School of Sciences and Engineering

TRANSCRIPTOMICS OF IRON LIMITATION IN *PHAEOCYSTIS ANTARCTICA*

A Thesis Submitted to

the Biotechnology Master's Program

in partial fulfillment of the requirements for

the degree in Master of Science

By: Mariam Reyad Rizkallah Issak

Under the co-supervision of:

Assist. Prof. Ahmed Moustafa

Biology Department

The American University in Cairo

Dr. Sára Beszteri

Polar Biological Oceanography

Alfred Wegener Institute for Polar and Marine Research

May / 2014

The American University in Cairo

TRANSCRIPTOMICS OF IRON LIMITATION IN *PHAEOCYSTIS ANTARCTICA*

A Thesis Submitted by
Mariam Reyad Rizkallah Issak

to the Biotechnology Master's Program
May / 2014

in partial fulfillment of the requirements for
the degree in Master of Science

Has been approved by

Thesis Committee Supervisor/Chair

_____

Affiliation _____

Thesis Committee Supervisor

_____

Affiliation _____

Thesis Committee Reader/Examiner

_____

Affiliation _____

Thesis Committee Reader/Examiner

_____

Affiliation _____

_____ _____ _____ _____

Dept. Chair/Director Date Dean Date

To the only constant in my life, my mother, my father and my brother.

ACKNOWLEDGEMENTS

I thank my supervisor, Dr. Ahmed Moustafa, Graduate Program Director, the American University in Cairo (AUC) for teaching me how a true scientist should be like, for saving the study and for making all what I achieved throughout my studentship possible. I thank my co-supervisor, Dr. Sára Beszteri, Alfred Wegener Institute (AWI) for training me on lab techniques, for the study design and moreover for her friendship inside and outside the lab. SB has conducted trace nutrient supplementation, RNA extraction and quantitative and qualitative assessments. I thank Dr. Steffi Gäbler-Schwarz for supplying the project with *P. antarctica* clones and for her valuable comments on *Phaeocystis* biology.

From AUC, I thank Dean Amr Sharaawi for AUC's grant office support for both my participation in the expedition ARK-XXVII/1 and my semester abroad. I thank Dr. Rania Siam, Biology Department Chair and former Graduate Program Director for her everlasting support, Dr. Arthur Bos and Dr. Walid Fouad for their support to my fellowship and grants applications. I thank Dr. Ari Ferreira, Hazem, Mustafa, Amged and Mr. Osama for training me on the university's computational facilities, their support and friendship. I thank my friends and colleagues Nahla, Yasmeen, Hadeel, Mohamed Lotaief, and Ali for being here for me even though I was not really here. I thank Laila, Sarah, Mona, Mohamed Abou El-Soud, Mahera, Noha, Rehab and Ghada for their valuable input to my research. I thank my Biotechnology program instructors, Dr. Hamza El-Dorry, Dr. Asma Amleh, Dr. Walid Fouad, Dr. Ahmed Said, and Dr. Edwin Rivera. I thank my colleagues Sarah Sonbol, Ahmed Samir, Momen, Aya, Mai, Mohamed Abdel-Wahed, Walid, and Mohamed Nagy. I thank Dr. Ramy Aziz, Faculty of Pharmacy, Cairo University, for his mentorship and support in college and up till this very day.

From AWI, I thank Dr. Klaus Valentin, Sea Ice Biology and Biochemistry, for his support throughout my participation in the expedition ARK-XXVII/1, my enrolment at Universität Bremen, my training at AWI and stay in Germany. I thank my expedition colleagues especially my science-mate, Katrin for the valuable discussions. I thank Dr. Bánk Beszteri and Dr. Barbara Niehoff for his support and guidance at AWI.

I am grateful to AWI Biosciences Division Director, Prof. Dr. Dieter Wolf-Gladrow for his support, and section leaders Prof. Dr. Bettina Meyer and Dr. Gerhard Dieckmann for allowing me to work in the sections labs and for their support with all the paperwork. I thank Dr. Bernd Krock from Ecological Chemistry section for conducting *Fragilariopsis kerguelensis* metabolic analysis.

