gene6795	-1.4	hypothetical protein BU14_0166s0020	F:GO:0003677 ; P:GO:0009058 ; F:GO:0016779 ; P:GO:0034641	F:DNA binding; P:biosynthetic process; F:nucleotidyltransferase activity; P:cellular nitrogen compound metabolic process
gene1611	-1.2	hypothetical protein BU14_0068s0010	F:GO:0005198 ; C:GO:0005575	F:structural molecule activity; C:cellular_component
gene446	-1.1	hypothetical protein BU14_0025s0070	C:GO:0005575	C:cellular_component
gene8926	-1.1	ABC transporter I family member 10, chloroplastic	F:GO:0016887 ; F:GO:0043167	F:ATPase activity; F:ion binding
gene4069	-1.1	hypothetical protein BU14_0148s0028	C:GO:0005575	C:cellular_component
gene8032	-1.0	hypothetical protein BU14_0352s0006	C:GO:0005575	C:cellular_component
gene10004	1.0	hypothetical protein BU14_0537s0017	F:GO:0004252 ; P:GO:0006508	F:serine-type endopeptidase activity; P:proteolysis
gene9525	1.0	hypothetical protein BU14_0492s0007	C:GO:0005575	C:cellular_component
gene3168	1.0	zinc finger protein OZF-like	F:GO:0003674	F:molecular_function
gene1710	1.0	beta-carotene isomerase D27, chloroplastic	F:GO:0005506 ; C:GO:0009507 ; P:GO:0010223 ; F:GO:0016853 ; F:GO:0016859 ; P:GO:1901601	F:iron ion binding; C:chloroplast; P:secondary shoot formation; F:isomerase activity; F:cis-trans isomerase activity; P:strigolactone biosynthetic process
gene5246	1.1	L-histidine N(alpha)-methyltransferase	F:GO:0004812 ; P:GO:0006479 ; F:GO:0008168	F:aminoacyl-tRNA ligase activity; P:protein methylation; F:methyltransferase activity; F:protein methyltransferase activity; F:transferase activity;

			; F:GO:0008276 ; F:GO:0016740 ; F:GO:0030745 ; P:GO:0032259 ; P:GO:0052699	F:dimethylhistidine N-methyltransferase activity; P:methylation; P:ergothioneine biosynthetic process
gene7927	1.1	hypothetical protein BU14_0341s0007	F:GO:0003674	F:molecular_function
gene1152	1.1	nucleolar RNA helicase 2-like isoform X1	F:GO:0003723 ; F:GO:0004386 ; C:GO:0005634 ; F:GO:0043167	F:RNA binding; F:helicase activity; C:nucleus; F:ion binding
gene6141	1.1	predicted protein	C:GO:0005575	C:cellular_component
gene12356	1.1	peroxisomal membrane protein PMP22	C:GO:0005575	C:cellular_component
gene11771	1.1	mitogen-activated protein kinase 17 isoform X2	P:GO:0006464 ; F:GO:0016301 ; F:GO:0043167	P:cellular protein modification process; F:kinase activity; F:ion binding
gene3460	1.1	aarF domain-containing kinase	P:GO:0006464 ; F:GO:0016301 ; F:GO:0043167	P:cellular protein modification process; F:kinase activity; F:ion binding
gene7984	1.1	respiratory burst oxidase-like protein	C:GO:0005575 ; F:GO:0016491	C:cellular_component; F:oxidoreductase activity
gene9310	1.1	hypothetical protein BU14_0466s0008	F:GO:0003700 ; P:GO:0009058 ; P:GO:0034641	F:DNA binding transcription factor activity; P:biosynthetic process; P:cellular nitrogen compound metabolic process
gene11644	1.2	hypothetical protein BU14_1241s0001	F:GO:0003700 ; P:GO:0009058	F:DNA binding transcription factor activity; P:biosynthetic

			; P:GO:0034641	process; P:cellular nitrogen compound metabolic process
gene3641	1.2	zinc finger protein 239-like	F:GO:0003674	F:molecular_function
gene9841	1.2	protein S-acyltransferase 8-like	C:GO:0016020 ; C:GO:0016021 ; F:GO:0016740 ; F:GO:0016746 ; F:GO:0019706	C:membrane; C:integral component of membrane; F:transferase activity; F:transferase activity, transferring acyl groups; F:protein-cysteine S-palmitoyltransferase activity
gene4718	1.2	hypothetical protein BU14_0176s0030	C:GO:0005575	C:cellular_component
gene4695	1.3	calcium/calmodulin-dependent protein kinase type IV-like	P:GO:0006464 ; F:GO:0016301 ; F:GO:0043167	P:cellular protein modification process; F:kinase activity; F:ion binding
gene11919	1.3	iron ABC transporter substrate-binding protein	P:GO:0006091 ; F:GO:0016491	P:generation of precursor metabolites and energy; F:oxidoreductase activity
gene4396	1.4	zinc finger protein 708-like	F:GO:0003674	F:molecular_function
gene5254	1.4	hypothetical protein BU14_0202s0013	C:GO:0005575	C:cellular_component
gene12686	1.5	gelsolin-like protein 2	C:GO:0005622 ; P:GO:0007010 ; F:GO:0008092 ; C:GO:0043226	C:intracellular; P:cytoskeleton organization; F:cytoskeletal protein binding; C:organelle
gene99	1.5	hypothetical protein BU14_0014s0100	F:GO:0008233	F:peptidase activity
gene2681	1.7	class I SAM-dependent methyltransferase	F:GO:0003674	F:molecular_function
gene1902	2.0	hypothetical protein BU14_0072s0103, partial	F:GO:0003674	F:molecular_function

gene2168	2.2	hypothetical protein BU14_0079s0007	F:GO:0008233	F:peptidase activity
gene4525	2.3	hypothetical protein BU14_0166s0032	F:GO:0008233	F:peptidase activity
gene12423	2.3	ER membrane protein complex subunit 6	F:GO:0043167	F:ion binding
gene2799	3.9	MEDEA	P:GO:0000003; F:GO:0003677; C:GO:0005634; P:GO:0009058; P:GO:0034641; P:GO:0048856	P:reproduction; F:DNA binding; C:nucleus; P:biosynthetic process; P:cellular nitrogen compound metabolic process; P:anatomical structure development
gene6154	4.0	hypothetical protein BU14_0240s0022	C:GO:0005575	C:cellular_component
gene6631	4.6	hypothetical protein BU14_0265s0010	F:GO:0003676	F:nucleic acid binding

Supplement Table 6: DEGs between rehydration and fresh with gene name, Log_2fold change (Log2FC), putative protein name and putative GO functions

Gene	Log2 FC	Protein name	GO term	Go Function
gene8115	-6.5	hypothetical protein BU14_0357s0016	F:GO:0003674	F:molecular_function
gene4684	-6.3	Bile salt-activated lipase	F:GO:0003674	F:molecular_function
gene4487	-5.9	hypothetical protein BU14_0283s0005	F:GO:0008270	F:zinc ion binding
gene6615	-5.8	hypothetical protein BU14_0033s0078	F:GO:0004386; P:GO:0006259; P:GO:0006950; P:GO:0042592; F:GO:0043167	F:helicase activity; P:DNA metabolic process; P:response to stress; P:homeostatic process; F:ion binding; P:chromosome organization