From Sea Ice Biology and Biochemistry, I thank Mrs. Erika Allhusen for training me on the labs facilities, photosynthesis measurements, and culturing techniques and for her kindness and support. I thank Dr. Maddalena Bayer for her valuable comments and support. From Polar Biological Oceanography section, I thank Dr. Christine Klaas for her help in Dark Microscopy labs and her rich discussions on phytoplankton ecology and iron enrichment. I thank Nike and Isabelle for training me on cell enumeration and Christian for supplying me with his latest published work. I thank Ms. Christiane Lorenzen and Ms. Sandra Murawski for their help with pigment and elemental analysis and Dr. Eva-Maria Nöthig for her valuable discussions. I thank Mrs. Tanja Glawatty for her friendship and help with the paperwork.

From Universität Bremen, I thank Dr. Holger Auel for accepting me as a guest student at his program and for his help with biological oceanography education principles, and I thank my Marine Biology and Erasmus Mundus colleagues. I thank marine microbiology instructors at MPI-MM and plankton ecology instructors from Universität Bremen and AWI. From POLMAR, I thank Dr. Claudia Hanfland for giving me the opportunity to attend POLMAR courses and seminars, and I thank Ella, Mirja and Clara, for their friendship and support.

I acknowledge the Data Intensive Academic Grid (DIAG) (http://diagcomputing.org/) for allowing me to conduct the assembly and partially the downstream analysis on it, and Google Drive for saving my thesis.

Finally, I express my gratitude to the American University in Cairo for awarding me Laboratory Instructions Fellowship for one year and Al-Alfi Foundation for awarding me its Biotechnology Fellowship for two years to purse my master's degree.

ABSTRACT

Phytoplankton, the hidden trees of the ocean, are responsible for nearly half of global oxygen production. Phytoplankton growth and productivity in high-nitrate low-chlorophyll (HNLC) regions have been demonstrated to be limited by iron availability. One of the most important HNLC regions is the Southern Ocean, as it regulates Earth system climate due to its unique hydrography and consequently its phytoplankton assemblage drives the global carbon cycle. Iron fertilization experiments were conducted to understand the response of its photosynthetic assemblage to iron addition, and the nanoflagellate *Phaeocystis antarctica* has been reported the first to bloom. *P. antarctica* is ecologically important due to its contribution in global sulfur gases emissions. Thus, *P. antarctica* fits as an ideal model organism in understanding phytoplankton adaptation to iron limitation and functional changes following iron addition.

Here the results of a transcriptomic study assessing the effect of iron repletion on *P. antarctica* in a time-series manner are firstly reported. A Ross Sea-endemic *P. antarctica* clone was acclimated under iron limitation and iron was supplemented to the cultures. RNA was extracted at 5 time-points before and after iron addition. The generated 389,846,414 paired-end Illumina reads were *de novo* assembled into ~88,000 putative genes, providing the first reported transcriptome of *P. antarctica*. Differential expression at each time-point was inferred at an adjusted p-value ≤ 0.001 and a log fold-change ≥ 5, revealing a total of 2,367 differentially expressed genes.

Iron-limited *P. antarctica* recovered its photosynthetic fitness, colony-forming ability, and chlorophyll *a*, particulate organic carbon and nitrogen contents shortly after iron addition comparable to the replete control and the reported values of healthy cells. A shift in expression from iron-economic reactive oxygen species defense and photosystem II to iron-dependent alternatives has been observed. In addition, a metabolic shift from structural carbon and nitrogen reallocation to anabolism has been observed. Transcriptomic data supports the previous studies that *P. antarctica* is successful in utilizing bound iron in a reductive non-ligand-dependent mechanism. Iron levels limits *P. antarctica* growth yet its iron requirements are the lowest among all phytoplankton species and its adaptation to iron limitation is well-established.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SUPPLEMENTARY MATERIALS[1]

Files names/types	**Files description**
.fastq.bz2	Thirty two files containing Illumina RNA-Seq paired-end raw reads of the of the 16 sequenced *Phaeocystis antarctica* samples
Trinity.fasta	Assembled transcripts by Trinity pipeline
.pep	Extracted and translated open reading frames by TransDecoder used in functional annotation
.xls	Trinotate annotation report
rRNA.fasta	Predicted rRNA genes sequences
.genes.results	Sixteen files containing gene abundance estimates for each sequenced sample
.isoforms.results	Sixteen files containing isoform abundance estimates for each sequenced sample
.filtered.txt	Ten files containing significantly differentially expressed components binomial test statistics sorted by decreasing *p*-value

[1] Available at https://copy.com/SFSCnYX0Aoon

LIST OF ABBREVIATIONS

ACC	Antarctic Circumpolar Current
BLAST	Basic Local Alignment Search Tool
cDNA	Complementary DNA
DFB	Desferrioxamine B
DMS	Dimethyl Sulfide
DMSP	Dimethylsulfoniopropionate
EDTA	Ethylenediaminetetraacetic acid
eggNOG	Evolutionary Genealogy of Genes: Non-Supervised Orthologous Groups
GO	Gene Ontology
HNLC	High-Nutrient/Nitrate Low-Chlorophyll
mV	Millivolt
ORF	Open Reading Frame
PAM	Pulse Amplitude Modulation
PCA	Principal Component Analysis
PFD	Proton Flux Density
PSI	Photosystem I
PSII	Photosystem II
rDNA	Ribosomal DNA
RNA-Seq	Ribonucleic acid (RNA) sequencing
SO	Southern Ocean
TCA	Tricarboxylic Acid
tRNA	Transfer RNA

CHAPTER 1: INTRODUCTION AND STUDY OBJECTIVES

1.1 Trees of the Ocean: Marine Phytoplankton Growth and Productivity

Phytoplankton, or the prokaryotic and unicellular eukaryotic photosynthetic drifting community, is long known as the trees of the ocean. It is responsible for half of the global net primary production [(Field et al. 1998) cited in the authors summary of (Marchetti et al. 2012)] (Figure 1). Marine diatoms solely contribute 40% of oceanic carbon production annually (Allen et al. 2008).

The assemblage of phytoplankton can be classified according to: (1) size [e.g., pico-(0.2-2 μm), nano-(2-20 μm) and microphytoplankton (20-200 μm based on Sieburth-scale (Lalli & Parsons 2004; Sieburth et al. 1978)], (2) structure [having a shell or scales mineralized by silica (e.g., diatoms and golden-brown algae) or calcium carbonate (e.g., coccolithophores) or organic cell wall], (3) taxonomy [e.g., Cyanobateria and unicellular eukaryotes including: green and red algae, stramenopiles such as diatoms and haptophytes (e.g., coccolithophores and *Phaeocystis*), alveolates (e.g., dinoflagellates and ciliates), and Rhizaria (e.g., Foraminifera and Acantharia)], (4) evolutionary origin and fate of their plastids [e.g., through primary (e.g., red and green algae), secondary (e.g., chromalveolates supergroup and Rhizaria) or tertiary endosymbiosis (e.g., dinoflagellates) (Delwiche & Palmer 1997; Reyes-Prieto et al. 2008; Moustafa et al. 2009)], and (5) ecological impact [e.g., toxin-producing dinoflagellate species 'red tides' and dimethylsulfoniopropionate (DMSP)-producing Haptophyta]. Given their diversity, different phytoplankton classes neither contribute equally in global oxygen production nor are distributed equally in the world ocean.

In the early 20th Century, Alfred Redfield has studied the elemental composition across the oceans correlating water carbon (C), nitrate (N) and phosphate (P) contents to that of the marine communities and elucidating the optimal nutrients ratio for plankton growth –in general— and photosynthetic plankton growth –in particular— to be: C:N:P = 106:16:1 (Redfield 1934). Sodium, calcium, potassium and other abundant minerals are required for phytoplankton growth, however, growth also depends on group-specific macronutrients (e.g., silica for diatoms). In addition to water macronutrients contents,

phytoplankton growth rate, and primary productivity and photosynthesis rate, are governed by various factors. Abiotic factors such as: water column depth and consequently light intensity variability, micronutrients concentrations [e.g., iron (de Baar et al. 2005) and vitamin B12 (Bertrand et al. 2011)], CO_2 levels (Hoppe et al. 2013), water temperature [according to Monod equation (de Baar et al. 2005)], and water layer mixing and subsequently nutrients availability, as well as biotic factors such as: pathogen and viral attacks (Suttle 2007), species competition [addressed through Tilman's resource competition model (Tilman 1994)], and grazing pressure together affect phytoplankton growth and productivity (Bakker et al. 2005).