			; P:GO:0051276	
gene4667	-1.7	elongation factor G	F:GO:0003924 ; P:GO:0006412 ; F:GO:0043167	F:GTPase activity; P:translation; F:ion binding
gene2953	-1.7	hypothetical protein BU14_0106s0006	P:GO:0005975 ; F:GO:0016798	P:carbohydrate metabolic process; F:hydrolase activity, acting on glycosyl bonds
gene5613	-1.5	STE family protein kinase	P:GO:0000278 ; C:GO:0005737 ; P:GO:0006464 ; P:GO:0006950 ; P:GO:0007165 ; P:GO:0008219 ; F:GO:0016301 ; F:GO:0043167	P:mitotic cell cycle; C:cytoplasm; P:cellular protein modification process; P:response to stress; P:signal transduction; P:cell death; F:kinase activity; F:ion binding
gene13084	-1.4	DNA polymerase epsilon catalytic subunit A	C:GO:0000228 ; P:GO:0000278 ; F:GO:0003677 ; P:GO:0006259 ; P:GO:0006950 ; P:GO:0009058 ; F:GO:0016779 ; C:GO:0032991	C:nuclear chromosome; P:mitotic cell cycle; F:DNA binding; P:DNA metabolic process; P:response to stress; P:biosynthetic process; F:nucleotidyltransferase activity; C:protein-containing complex
gene11490	-1.4	hypothetical protein BU14_1041s0003	F:GO:0043167	F:ion binding
gene6114	-1.3	hypothetical protein BU14_0239s0013	C:GO:0016020 ; C:GO:0016021	C:membrane; C:integral component of membrane

gene9570	-1.3	protein CHROMATIN REMODELING 24	F:GO:0043167	F:ion binding
gene8145	-1.3	hypothetical protein BU14_0361s0005	P:GO:0005975 ; F:GO:0016301 ; P:GO:0044281	P:carbohydrate metabolic process; F:kinase activity; P:small molecule metabolic process
gene10330	-1.3	hypothetical protein BU14_0608s0014	C:GO:0000228 ; P:GO:0006950 ; P:GO:0007049 ; C:GO:0032991	C:nuclear chromosome; P:response to stress; P:cell cycle; C:protein-containing complex
gene10540	-1.2	alpha carbonic anhydrase 1, chloroplastic	F:GO:0016829 ; F:GO:0043167	F:lyase activity; F:ion binding
gene8152	-1.2	hypothetical protein BU14_0361s0012	P:GO:0005975 ; F:GO:0016301 ; P:GO:0044281	P:carbohydrate metabolic process; F:kinase activity; P:small molecule metabolic process
gene12887	-1.2	FAD:protein FMN transferase	C:GO:0005886 ; F:GO:0010181 ; F:GO:0016740 ; P:GO:0017013 ; F:GO:0046872 ; F:GO:0048037	C:plasma membrane; F:FMN binding; F:transferase activity; P:protein flavinylation; F:metal ion binding; F:cofactor binding
gene2779	-1.1	hypothetical protein BU14_0099s0022	F:GO:0003723 ; F:GO:0004386 ; F:GO:0043167	F:RNA binding; F:helicase activity; F:ion binding
gene6586	-1.1	Centromere-associated protein E	F:GO:0003674 ; C:GO:0005622 ; C:GO:0043226	F:molecular_function; C:intracellular; C:organelle

gene9173	-1.1	hypothetical protein BU14_0453s0012	C:GO:0005575 ; F:GO:0043167	C:cellular_component; F:ion binding
gene7433	-1.1	DNA topoisomerase 2	F:GO:0003677 ; P:GO:0006259 ; P:GO:0007049 ; P:GO:0007059 ; F:GO:0016853 ; F:GO:0016887 ; F:GO:0043167 ; P:GO:0051276	F:DNA binding; P:DNA metabolic process; P:cell cycle; P:chromosome segregation; F:isomerase activity; F:ATPase activity; F:ion binding; P:chromosome organization
gene13000	-1.1	kinetochore protein NUF2-like protein	C:GO:0005694 ; C:GO:0032991	C:chromosome; C:protein-containing complex
gene10873	-1.1	DNA replication licensing factor MCM5	C:GO:0000228 ; F:GO:0003677 ; F:GO:0004386 ; C:GO:0005654 ; C:GO:0005737 ; P:GO:0006259 ; P:GO:0007049 ; P:GO:0009058 ; C:GO:0032991 ; F:GO:0043167 ; P:GO:0051276	C:nuclear chromosome; F:DNA binding; F:helicase activity; C:nucleoplasm; C:cytoplasm; P:DNA metabolic process; P:cell cycle; P:biosynthetic process; C:protein-containing complex; F:ion binding; P:chromosome organization
gene7067	-1.1	urea carboxylase	F:GO:0043167	F:ion binding
gene10456	-1.1	oxidoreductase, FAD-binding	F:GO:0016491	F:oxidoreductase activity

gene10967	-1.1	high-affinity iron permease	C:GO:0005886 ; F:GO:0022857 ; C:GO:0032991	C:plasma membrane; F:transmembrane transporter activity; C:protein-containing complex
gene12613	-1.0	DNA replication licensing factor mcm7	F:GO:0003677 ; F:GO:0004386 ; C:GO:0005634 ; C:GO:0005829 ; P:GO:0006259 ; P:GO:0007049 ; P:GO:0007165 ; P:GO:0009058 ; C:GO:0032991 ; F:GO:0043167 ; P:GO:0051276	F:DNA binding; F:helicase activity; C:nucleus; C:cytosol; P:DNA metabolic process; P:cell cycle; P:signal transduction; P:biosynthetic process; C:protein-containing complex; F:ion binding; P:chromosome organization
gene9111	-1.0	DNA replication licensing factor MCM3	C:GO:0000228 ; F:GO:0003677 ; F:GO:0004386 ; P:GO:0006259 ; P:GO:0009058 ; C:GO:0032991 ; F:GO:0043167	C:nuclear chromosome; F:DNA binding; F:helicase activity; P:DNA metabolic process; P:biosynthetic process; C:protein-containing complex; F:ion binding
gene12280	-1.0	DNA repair protein rad32	P:GO:0000003 ; F:GO:0004518 ; C:GO:0005634 ; P:GO:0006259 ; P:GO:0006950	P:reproduction; F:nuclease activity; C:nucleus; P:DNA metabolic process; P:response to stress; P:cell cycle; C:protein-containing complex

			; P:GO:0007049 ; C:GO:0032991	
gene783	-1.0	hyalin	P:GO:0007154 ; P:GO:0007229 ; C:GO:0016021	P:cell communication; P:integrin-mediated signaling pathway; C:integral component of membrane
gene487	1.0	hypothetical protein BU14_0027s0035	C:GO:0005575	C:cellular_component
gene6620	1.0	hypothetical protein BU14_0264s0014	C:GO:0016020 ; C:GO:0016021	C:membrane; C:integral component of membrane
gene12120	1.0	putative phospholipid-transporting ATPase 9	C:GO:0005575 ; P:GO:0006810 ; F:GO:0016887 ; F:GO:0043167 ; P:GO:0061024	C:cellular_component; P:transport; F:ATPase activity; F:ion binding; P:membrane organization
gene2722	1.0	hypothetical protein BU14_0098s0038	C:GO:0005575	C:cellular_component
gene7630	1.0	DEAD-box ATP-dependent RNA helicase 7	F:GO:0003723 ; F:GO:0004386 ; C:GO:0005634 ; C:GO:0005737 ; F:GO:0043167	F:RNA binding; F:helicase activity; C:nucleus; C:cytoplasm; F:ion binding
gene10607	1.0	histidinol-phosphate aminotransferase	P:GO:0009058 ; F:GO:0043167	P:biosynthetic process; F:ion binding
gene9310	1.0	hypothetical protein BU14_0466s0008	F:GO:0003700 ; P:GO:0009058 ; P:GO:0034641	F:DNA binding transcription factor activity; P:biosynthetic process; P:cellular nitrogen compound metabolic process

gene9498	1.0	hypothetical protein BU14_0490s0010	F:GO:0046983	F:protein dimerization activity
gene13031	1.0	ribosomal protein L1	C:GO:0032991	C:protein-containing complex
gene10946	1.0	predicted protein	C:GO:0005730 ; C:GO:0032991 ; P:GO:0034641 ; P:GO:0042254	C:nucleolus; C:protein-containing complex; P:cellular nitrogen compound metabolic process; P:ribosome biogenesis
gene8849	1.0	ATPase family AAA domain-containing protein 3A	F:GO:0043167	F:ion binding
gene8498	1.0	hypothetical protein BU14_0393s0017	F:GO:0003674	F:molecular_function
gene8037	1.1	alpha-galactosidase	P:GO:0005975 ; F:GO:0016798	P:carbohydrate metabolic process; F:hydrolase activity, acting on glycosyl bonds
gene5701	1.1	beta and beta-prime subunits of DNA dependent RNA-polymerase	F:GO:0003677 ; C:GO:0005634 ; P:GO:0009058 ; F:GO:0016779 ; P:GO:0034641 ; F:GO:0043167	F:DNA binding; C:nucleus; P:biosynthetic process; F:nucleotidyltransferase activity; P:cellular nitrogen compound metabolic process; F:ion binding
gene7258	1.1	hypothetical protein BU14_0303s0006	C:GO:0005575	C:cellular_component
gene9941	1.1	Polyribonucleotide nucleotidyltransferase 2, mitochondrial	F:GO:0003723 ; C:GO:0005737 ; F:GO:0016779 ; P:GO:0034655	F:RNA binding; C:cytoplasm; F:nucleotidyltransferase activity; P:nucleobase-containing compound catabolic process
gene12172	1.1	hypothetical conserved protein	C:GO:0005575	C:cellular_component
gene10087	1.1	hypothetical protein BU14_0560s0003	F:GO:0003674	F:molecular_function

gene6749	1.1	carbonic anhydrase	F:GO:0016829 ; F:GO:0043167	F:lyase activity; F:ion binding
gene12425	1.1	hypothetical protein BU14_2083s0001	C:GO:0005575	C:cellular_component
gene9265	1.1	hypothetical protein BU14_0463s0012	C:GO:0005737 ; P:GO:0006457 ; F:GO:0043167	C:cytoplasm; P:protein folding; F:ion binding
gene4371	1.1	Na(+)/H(+) antiporter NhaD-like	C:GO:0005575 ; P:GO:0055085	C:cellular_component; P:transmembrane transport
gene10290	1.1	Glycosyltransferase, family GT90	C:GO:0016020 ; C:GO:0016021 ; F:GO:0016740	C:membrane; C:integral component of membrane; F:transferase activity
gene1023	1.1	charged multivesicular body protein 2A	P:GO:0007034	P:vacuolar transport
gene6219	1.1	hypothetical protein BU14_0244s0008	F:GO:0003700 ; P:GO:0009058 ; P:GO:0034641	F:DNA binding transcription factor activity; P:biosynthetic process; P:cellular nitrogen compound metabolic process
gene10674	1.1	LOW QUALITY PROTEIN: hypothetical protein BU14_0686s0002	F:GO:0016209 ; F:GO:0016491 ; P:GO:0055114 ; P:GO:0098869	F:antioxidant activity; F:oxidoreductase activity; P:oxidation-reduction process; P:cellular oxidant detoxification
gene8253	1.1	pyrroline-5-carboxylate reductase	P:GO:0008150	P:biological_process
gene7789	1.1	major intrinsic protein AQP1	C:GO:0005575 ; F:GO:0022857	C:cellular_component; F:transmembrane transporter activity
gene3394	1.1	pre-60S factor REI1	F:GO:0043167	F:ion binding
gene6685	1.1	cyclic pyranopterin monophosphate synthase,	C:GO:0005623 ; P:GO:0009058 ;	C:cell; P:biosynthetic process; C:protein-containing complex;

		mitochondrial isoform X1	C:GO:0032991 ; F:GO:0043167 ; P:GO:0051186	F:ion binding; P:cofactor metabolic process
gene9667	1.1	carbohydrate-binding module family 48 protein	C:GO:0005622 ; P:GO:0006464 ; P:GO:0007165 ; C:GO:0032991	C:intracellular; P:cellular protein modification process; P:signal transduction; C:protein-containing complex
gene6602	1.1	hypothetical protein BU14_0263s0016	C:GO:0016020 ; C:GO:0016021	C:membrane; C:integral component of membrane
gene8448	1.1	glutathione S-transferase Mu 1-like	F:GO:0003674	F:molecular_function
gene7450	1.1	uridine phosphorylase 1 isoform X2	F:GO:0016757 ; P:GO:0034641 ; P:GO:0044281	F:transferase activity, transferring glycosyl groups; P:cellular nitrogen compound metabolic process; P:small molecule metabolic process
gene9160	1.1	Transcription factor YY2	F:GO:0003674	F:molecular_function
gene5173	1.1	Mitogen-activated protein kinase 15-2 MPK15-2	P:GO:0006464 ; F:GO:0016301 ; F:GO:0043167	P:cellular protein modification process; F:kinase activity; F:ion binding
gene5700	1.1	beta and beta-prime subunits of DNA dependent RNA-polymerase	F:GO:0003677 ; C:GO:0005730 ; P:GO:0009058 ; F:GO:0016779 ; C:GO:0032991 ; P:GO:0034641 ; F:GO:0043167	F:DNA binding; C:nucleolus; P:biosynthetic process; F:nucleotidyltransferase activity; C:protein-containing complex; P:cellular nitrogen compound metabolic process; F:ion binding
gene854	1.1	ferritin-3, chloroplastic	C:GO:0005623 ; P:GO:0006810	C:cell; P:transport; F:oxidoreductase activity;

			; F:GO:0016491 ; P:GO:0042592 ; F:GO:0043167	P:homeostatic process; F:ion binding
gene9629	1.1	hypothetical protein BU14_0502s0001	C:GO:0005575	C:cellular_component
gene8965	1.1	predicted protein	F:GO:0016887 ; F:GO:0043167	F:ATPase activity; F:ion binding
gene11468	1.2	P-loop containing nucleoside triphosphate hydrolase protein	F:GO:0016887 ; F:GO:0043167	F:ATPase activity; F:ion binding
gene1194	1.2	hypothetical protein BU14_0053s0021	C:GO:0005575	C:cellular_component
gene11245	1.2	predicted protein	F:GO:0004386 ; C:GO:0005622 ; F:GO:0043167	F:helicase activity; C:intracellular; F:ion binding
gene12360	1.2	probable S-adenosylmethionine carrier 2, chloroplastic isoform X2	C:GO:0005575	C:cellular_component
gene11258	1.2	ankryin	F:GO:0003824 ; P:GO:0009116 ; C:GO:0016020 ; C:GO:0016021	F:catalytic activity; P:nucleoside metabolic process; C:membrane; C:integral component of membrane
gene13030	1.2	pre-rRNA-processing protein PNO1	F:GO:0003723 ; C:GO:0005730 ; P:GO:0034641 ; P:GO:0042254 ; F:GO:0043167 ; F:GO:0051082	F:RNA binding; C:nucleolus; P:cellular nitrogen compound metabolic process; P:ribosome biogenesis; F:ion binding; F:unfolded protein binding; P:protein-containing complex assembly

			; P:GO:0065003	
gene5857	1.2	ABC2 type transporter superfamily protein	C:GO:0005575 ; F:GO:0016887 ; F:GO:0043167	C:cellular_component; F:ATPase activity; F:ion binding
gene9024	1.2	folate/biopterin family MFS transporter	C:GO:0005575 ; F:GO:0022857	C:cellular_component; F:transmembrane transporter activity
gene7431	1.2	cytochrome P450	F:GO:0016491 ; F:GO:0043167	F:oxidoreductase activity; F:ion binding
gene3738	1.2	hypothetical protein BU14_0134s0034	C:GO:0005623 ; P:GO:0042592	C:cell; P:homeostatic process
gene5345	1.2	fatty acid desaturase 4, chloroplastic-like	P:GO:0006629 ; P:GO:0009058 ; C:GO:0009536 ; F:GO:0016491 ; P:GO:0044281	P:lipid metabolic process; P:biosynthetic process; C:plastid; F:oxidoreductase activity; P:small molecule metabolic process
gene4219	1.3	acyl-CoA desaturase	C:GO:0005575 ; P:GO:0006629 ; P:GO:0009058 ; F:GO:0016491 ; F:GO:0043167 ; P:GO:0044281	C:cellular_component; P:lipid metabolic process; P:biosynthetic process; F:oxidoreductase activity; F:ion binding; P:small molecule metabolic process
gene12988	1.3	dead deah box rna	F:GO:0004386 ; C:GO:0005730 ; C:GO:0005737 ; F:GO:0016887 ; P:GO:0034641	F:helicase activity; C:nucleolus; C:cytoplasm; F:ATPase activity; P:cellular nitrogen compound metabolic process; P:ribosome biogenesis; F:ion binding