As part of their everlasting arms race (Smetacek 2001), different phytoplankton groups have evolved developing different strategies and adaptive mechanisms to overcome such growth- and productivity-limiting factors. Anti-grazing strategies are the most sound. Flagellated phytoplankton, for instance, use their flagella to escape grazing and move towards light and nutrients, diatoms were found to utilize ferritin to store iron (Marchetti et al. 2009) or domoic acid to bind beneficial iron or harmful copper (Wells et al. 2005). While cyanobacteria use gas vacuoles to regulate its buoyancy and consequently its access to light (Reynolds et al. 1987), *Phaeocystis* colonies mucilage regulates its buoyancy. Moreover, morphological changes such as increase in cell size, chain- and colony-formation, and increase in silicification in diatoms are among defense strategies against grazers (Smetacek et al. 2004; Marchetti & Cassar 2009). Phytoplankton productivity –or limited productivity— does not only affect the regional food chain the algal class is part of, but it also affects the global biogeochemical recycling of nutrients by bacterioplankton (Azam & Malfatti 2007) [modeled by (Boyd & Doney 2002)].

1.2 The Iron Hypothesis

Three (Smetacek et al. 2004; Dugdale & Wilkerson 1991; Assmy et al. 2013; de Baar et al. 2005) of the largest oceanic ecosystems, and collectively form up to one-third of the world ocean area (Boyd et al. 2007), have been stamped as High-Nutrient Low-Chlorophyll (HNLC) regions, which, as the name implies, are characterized by high

concentration of nutrients, N in particular, yet show low phytoplankton nitrate uptake rates (Dugdale & Wilkerson 1991) and consequently decreased growth rates and productivity. Grazing pressure has long been suggested as the most productivity-limiting factor until the work of Martin *et al.* demonstrated that iron is the key limiting micronutrient that drives N uptake by endemic phytoplankton community members of the three HNLC regions, the subarctic (Martin & Fitzwater 1988) (*in vitro*) and equatorial Pacific Ocean (Martin et al. 1994) and the Antarctic/Southern Ocean (Martin et al. 1990) (*in situ*), structuring the foundation of the "Iron Hypothesis (Martin 1990)" (Marchetti & Cassar 2009; de Baar et al. 2005). Iron fertilization experiments demonstrated the increase of carbon fixation and export (Smetacek et al. 2012), and nitrate and silica utilization (Assmy et al. 2013; Hutchins & Bruland 1998; Whitney et al. 2005) following iron supplementation of HNLC regions phytoplankton assemblage linking the global climate with oceanic biogeochemistry (Boyd 2002a; Boyd & Ellwood 2010; de Baar et al. 2005; Charles et al. 1991; Mortlock et al. 1991). Dissolved iron concentration in HNLC regions can be as low as < 1 nM in the Southern Ocean (Smetacek et al. 1997).

Iron biogeochemical cycle has been extensively studied and modeled [reviewed in (Boyd & Ellwood 2010)]. Iron supply to the ocean includes: atmospheric dust deposition, shallow sediments mixing and hydrothermal fluids, while coastal iron supply includes continental and ice depositions (Boyd & Ellwood 2010). The major source of iron in the offshore environments is dust deposition, which is particularly low in the Southern Ocean (Strzepek et al. 2011), for instance, leading to dependence on upwelling of iron through deep mixing. Iron forms are: the bioavailable bound to bacterially secreted siderophores iron, and the unavailable trivalent or divalent inorganic iron species. The complexation of iron to organic ligands keeps iron concentration at near constancy along depth (Boyd & Ellwood 2010). The fate of ocean trivalent iron is either to be captured by bacterial ligands and kept in solution or reduced by algal membrane reductases, or to be exported to the sea bottom in an aggregated form that can be also recycled by viruses, heterotrophs, mixotrophs and microzooplankton grazers depending on the depth (Boyd & Ellwood 2010; Strzepek et al. 2011).