			; P:GO:0042254 ; F:GO:0043167	
gene3974	1.3	hypothetical protein BU14_0144s0021	P:GO:0006629	P:lipid metabolic process
gene6314	1.3	emopamil-binding protein	C:GO:0005575	C:cellular_component
gene7026	1.3	RNA polymerase sigma factor RpoD	F:GO:0003700 ; P:GO:0009058 ; P:GO:0034641	F:DNA binding transcription factor activity; P:biosynthetic process; P:cellular nitrogen compound metabolic process
gene277	1.3	Cyclin-P3-1	P:GO:0006464 ; P:GO:0007049 ; F:GO:0019899	P:cellular protein modification process; P:cell cycle; F:enzyme binding
gene9732	1.3	Ferredoxin	P:GO:0006091 ; F:GO:0016491	P:generation of precursor metabolites and energy; F:oxidoreductase activity
gene10717	1.4	hypothetical protein BU14_0692s0013	C:GO:0005575	C:cellular_component
gene9867	1.4	mitochondrial carrier domain-containing protein	C:GO:0005575 ; P:GO:0055085	C:cellular_component; P:transmembrane transport
gene11087	1.4	CRTAC1 family protein	C:GO:0016020 ; C:GO:0016021	C:membrane; C:integral component of membrane
gene12180	1.4	hypothetical protein BU14_1770s0001, partial	F:GO:0016740	F:transferase activity
gene1058	1.4	tyrosine transporter	C:GO:0005886 ; C:GO:0009536 ; F:GO:0022857	C:plasma membrane; C:plastid; F:transmembrane transporter activity
gene466	1.4	S-adenosylmethionine carrier 1, chloroplastic/mitochondrial isoform X1	C:GO:0005737 ; P:GO:0006810 ; C:GO:0043226	C:cytoplasm; P:transport; C:organelle

gene11207	1.4	predicted protein	C:GO:0005575	C:cellular_component
gene3168	1.4	zinc finger protein OZF-like	F:GO:0003674	F:molecular_function
gene3747	1.4	flotillin-2 isoform X2	P:GO:0002376 ; C:GO:0005623 ; P:GO:0006810 ; P:GO:0006950 ; P:GO:0007165 ; P:GO:0007267 ; P:GO:0019748	P:immune system process; C:cell; P:transport; P:response to stress; P:signal transduction; P:cell-cell signaling; P:secondary metabolic process
gene693	1.4	putative ABC1 protein At2g40090	C:GO:0005575	C:cellular_component
gene3653	1.5	hypothetical protein BU14_0130s0016	F:GO:0043167	F:ion binding
gene11771	1.5	mitogen-activated protein kinase 17 isoform X2	P:GO:0006464 ; F:GO:0016301 ; F:GO:0043167	P:cellular protein modification process; F:kinase activity; F:ion binding
gene2098	1.5	glutathione S-transferase	F:GO:0004364 ; F:GO:0016740	F:glutathione transferase activity; F:transferase activity
gene3090	1.5	hypothetical protein BU14_0109s0016	F:GO:0005509	F:calcium ion binding
gene4199	1.5	catalase family protein	P:GO:0006950 ; F:GO:0016491 ; F:GO:0043167	P:response to stress; F:oxidoreductase activity; F:ion binding
gene8480	1.5	phytoene synthase	C:GO:0005575 ; P:GO:0006629 ; P:GO:0009058 ; F:GO:0016765	C:cellular_component; P:lipid metabolic process; P:biosynthetic process; F:transferase activity, transferring alkyl or aryl (other than methyl) groups; P:small molecule metabolic process

			; P:GO:0044281	
gene6656	1.5	multidrug resistance protein homolog 49	C:GO:0005575 ; F:GO:0016887 ; F:GO:0022857 ; F:GO:0043167	C:cellular_component; F:ATPase activity; F:transmembrane transporter activity; F:ion binding
gene9630	1.5	hypothetical protein BU14_0502s0002	C:GO:0005575	C:cellular_component
gene12018	1.5	serine/threonine-protein kinase HT1-like	P:GO:0006464 ; F:GO:0016301 ; F:GO:0043167	P:cellular protein modification process; F:kinase activity; F:ion binding
gene11038	1.5	D-alanine--D-alanine ligase	F:GO:0016874 ; F:GO:0043167	F:ligase activity; F:ion binding
gene11631	1.6	hypothetical protein BU14_1226s0001	C:GO:0005575	C:cellular_component
gene4839	1.6	hypothetical protein BU14_0183s0003	F:GO:0008233	F:peptidase activity
gene6412	1.6	hypothetical protein BU14_0256s0035	F:GO:0003674	F:molecular_function
gene5076	1.6	catalase	P:GO:0006950 ; F:GO:0016491 ; F:GO:0043167	P:response to stress; F:oxidoreductase activity; F:ion binding
gene9657	1.6	hypothetical protein BU14_0504s0015	C:GO:0005623 ; P:GO:0042592	C:cell; P:homeostatic process
gene1227	1.6	hypothetical protein BU14_0055s0007	C:GO:0005575	C:cellular_component
gene3460	1.7	aarF domain-containing kinase	P:GO:0006464 ; F:GO:0016301 ; F:GO:0043167	P:cellular protein modification process; F:kinase activity; F:ion binding

gene10893	1.8	LOW QUALITY PROTEIN: hypothetical protein BU14_0742s0001	F:GO:0003723	F:RNA binding
gene2681	1.8	class I SAM-dependent methyltransferase	F:GO:0003674	F:molecular_function
gene6488	1.9	ribosome biogenesis ATPase RIX7	P:GO:0008150 ; F:GO:0043167	P:biological_process; F:ion binding
gene671	1.9	predicted protein	C:GO:0005575	C:cellular_component
gene12686	1.9	gelsolin-like protein 2	C:GO:0005622 ; P:GO:0007010 ; F:GO:0008092 ; C:GO:0043226	C:intracellular; P:cytoskeleton organization; F:cytoskeletal protein binding; C:organelle
gene12683	1.9	hypothetical protein BU14_0104s0024	F:GO:0003700 ; P:GO:0006355	F:DNA binding transcription factor activity; P:regulation of transcription, DNA-templated
gene7628	1.9	hypothetical protein BU14_0326s0004	F:GO:0008168	F:methyltransferase activity
gene11952	1.9	hypothetical protein BU14_1542s0001, partial	F:GO:0003674 ; P:GO:0006629	F:molecular_function; P:lipid metabolic process
gene6565	1.9	predicted protein	F:GO:0003824 ; F:GO:0016740	F:catalytic activity; F:transferase activity
gene5303	2.0	hemolysin III-like protein isoform 2	C:GO:0016020 ; C:GO:0016021	C:membrane; C:integral component of membrane
gene9841	2.0	protein S-acyltransferase 8-like	C:GO:0016020 ; C:GO:0016021 ; F:GO:0016740 ; F:GO:0016746 ; F:GO:0019706	C:membrane; C:integral component of membrane; F:transferase activity; F:transferase activity, transferring acyl groups; F:protein-cysteine S-palmitoyltransferase activity

gene12000	2.0	zn-dependent exopeptidase	F:GO:0003674 ; P:GO:0008150	F:molecular_function; P:biological_process
gene9832	2.0	predicted protein	F:GO:0004672 ; F:GO:0005524 ; P:GO:0006468 ; F:GO:0016301 ; P:GO:0016310	F:protein kinase activity; F:ATP binding; P:protein phosphorylation; F:kinase activity; P:phosphorylation
gene4396	2.1	zinc finger protein 708-like	F:GO:0003674	F:molecular_function
gene9631	2.1	Endoribonuclease XendoU	F:GO:0004518	F:nuclease activity
gene8340	2.1	RNA polymerase sigma factor RpoD	F:GO:0003700 ; P:GO:0009058 ; P:GO:0034641	F:DNA binding transcription factor activity; P:biosynthetic process; P:cellular nitrogen compound metabolic process
gene12356	2.2	peroxisomal membrane protein PMP22	C:GO:0005575	C:cellular_component
gene11687	2.2	zinc finger protein OZF-like	F:GO:0003674	F:molecular_function
gene5929	2.3	carotenoid cleavage dioxygenase	F:GO:0016491 ; F:GO:0043167	F:oxidoreductase activity; F:ion binding
gene1152	2.3	nucleolar RNA helicase 2-like isoform X1	F:GO:0003723 ; F:GO:0004386 ; C:GO:0005634 ; F:GO:0043167	F:RNA binding; F:helicase activity; C:nucleus; F:ion binding
gene6141	2.3	predicted protein	C:GO:0005575	C:cellular_component
gene2054	2.3	Monogalactosyldiacylglycerol synthase fragment, family GT28	F:GO:0016757	F:transferase activity, transferring glycosyl groups

gene12890	2.4	Serine/threonine protein kinase	P:GO:0006464 ; F:GO:0016301 ; F:GO:0043167	P:cellular protein modification process; F:kinase activity; F:ion binding
gene1902	2.5	hypothetical protein BU14_0072s0103, partial	F:GO:0003674	F:molecular_function
gene1710	2.5	beta-carotene isomerase D27, chloroplastic	F:GO:0005506 ; C:GO:0009507 ; P:GO:0010223 ; F:GO:0016853 ; F:GO:0016859 ; P:GO:1901601	F:iron ion binding; C:chloroplast; P:secondary shoot formation; F:isomerase activity; F:cis-trans isomerase activity; P:strigolactone biosynthetic process
gene6142	2.5	hypothetical protein BU14_0240s0010	C:GO:0005575	C:cellular_component
gene5254	2.5	hypothetical protein BU14_0202s0013	C:GO:0005575	C:cellular_component
gene10145	2.5	hypothetical protein BU14_0572s0009	F:GO:0008233	F:peptidase activity
gene11919	2.6	iron ABC transporter substrate-binding protein	P:GO:0006091 ; F:GO:0016491	P:generation of precursor metabolites and energy; F:oxidoreductase activity
gene6133	2.7	Xylosyltransferase, family GT14	C:GO:0005575 ; F:GO:0016757	C:cellular_component; F:transferase activity, transferring glycosyl groups
gene13076	2.7	RNA polymerase sigma factor, RpoD/SigA family	F:GO:0003677 ; F:GO:0003700 ; P:GO:0006351 ; P:GO:0006352 ; P:GO:0006355 ; F:GO:0016987 ; P:GO:2000142	F:DNA binding; F:DNA binding transcription factor activity; P:transcription, DNA-templated; P:DNA-templated transcription, initiation; P:regulation of transcription, DNA-templated; F:sigma factor activity; P:regulation of DNA-templated transcription, initiation

gene4525	2.8	hypothetical protein BU14_0166s0032	F:GO:0008233	F:peptidase activity
gene12394	2.9	fructose-1,6-bisphosphate aldolase	C:GO:0005829; P:GO:0005975; P:GO:0006091; P:GO:0009058; F:GO:0016829; P:GO:0034655; P:GO:0044281; P:GO:0051186	C:cytosol; P:carbohydrate metabolic process; P:generation of precursor metabolites and energy; P:biosynthetic process; F:lyase activity; P:nucleobase-containing compound catabolic process; P:small molecule metabolic process; P:cofactor metabolic process
gene11909	2.9	alpha-1,2-fucosyltransferase	C:GO:0005575; P:GO:0005975; F:GO:0016757	C:cellular_component; P:carbohydrate metabolic process; F:transferase activity, transferring glycosyl groups
gene11644	3.2	hypothetical protein BU14_1241s0001	F:GO:0003700; P:GO:0009058; P:GO:0034641	F:DNA binding transcription factor activity; P:biosynthetic process; P:cellular nitrogen compound metabolic process
gene6143	3.2	hypothetical protein BU14_0109s0016	F:GO:0005509	F:calcium ion binding
gene3663	3.4	circadian clock associated 1	F:GO:0003677; C:GO:0005634; P:GO:0009058; P:GO:0034641	F:DNA binding; C:nucleus; P:biosynthetic process; P:cellular nitrogen compound metabolic process
gene6154	4.2	hypothetical protein BU14_0240s0022	C:GO:0005575	C:cellular_component

Supplement Table 7: DEGs between rehydration and dehydration with gene name, Log_2fold change (Log2FC), putative protein name and putative GO functions

Gene	Log2FC	Protein name	GO term	Go Function
gene6589	-7.2	hypothetical protein BU14_0263s0003	C:GO:0005575	C:cellular_component
gene2909	-3.0	hypothetical protein BU14_0238s0011	C:GO:0005575	C:cellular_component
gene2721	-1.6	hypothetical protein BU14_0098s0037	C:GO:0005575	F:oxidoreductase activity; F:ion binding
gene1859	-1.3	copper amine oxidase	F:GO:0016491; F:GO:0043167	F:oxidoreductase activity
gene4876	-1.3	polyamine oxidase-like	F:GO:0016491	C:cellular_component
gene9525	-1.3	hypothetical protein BU14_0492s0007	C:GO:0005575	C:nucleoplasm; P:cellular protein modification process; F:kinase activity; C:protein-containing complex; P:chromosome organization
gene5655	-1.2	transformation/transcription domain-associated protein-like	C:GO:0005654; P:GO:0006464; F:GO:0016301; C:GO:0032991; P:GO:0051276	F:molecular_function
gene8856	-1.2	hypothetical protein BU14_0424s0001	F:GO:0003674	C:membrane; C:integral component of membrane; F:kinase activity; P:phosphorylation
gene5068	-1.1	protein kinase	C:GO:0016020; C:GO:0016021; F:GO:0016301; P:GO:0016310	F:molecular_function; C:extracellular region; P:cellular protein modification process; P:growth
gene1263	-1.1	aspartyl/asparaginyl beta-hydroxylase domain-containing protein	F:GO:0003674; C:GO:0005576; P:GO:0006464	C:plastid; C:thylakoid; F:transmembrane transporter activity

			; P:GO:0040007	
gene1897	-1.1	HAD-superfamily hydrolase	C:GO:0009536 ; C:GO:0009579 ; F:GO:0022857	F:molecular_function
gene9160	-1.1	Transcription factor YY2	F:GO:0003674	C:cellular_component; F:oxidoreductase activity
gene7984	-1.0	respiratory burst oxidase-like protein	C:GO:0005575 ; F:GO:0016491	C:membrane; C:integral component of membrane
gene7986	-1.0	keratin-associated protein 16-1-like	C:GO:0016020 ; C:GO:0016021	C:cellular_component; P:lipid metabolic process; P:biosynthetic process; F:transferase activity, transferring alkyl or aryl (other than methyl) groups; P:small molecule metabolic process
gene8480	1.0	phytoene synthase	C:GO:0005575 ; P:GO:0006629 ; P:GO:0009058 ; F:GO:0016765 ; P:GO:0044281	C:membrane; C:integral component of membrane; F:transferase activity
gene12073	1.0	sulfotransferase family protein	C:GO:0016020 ; C:GO:0016021 ; F:GO:0016740	F:catalytic activity; P:nucleoside metabolic process; C:membrane; C:integral component of membrane
gene11258	1.0	ankryin	F:GO:0003824 ; P:GO:0009116 ; C:GO:0016020 ; C:GO:0016021	C:cellular_component
gene11631	1.0	hypothetical protein BU14_1226s0001	C:GO:0005575	C:cellular_component; P:lipid metabolic process; P:biosynthetic process; F:oxidoreductase activity; F:ion

				binding; P:small molecule metabolic process
gene4219	1.0	acyl-CoA desaturase	C:GO:0005575; P:GO:0006629; P:GO:0009058; F:GO:0016491; F:GO:0043167; P:GO:0044281	C:cell; P:homeostatic process
gene9657	1.0	hypothetical protein BU14_0504s0015	C:GO:0005623; P:GO:0042592	F:oxidoreductase activity; F:ion binding
gene7431	1.1	cytochrome P450	F:GO:0016491; F:GO:0043167	C:intracellular; C:plasma membrane; P:cellular protein modification process; P:signal transduction; P:cell-cell signaling; F:kinase activity; P:locomotion; F:ion binding; P:anatomical structure development
gene11136	1.1	serine/threonine protein kinase	C:GO:0005622; C:GO:0005886; P:GO:0006464; P:GO:0007165; P:GO:0007267; F:GO:0016301; P:GO:0040011; F:GO:0043167; P:GO:0048856	C:cell; P:transport; F:oxidoreductase activity; P:homeostatic process; F:ion binding
gene854	1.1	ferritin-3, chloroplastic	C:GO:0005623; P:GO:0006810; F:GO:0016491;	F:ligase activity; F:ion binding

			P:GO:0042592 ; F:GO:0043167	
gene11038	1.1	D-alanine--D-alanine ligase	F:GO:0016874 ; F:GO:0043167	C:membrane; C:integral component of membrane
gene11087	1.1	CRTAC1 family protein	C:GO:0016020 ; C:GO:0016021	P:cellular protein modification process; F:kinase activity; F:ion binding
gene3460	1.1	aarF domain-containing kinase	P:GO:0006464 ; F:GO:0016301 ; F:GO:0043167	F:molecular_function
gene8448	1.1	glutathione S-transferase Mu 1-like	F:GO:0003674	F:molecular_function
gene1101	1.1	hypothetical protein BU14_0052s0021	F:GO:0003674	P:biological_process; F:ion binding
gene6488	1.1	ribosome biogenesis ATPase RIX7	P:GO:0008150 ; F:GO:0043167	C:cellular_component
gene424	1.1	RND family transporter: Niemann-Pick type C1 disease protein-like protein	C:GO:0005575	C:cellular_component
gene6314	1.1	emopamil-binding protein	C:GO:0005575	F:calcium ion binding
gene3090	1.1	hypothetical protein BU14_0109s0016	F:GO:0005509	C:intracellular; P:cellular protein modification process; P:signal transduction; C:protein-containing complex
gene9667	1.1	carbohydrate-binding module family 48 protein	C:GO:0005622 ; P:GO:0006464 ; P:GO:0007165 ; C:GO:0032991	P:generation of precursor metabolites and energy; F:oxidoreductase activity
gene11919	1.2	iron ABC transporter substrate-binding protein	P:GO:0006091 ; F:GO:0016491	C:membrane; C:integral component of membrane