1.2.1. The Southern Ocean the Largest HNLC Region: Its Oceanography and Phytoplankton

The Southern Ocean (SO) evolved throughout the past 35 million years into a highly productive region through the development of deep circulations upwelling nutrients to the surface (Kennett 1977; Sigman & Hain 2012). The SO is the largest (Smetacek et al. 2004) typical iron-availability-driven HNLC region [supported by (Charles et al. 1991; Mortlock et al. 1991)].The SO phytoplankton assemblage productivity regulate the sinking and production of the Earth's greenhouse gases carbon dioxide (CO_2) and dimethyl sulfide (DMS), respectively (Boyd 2002a). Although it is the 4th largest ocean, comprising 20% of the world ocean area, the SO directly regulates both the global climate and geochemistry (Boyd 2002a). Understanding the physical oceanography of the SO helped resolving its global impact.

The SO extends from 60° South circulating Antarctica with an area of 20,327,000 km^2. It comprises water masses from the Atlantic, Indian and Pacific Oceans and includes Amundsen Sea, Bellingshausen Sea, Ross Sea, Scotia Sea, and the Weddell Sea (Stewart 2009; Gaebler-Schwarz 2009) (Figure 2). The main and most prominent circulation feature of the SO is the wind-driven and seafloor topology-affected Antarctic Circumpolar Current (ACC) (Figure 2A), which transfers 95-158 million m^3 per second of the SO water masses running clockwise from west to east reaching its maximum around July (i.e., the late Antarctic winter and early spring) (Stewart 2009). The distinctly physically and chemically different water masses of the SO are separated by fronts (Figure 2B). The temperature of Antarctic Surface Water (ASW) ranges from, from the north to the south, 5 °C to -1.86 °C (Gaebler-Schwarz 2009) and its salinity ranges with depth from 33.5 to 34.0 psu (i.e., Practical Salinity Units) decreasing in the summer following ice melting (Tomczak & Godfrey 2004; Riffenburgh 2007).

Two circulation systems affect ASW characteristics; the strongest deepest circulation system, ACC, that mixes and redistributes the deep water from all oceans, and the Antarctic Convergence that directs the flow of the ASW towards the warmer less dense SASW. The circulation systems mix the cold oxygen-poor ASW with the high-salinity Atlantic deep water as well as the near freezing point highest-salinity Ross and

Weddell gyres waters and ABW leading to the upwelling of CO_2 and nutrient-rich as the ASW at the Antarctic Divergence and downwelling of the dense salty water (Tomczak & Godfrey 2004; Stewart 2009). The nutrient-rich ASW next mixes with the Subantarctic Surface Water (SASW) and sinks to become the Antarctic Intermediate Water (AAIW). Both the Antarctic Bottom Water (ABW) and AAIW are distributed by ACC to all the oceans distributing both water mass and heat to the northern hemisphere.

The high sensitivity of the SO suggests that changes in climate and atmospheric CO_2 affect its water masses characteristics as well as the ACC velocity. Consequently, these changed waters will be distributed to the global ocean [reviewed in (IPCC Working Group II 2001)].

The climatological influence of the SO which was prominent in the past (Sarmiento et al. 1998), the present (Broecker & Henderson 1998; Moore et al. 2000) cited in (Boyd 2002a; Gaebler-Schwarz 2009)], and would extend for centuries to the future time (Takahashi et al. 2009; IPCC Working Group II 2001; Hoppe et al. 2013) is a direct result of its oceanographic characteristics. Because it drives the world ocean's circulation and water mass and heat exchange, the SO is one of the most important regulators of the Earth system (Boyd et al. 2007). Back to and in support of the iron hypothesis, iron supply either by the atmospheric dust in the past at the mid-Pleistocene (1.25 million years ago) (Martínez-Garcia et al. 2011) and the Last Glacial Maximum (LGM) (21,500 years ago) (Moore et al. 2000; Martinez-Garcia et al. 2014) or at the present days by winter ACC-mixed subsurface dissolved iron (Tagliabue et al. 2014) has been observed to drive the SO productivity, the increase in CO_2 export and the decrease in atmospheric CO_2 [(Anderson et al. 2009; Sigman et al. 2010) cited in (Smetacek et al. 2012)].