gene6620	1.2	hypothetical protein BU14_0264s0014	C:GO:0016020 ; C:GO:0016021	C:cellular_component; F:ATPase activity; F:transmembrane transporter activity; F:ion binding
gene6656	1.2	multidrug resistance protein homolog 49	C:GO:0005575 ; F:GO:0016887 ; F:GO:0022857 ; F:GO:0043167	F:ATPase activity; F:ion binding
gene11468	1.2	P-loop containing nucleoside triphosphate hydrolase protein	F:GO:0016887 ; F:GO:0043167	C:cytoplasm; P:transport; C:organelle
gene466	1.2	S-adenosylmethionine carrier 1, chloroplastic/mitochondrial isoform X1	C:GO:0005737 ; P:GO:0006810 ; C:GO:0043226	F:molecular_function
gene11687	1.2	zinc finger protein OZF-like	F:GO:0003674	C:cellular_component
gene12360	1.2	probable S-adenosylmethionine carrier 2, chloroplastic isoform X2	C:GO:0005575	P:cellular protein modification process; F:kinase activity; F:ion binding
gene12890	1.2	Serine/threonine protein kinase	P:GO:0006464 ; F:GO:0016301 ; F:GO:0043167	C:cellular_component
gene10717	1.2	hypothetical protein BU14_0692s0013	C:GO:0005575	F:ion binding
gene3653	1.2	hypothetical protein BU14_0130s0016	F:GO:0043167	C:cellular_component
gene9630	1.2	hypothetical protein BU14_0502s0002	C:GO:0005575	F:molecular_function; P:biological_process
gene12000	1.3	zn-dependent exopeptidase	F:GO:0003674 ; P:GO:0008150	P:lipid metabolic process; P:biosynthetic process; C:plastid; F:oxidoreductase activity; P:small molecule metabolic process

gene5345	1.3	fatty acid desaturase 4, chloroplastic-like	P:GO:0006629 ; P:GO:0009058 ; C:GO:0009536 ; F:GO:0016491 ; P:GO:0044281	P:response to stress; F:oxidoreductase activity; F:ion binding
gene5076	1.3	catalase	P:GO:0006950 ; F:GO:0016491 ; F:GO:0043167	F:lyase activity; F:ion binding
gene6749	1.3	carbonic anhydrase	F:GO:0016829 ; F:GO:0043167	C:cellular_component
gene11207	1.3	predicted protein	C:GO:0005575	C:membrane; C:integral component of membrane; F:transferase activity
gene10290	1.4	Glycosyltransferase, family GT90	C:GO:0016020 ; C:GO:0016021 ; F:GO:0016740	C:cellular_component; F:ATPase activity; F:transmembrane transporter activity; F:ion binding
gene12825	1.4	multidrug resistance protein 3-like	C:GO:0005575 ; F:GO:0016887 ; F:GO:0022857 ; F:GO:0043167	C:cellular_component
gene671	1.4	predicted protein	C:GO:0005575	F:glutathione transferase activity; F:transferase activity
gene2098	1.4	glutathione S-transferase	F:GO:0004364 ; F:GO:0016740	F:RNA binding; F:helicase activity; C:nucleus; F:ion binding
gene1152	1.4	nucleolar RNA helicase 2-like isoform X1	F:GO:0003723 ; F:GO:0004386 ; C:GO:0005634	P:cellular protein modification process; F:kinase activity; F:ion binding

			; F:GO:0043167	
gene12018	1.5	serine/threonine-protein kinase HT1-like	P:GO:0006464 ; F:GO:0016301 ; F:GO:0043167	F:peptidase activity; F:oxidoreductase activity; F:ion binding
gene12687	1.5	hypothetical protein BU14_2421s0001, partial	F:GO:0008233 ; F:GO:0016491 ; F:GO:0043167	C:cellular_component
gene6142	1.5	hypothetical protein BU14_0240s0010	C:GO:0005575	P:response to stress; F:oxidoreductase activity; F:ion binding
gene4199	1.5	catalase family protein	P:GO:0006950 ; F:GO:0016491 ; F:GO:0043167	C:membrane; C:integral component of membrane; F:transferase activity; F:transferase activity, transferring acyl groups; F:protein-cysteine S-palmitoyltransferase activity
gene9841	1.5	protein S-acyltransferase 8-like	C:GO:0016020 ; C:GO:0016021 ; F:GO:0016740 ; F:GO:0016746 ; F:GO:0019706	F:protein kinase activity; F:ATP binding; P:protein phosphorylation; F:kinase activity; P:phosphorylation
gene9832	1.6	predicted protein	F:GO:0004672 ; F:GO:0005524 ; P:GO:0006468 ; F:GO:0016301 ; P:GO:0016310	F:oxidoreductase activity; F:ion binding
gene5929	1.7	carotenoid cleavage dioxygenase	F:GO:0016491 ; F:GO:0043167	C:cellular_component

gene1227	1.7	hypothetical protein BU14_0055s0007	C:GO:0005575	C:cellular_component
gene5254	1.7	hypothetical protein BU14_0202s0013	C:GO:0005575	F:molecular_function; C:intracellular; C:organelle
gene654	1.7	hypothetical protein BU14_0031s0064	F:GO:0003674 ; C:GO:0005622 ; C:GO:0043226	C:cellular_component
gene12356	1.7	peroxisomal membrane protein PMP22	C:GO:0005575	F:iron ion binding; C:chloroplast; P:secondary shoot formation; F:isomerase activity; F:cis-trans isomerase activity; P:strigolactone biosynthetic process
gene1710	1.7	beta-carotene isomerase D27, chloroplastic	F:GO:0005506 ; C:GO:0009507 ; P:GO:0010223 ; F:GO:0016853 ; F:GO:0016859 ; P:GO:1901601	C:cellular_component; F:kinase activity
gene12792	1.7	predicted protein	C:GO:0005575 ; F:GO:0016301	C:cellular_component
gene6141	1.7	predicted protein	C:GO:0005575	F:molecular_function
gene6412	1.8	hypothetical protein BU14_0256s0035	F:GO:0003674	F:peptidase activity
gene10145	1.8	hypothetical protein BU14_0572s0009	F:GO:0008233	C:cellular_component; P:carbohydrate metabolic process; F:transferase activity, transferring glycosyl groups
gene11909	1.9	alpha-1,2-fucosyltransferase	C:GO:0005575 ; P:GO:0005975 ; F:GO:0016757	F:DNA binding transcription factor activity; P:biosynthetic process; P:cellular nitrogen compound metabolic process

gene8340	1.9	RNA polymerase sigma factor RpoD	F:GO:0003700 ; P:GO:0009058 ; P:GO:0034641	C:membrane; C:integral component of membrane
gene5303	2.1	hemolysin III-like protein isoform 2	C:GO:0016020 ; C:GO:0016021	F:DNA binding transcription factor activity; P:regulation of transcription, DNA-templated
gene12683	2.1	hypothetical protein BU14_0104s0024	F:GO:0003700 ; P:GO:0006355	C:cellular_component; F:transferase activity, transferring glycosyl groups
gene6133	2.2	Xylosyltransferase, family GT14	C:GO:0005575 ; F:GO:0016757	F:DNA binding transcription factor activity; P:biosynthetic process; P:cellular nitrogen compound metabolic process
gene11644	2.3	hypothetical protein BU14_1241s0001	F:GO:0003700 ; P:GO:0009058 ; P:GO:0034641	C:cytosol; P:carbohydrate metabolic process; P:generation of precursor metabolites and energy; P:biosynthetic process; F:lyase activity; P:nucleobase-containing compound catabolic process; P:small molecule metabolic process; P:cofactor metabolic process
gene12394	2.3	fructose-1,6-bisphosphate aldolase	C:GO:0005829 ; P:GO:0005975 ; P:GO:0006091 ; P:GO:0009058 ; F:GO:0016829 ; P:GO:0034655 ; P:GO:0044281 ; P:GO:0051186	F:molecular_function
gene12060	2.3	hypothetical protein BU14_1639s0001	F:GO:0003674	F:oxidoreductase activity; F:ion binding
gene11580	2.5	hypothetical protein BU14_1142s0001, partial	F:GO:0016491 ; F:GO:0043167	F:oxidoreductase activity; F:ion binding

gene11547	2.7	hypothetical protein BU14_1090s0001, partial	F:GO:0016491 ; F:GO:0043167	F:oxidoreductase activity
gene11995	2.8	oxidase	F:GO:0016491	F:DNA binding; C:nucleus; P:biosynthetic process; P:cellular nitrogen compound metabolic process
gene3663	3.4	circadian clock associated 1	F:GO:0003677 ; C:GO:0005634 ; P:GO:0009058 ; P:GO:0034641	F:hydrolase activity