Directed by importance of its ecological niche, the SO phytoplankton community is a major player in the global biogeochemical cycle of nutrients. In order to resolve such impact, the SO phytoplankton assemblage, its key species, and the factors affect its formation, distribution and processes needs to be studied. Boyd reviewed the history of studying the SO assemblage (Boyd 2002a) demonstrating that the assemblage can only be studied in the light of the SO biogeochemistry, physics, and hydrology. The SO

phytoplankton growth is affected by the light intensity (Antarctic water tongue), grazing (in coastal waters), micronutrients (iron in the SO and silica in the northern part of the Antarctic Zone) availability, and optimum growth temperature (Boyd 2002a), therefore phytoplankton flourish during the Antarctic spring and bloom afterwards providing food for the grazers to grow during the summer. This review focuses on the assemblages of the Atlantic and Pacific sectors in comparison to that of the Ross Sea as it is the ecological niche of the subject species of the study.

In regard to the Atlantic sector of the SO, Smetacek *et al.* have assessed the biodiversity of the ACC phytoplankton assemblage in the Antarctic spring (October and November) in 1992 (Smetacek et al. 1997) and reported that it was diatoms-dominated (~130 species) at the highest SO iron concentration (1.14-1.87 nM) in the Polar Frontal Zone in addition to other algal groups such as (with focus on eukaryotes): pico- and nanophtoplankton, prymnesiophytes, autotrophic dinoflagellates, cryptophytes, prasinophytes and chlorophytes (Smetacek et al. 1997). Diatom blooms of *Fragilariopsis kerguelensis* and *Corethron* spp. were reported in the Polar Frontal Zone in addition to the weakly-silicified closer to the shore *Pseudo-nitzschia* spp., *Chaetoceros* spp., and *Thalassiosira* and the heavily-silicified *Thalassiothrix antarctica* (Assmy et al. 2013; Smetacek et al. 1997). N:P water ratios were found to be low in the iron limitation-adapted diatom blooming (*F. kerguelensis*) sites. Limitation in iron impaired N utilization (Smetacek et al. 1997). Supporting the iron hypothesis, CO_2 export levels were observed to be the highest in the Polar Frontal Zone (Smetacek et al. 1997) by the ubiquitous diatom *Chaetoceros* (Assmy et al. 2013), while the major contributors of silica export were the heavily-silicified *F. kerguelensis* and *T. antarctica* (Assmy et al. 2013).

Further to the north of the slightly iron-rich Polar Frontal Zone, pico- and nanophyoplankton dominated as iron becomes very limited for diatoms to grow. Moving upwards along the oceanic food chain, mesozooplankton (e.g., copepods in the Polar Frontal Zone and salps in the northern direction) were reported to be the grazers of the microphytoplankton species (Smetacek et al. 1997) adding another layer of limitation to the SO primary productivity and growth. DMSP production levels were reported and served as a proxy for the other bloom-forming algal group, *Phaeocystis antarctica*,

ecology in the ACC. The highest abundance of *P. antarctica* was reported to be in the Polar Frontal Zone in the sites dominated by the small algal classes of sizes ranging 5-20 μm (Smetacek et al. 1997). Similar findings were reported by Bracher *et al.* of the SO summer assemblages indicating that diatoms dominated the Polar Frontal Zone, however, *P. antarctica* dominated the Marginal Ice Zone (Bracher et al. 1999).

In contrast to the Polar Frontal Zone, in a comparable region, the Ross Sea, *P. antarctica* dominated the phytoplankton assemblage of the deeply mixed water column, marginal ice zones and polynyas (Arrigo et al. 2010) due to their ability to maintain their photosynthetic capacity at lower light intensities than diatoms. In addition, *P. antarctica* less adaptability to higher intensities (Arrigo 1999; Smith et al. 2003; Boyd 2002a; Arrigo et al. 2010) drove its quick blooming before diatoms in the late spring (December) (Arrigo 1999). *P. antarctica* exports C utilizing the available NO_3 reserves at faster rates than that of diatoms (Arrigo 1999) suggesting that, unlike diatoms, iron might not be required for N utilization in *Phaeocystis* specially that the effect of iron supply on *Phaeocystis* is yet to be investigated (Olson et al. 2000; Boyd 2002a). Other factors controlling the blooming of *P. antarctica* in the Ross Sea and the blooming of diatoms in the Polar Frontal Zone in addition to iron, nutrients and light intensity are also yet to be investigated (Boyd 2002a).