Supplement Table 8: DEGs between rehydration and desiccation with gene name, Log_2fold change (Log2FC), putative protein name and putative GO functions

Gene	Log2FC	Protein name	GO term	Go Function
gene4487	-7.4	hypothetical protein BU14_0283s0005	F:GO:0008270	F:nucleic acid binding
gene6631	-7.2	hypothetical protein BU14_0265s0010	F:GO:0003676	F:helicase activity; P:DNA metabolic process; P:response to stress; P:homeostatic process; F:ion binding; P:chromosome organization
gene6615	-4.8	hypothetical protein BU14_0033s0078	F:GO:0004386; P:GO:0006259; P:GO:0006950; P:GO:0042592; F:GO:0043167; P:GO:0051276	C:cellular_component
gene6776	-4.1	hypothetical protein BU14_0272s0017	C:GO:0005575	C:cytoplasm; P:tRNA metabolic process; P:translation; P:cellular amino acid metabolic process; F:ligase activity; F:ion binding
gene13323	-3.7	ycf68 protein	C:GO:0005737 ; P:GO:0006399; P:GO:0006412; P:GO:0006520;	C:cytoplasm; P:tRNA metabolic process; P:translation; P:cellular amino

			F:GO:0016874; F:GO:0043167	acid metabolic process; F:ligase activity; F:ion binding
gene13370	-3.6	ycf68 protein	C:GO:0005737 ; P:GO:0006399; P:GO:0006412; P:GO:0006520; F:GO:0016874; F:GO:0043167	C:plastid
gene13326	-3.2	Uncharacterised protein	C:GO:0009536	C:plastid
gene13373	-3.2	Uncharacterised protein	C:GO:0009536	P:reproduction; F:DNA binding; C:nucleus; P:biosynthetic process; P:cellular nitrogen compound metabolic process; P:anatomical structure development
gene2799	-2.6	MEDEA	P:GO:0000003; F:GO:0003677; C:GO:0005634 ; P:GO:0009058; P:GO:0034641; P:GO:0048856	F:ATP binding; C:membrane; F:phosphotransferase activity, alcohol group as acceptor
gene12428	-2.4	hypothetical protein BU14_2086s0001	F:GO:0005524; C:GO:0016020 ; F:GO:0016773	C:membrane; C:integral component of membrane
gene7738	-2.3	hypothetical protein BU14_0331s0006	C:GO:0016020 ; C:GO:0016021	C:cellular_component
gene2997	-2.2	hypothetical protein BU14_0107s0003	C:GO:0005575	F:peptidase activity
gene2168	-2.2	hypothetical protein BU14_0079s0007	F:GO:0008233	C:cellular_component
gene4490	-2.0	hypothetical protein BU14_0165s0028	C:GO:0005575	F:nucleic acid binding; F:RNA-DNA hybrid ribonuclease activity; F:zinc ion binding; P:DNA integration; P:RNA phosphodiester bond hydrolysis, endonucleolytic

gene292	-1.9	hypothetical protein BU14_0022s0113	F:GO:0003676; F:GO:0004523; F:GO:0008270; P:GO:0015074; P:GO:0090502	F:nucleic acid binding; C:membrane; C:integral component of membrane
gene6526	-1.9	hypothetical protein BU14_0260s0022	F:GO:0003676; C:GO:0016020; C:GO:0016021	F:DNA binding; P:DNA recombination; P:DNA integration; C:membrane; C:integral component of membrane
gene11045	-1.8	hypothetical protein BU14_0784s0002	F:GO:0003677; P:GO:0006310; P:GO:0015074; C:GO:0016020; C:GO:0016021	C:membrane; C:integral component of membrane
gene5823	-1.8	hypothetical protein BU14_0228s0005	C:GO:0016020; C:GO:0016021	C:cellular_component
gene9357	-1.8	hypothetical protein BU14_0474s0007	C:GO:0005575	F:peptidase activity
gene99	-1.4	hypothetical protein BU14_0014s0100	F:GO:0008233	F:zinc ion binding
gene11921	-1.4	hypothetical protein BU14_1511s0001	F:GO:0008270	F:ion binding
gene11490	-1.4	hypothetical protein BU14_1041s0003	F:GO:0043167	F:peptidase activity
gene5824	-1.3	hypothetical protein BU14_0228s0006	F:GO:0008233	C:nuclear chromosome; P:response to stress; P:cell cycle; C:protein-containing complex
gene10330	-1.3	hypothetical protein BU14_0608s0014	C:GO:0000228; P:GO:0006950; P:GO:0007049; C:GO:0032991	F:GTPase activity; P:translation; F:ion binding
gene4667	-1.2	elongation factor G	F:GO:0003924; P:GO:0006412; F:GO:0043167	C:cellular_component
gene9525	-1.2	hypothetical protein BU14_0492s0007	C:GO:0005575	F:molecular_function; C:extracellular region;

				P:cellular protein modification process; P:growth
gene1263	-1.1	aspartyl/asparaginyl beta-hydroxylase domain-containing protein	F:GO:0003674; C:GO:0005576; P:GO:0006464; P:GO:0040007	C:cellular_component
gene4718	-1.1	hypothetical protein BU14_0176s0030	C:GO:0005575	C:cellular_component; F:transmembrane transporter activity
gene11273	1.0	Glycerol uptake facilitator protein	C:GO:0005575; F:GO:0022857	C:cytoplasm; P:transport; C:organelle
gene466	1.0	S-adenosylmethionine carrier 1, chloroplastic/mitochondrial isoform X1	C:GO:0005737; P:GO:0006810; C:GO:0043226	C:cellular_component
gene12257	1.0	prohibitin-3, mitochondrial-like	C:GO:0005575	C:cellular_component; F:peptidase activity
gene645	1.0	serine protease 27-like	C:GO:0005575; F:GO:0008233	C:membrane; C:integral component of membrane
gene6620	1.0	hypothetical protein BU14_0264s0014	C:GO:0016020; C:GO:0016021	P:carbohydrate transport; F:pyrimidine nucleotide-sugar transmembrane transporter activity; C:membrane; C:integral component of membrane; P:pyrimidine nucleotide-sugar transmembrane transport
gene6107	1.1	UDP-N-acetylglucosamine/UDP-glucose/GDP-mannose transporter	P:GO:0008643; F:GO:0015165; C:GO:0016020; C:GO:0016021; P:GO:0090481	P:cellular protein modification process; F:kinase activity; F:ion binding
gene12018	1.1	serine/threonine-protein kinase HT1-like	P:GO:0006464; F:GO:0016301; F:GO:0043167	F:phosphatase activity
gene9566	1.1	predicted protein	F:GO:0016791	F:structural molecule activity; C:cellular_component

gene1611	1.1	hypothetical protein BU14_0068s0010	F:GO:0005198; C:GO:0005575	F:ATPase activity; F:ion binding
gene11468	1.1	P-loop containing nucleoside triphosphate hydrolase protein	F:GO:0016887; F:GO:0043167	F:molecular_function; P:biological_process
gene10556	1.1	hypothetical protein BU14_0654s0003	F:GO:0003674; P:GO:0008150	F:molecular_function
gene8448	1.1	glutathione S-transferase Mu 1-like	F:GO:0003674	F:molecular_function; C:intracellular; P:cellular protein modification process; P:response to stress
gene2932	1.1	ubiquitin-conjugating enzyme E2 T	F:GO:0003674; C:GO:0005622 ; P:GO:0006464; P:GO:0006950	C:cellular_component
gene11631	1.1	hypothetical protein BU14_1226s0001	C:GO:0005575	C:cellular_component
gene12360	1.1	probable S-adenosylmethionine carrier 2, chloroplastic isoform X2	C:GO:0005575	F:molecular_function; P:biological_process
gene12000	1.1	zn-dependent exopeptidase	F:GO:0003674; P:GO:0008150	C:cellular_component
gene8032	1.1	hypothetical protein BU14_0352s0006	C:GO:0005575	C:cellular_component; F:ATPase activity; F:transmembrane transporter activity; F:ion binding
gene12825	1.1	multidrug resistance protein 3-like	C:GO:0005575 ; F:GO:0016887; F:GO:0022857; F:GO:0043167	F:molecular_function
gene1101	1.1	hypothetical protein BU14_0052s0021	F:GO:0003674	C:cellular_component
gene12356	1.2	peroxisomal membrane protein PMP22	C:GO:0005575	C:intracellular; P:cellular protein modification process; P:signal transduction; C:protein-containing complex

gene9667	1.2	carbohydrate-binding module family 48 protein	C:GO:0005622; P:GO:0006464; P:GO:0007165; C:GO:0032991	C:cell; P:homeostatic process
gene9657	1.2	hypothetical protein BU14_0504s0015	C:GO:0005623; P:GO:0042592	C:cellular_component
gene5254	1.2	hypothetical protein BU14_0202s0013	C:GO:0005575	C:cellular_component
gene7007	1.2	hypothetical protein BU14_0288s0028	C:GO:0005575	C:cellular_component
gene10717	1.2	hypothetical protein BU14_0692s0013	C:GO:0005575	F:RNA binding; F:helicase activity; C:nucleus; F:ion binding
gene1152	1.2	nucleolar RNA helicase 2-like isoform X1	F:GO:0003723; F:GO:0004386; C:GO:0005634; F:GO:0043167	F:ATPase activity; F:ion binding
gene8926	1.2	ABC transporter I family member 10, chloroplastic	F:GO:0016887; F:GO:0043167	F:DNA binding; P:biosynthetic process; F:nucleotidyltransferase activity; P:cellular nitrogen compound metabolic process
gene6795	1.2	hypothetical protein BU14_0166s0020	F:GO:0003677; P:GO:0009058; F:GO:0016779; P:GO:0034641	C:cellular_component
gene6141	1.2	predicted protein	C:GO:0005575	C:cellular_component
gene8149	1.2	hypothetical protein BU14_0361s0009	C:GO:0005575	C:cellular_component; F:oxidoreductase activity
gene5418	1.2	FAD-binding domain-containing protein	C:GO:0005575; F:GO:0016491	C:intracellular; P:lipid metabolic process; P:signal transduction; F:kinase activity
gene4936	1.2	hypothetical protein BU14_0185s0021	C:GO:0005622; P:GO:0006629; P:GO:0007165; F:GO:0016301	C:cellular_component

gene4069	1.3	hypothetical protein BU14_0148s0028	C:GO:0005575	C:cellular_component
gene3169	1.3	hypothetical protein BU14_0111s0045	C:GO:0005575	P:generation of precursor metabolites and energy; F:oxidoreductase activity
gene11919	1.3	iron ABC transporter substrate-binding protein	P:GO:0006091; F:GO:0016491	C:cellular_component; F:ATPase activity; F:transmembrane transporter activity; F:ion binding
gene7743	1.3	ABC transporter A family member 2	C:GO:0005575 ; F:GO:0016887; F:GO:0022857; F:GO:0043167	F:molecular_function; C:intracellular; P:cellular protein modification process; C:protein-containing complex; C:organelle
gene7508	1.3	kelch-like ECH-associated protein 1	F:GO:0003674; C:GO:0005622 ; P:GO:0006464; C:GO:0032991 ; C:GO:0043226	C:cellular_component
gene7415	1.4	hypothetical protein BU14_0313s0013	C:GO:0005575	C:cellular_component; F:kinase activity
gene12792	1.4	predicted protein	C:GO:0005575 ; F:GO:0016301	F:iron ion binding; C:chloroplast; P:secondary shoot formation; F:isomerase activity; F:cis-trans isomerase activity; P:strigolactone biosynthetic process
gene1710	1.4	beta-carotene isomerase D27, chloroplastic	F:GO:0005506; C:GO:0009507 ; P:GO:0010223; F:GO:0016853; F:GO:0016859; P:GO:1901601	F:catalytic activity; F:transferase activity
gene6565	1.4	predicted protein	F:GO:0003824; F:GO:0016740	C:cytoplasm; F:methyltransferase activity; C:organelle
gene12807	1.4	predicted protein	C:GO:0005737 ;	F:oxidoreductase activity

			F:GO:0008168; C:GO:0043226	
gene1516	1.5	alcohol dehydrogenase	F:GO:0016491	C:cellular_component
gene1227	1.5	hypothetical protein BU14_0055s0007	C:GO:0005575	F:protein kinase activity; F:ATP binding; P:protein phosphorylation; F:kinase activity; P:phosphorylation
gene9832	1.5	predicted protein	F:GO:0004672; F:GO:0005524; P:GO:0006468; F:GO:0016301; P:GO:0016310	C:cellular_component
gene6142	1.5	hypothetical protein BU14_0240s0010	C:GO:0005575	C:membrane; C:integral component of membrane
gene5303	1.5	hemolysin III-like protein isoform 2	C:GO:0016020 ; C:GO:0016021	F:glutathione transferase activity; F:transferase activity
gene2098	1.5	glutathione S-transferase	F:GO:0004364; F:GO:0016740	F:DNA binding transcription factor activity; P:regulation of transcription, DNA-templated
gene12683	1.5	hypothetical protein BU14_0104s0024	F:GO:0003700; P:GO:0006355	P:cellular protein modification process; F:kinase activity; F:ion binding
gene12890	1.6	Serine/threonine protein kinase	P:GO:0006464; F:GO:0016301; F:GO:0043167	F:DNA binding transcription factor activity; P:biosynthetic process; P:cellular nitrogen compound metabolic process
gene8340	1.7	RNA polymerase sigma factor RpoD	F:GO:0003700; P:GO:0009058; P:GO:0034641	C:membrane; C:integral component of membrane; F:transferase activity
gene10290	1.7	Glycosyltransferase, family GT90	C:GO:0016020 ; C:GO:0016021 ; F:GO:0016740	F:oxidoreductase activity; F:ion binding
gene5929	1.7	carotenoid cleavage dioxygenase	F:GO:0016491; F:GO:0043167	C:cellular_component
gene671	1.7	predicted protein	C:GO:0005575	F:peptidase activity; F:oxidoreductase activity; F:ion binding

gene12687	1.8	hypothetical protein BU14_2421s0001, partial	F:GO:0008233; F:GO:0016491; F:GO:0043167	F:molecular_function
gene11687	1.8	zinc finger protein OZF-like	F:GO:0003674	F:DNA binding; F:DNA binding transcription factor activity; P:transcription, DNA-templated; P:DNA-templated transcription, initiation; P:regulation of transcription, DNA-templated; F:sigma factor activity; P:regulation of DNA-templated transcription, initiation
gene13076	1.8	RNA polymerase sigma factor, RpoD/SigA family	F:GO:0003677; F:GO:0003700; P:GO:0006351; P:GO:0006352; P:GO:0006355; F:GO:0016987; P:GO:2000142	F:molecular_function; P:DNA metabolic process
gene7632	1.8	hypothetical protein BU14_0131s0014	F:GO:0003674; P:GO:0006259	F:molecular_function
gene6412	1.9	hypothetical protein BU14_0256s0035	F:GO:0003674	C:cellular_component
gene4694	1.9	hypothetical protein BU14_0176s0006	C:GO:0005575	F:DNA binding transcription factor activity; P:biosynthetic process; P:cellular nitrogen compound metabolic process
gene11644	2.0	hypothetical protein BU14_1241s0001	F:GO:0003700; P:GO:0009058; P:GO:0034641	C:cellular_component; F:transferase activity, transferring glycosyl groups
gene6133	2.1	Xylosyltransferase, family GT14	C:GO:0005575; F:GO:0016757	C:cytosol; P:carbohydrate metabolic process; P:generation of precursor metabolites and energy; P:biosynthetic process; F:lyase activity; P:nucleobase-containing compound catabolic process; P:small molecule metabolic process; P:cofactor metabolic process
gene12394	2.1	fructose-1,6-bisphosphate aldolase	C:GO:0005829; P:GO:0005975; P:GO:0006091;	C:cellular_component

			P:GO:0009058; F:GO:0016829; P:GO:0034655; P:GO:0044281; P:GO:0051186	
gene2463	2.2	hypothetical protein BU14_0713s0004	C:GO:0005575	C:cellular_component; P:carbohydrate metabolic process; F:transferase activity, transferring glycosyl groups
gene11909	2.2	alpha-1,2-fucosyltransferase	C:GO:0005575; P:GO:0005975; F:GO:0016757	F:oxidoreductase activity; F:ion binding
gene11580	2.7	hypothetical protein BU14_1142s0001, partial	F:GO:0016491; F:GO:0043167	F:hydrolase activity
gene1533	2.9	hypothetical protein BU14_0065s0018	F:GO:0016787	