In regard to the phytoplankton assemblage of the Pacific sector of the SO, during the late summer of 2010, automated ribosomal intergenic spacer analysis (ARISA) and 18S rRNA gene 454-pyroequencing showed that Pacific sector of the SO [i.e., the iron-rich Amundsen Sea (Alderkamp et al. 2012)] was dominated by diatoms (*Eucampia* spp., *Pseudo-nitzschia* spp. and *Chaetoceros* spp.), *P. antarctica* dominating the deeply mixed regions (or perhaps after the diatom bloom took place), while ciliates dominated underneath the iced regions (Wolf et al. 2013).

1.2.2. Iron Fertilization Experiments in the SO: Phytoplankton Behavior *In Situ* and *In Vitro* under Iron Enrichment

Since the establishment of the iron hypothesis, the effect of iron availability on phytoplankton growth–in general— and diatoms growth –in particular— in the HNLC

regions has been extensively investigated. The role of iron in photosynthesis, respiration, elemental composition (Marchetti & Cassar 2009), grazing (Assmy et al. 2013), N utilization (Dugdale & Wilkerson 1991; Geider & La Roche 1994) and other crucial metabolic processes in diatoms and Haptophyta have been investigated *in situ* (de Baar et al. 2005) and *in vitro* either at a single species level (Allen et al. 2008; Lommer et al. 2012; Hoffmann et al. 2007; Strzepek et al. 2011) or at a mesocosm level (Marchetti et al. 2012; Bertrand et al. 2011).

The primary aims of iron fertilization experiments are: (1) testing the iron hypothesis and whether iron is the main limiting factor for oceanic primary productivity, (2) elucidating the mode through which iron enhances nutrient utilization and C export (Boyd et al. 2007; Smetacek et al. 2012), (3) understanding the ocean biota dynamics upon iron addition and its underlying physics, (4) constructing the ocean biogeochemical cycles under iron enrichment, and (5) understanding phytoplankton the evolution of iron scavengers in pre- and post-iron addition (Boyd et al. 2007). Nevertheless, the history of iron fertilization as a potential geoengineering approach has been discussed (AL Strong et al. 2009) as an attempt to demolish its rationale (Aaron Strong et al. 2009; AL Strong et al. 2009).

Following *in situ* ocean fertilization attempts of Martin *et al.*, to date, thirteen iron fertilization experiments have been conducted in all three HNLC regions (de Baar et al. 2005; Smetacek et al. 2012; Boyd et al. 2007). Phytoplankton dynamics of the eastern equatorial Pacific Ocean experiments are reported (Landry, Ondrusek, et al. 2000; Landry, Constantinou, et al. 2000) reviewed compared to the SO in (Boyd 2002b). In addition, dynamics of the western (Tsuda et al. 2003; Tsuda et al. 2007) and eastern subarctic Pacific Ocean experiments are reported (Marchetti et al. 2006) and compared to the SO experiments (de Baar et al. 2005). Collaborative efforts of the Scientific Committee on Oceanic Research resulted into an open-access database of the findings of iron enrichment experiments available at the Biological and Chemical Oceanography Data Management Office (Program Iron synthesis http://www.bco-dmo.org/program/2017) (Boyd et al. 2012). Here iron fertilization experiments in the SO and *P. antarctica* dynamics is reviewed in detail.

Five iron fertilization experiments took place in the SO to date (Table 1) in addition to eight experiments carried out in the Ross Sea (Olson et al. 2000). The first four experiments were extensively reviewed and compared to the experiments in other niche (Boyd et al. 2007; de Baar et al. 2005; Marchetti et al. 2006; Trick et al. 2010; Boyd 2002b). As a general protocol, iron has been supplied mimicking the natural iron supply (Boyd et al. 2007) as an acidified sulfurhexafluoride (SF_6) tracer-labeled $FeSO_4$ solution (de Baar et al. 2005; Boyd et al. 2007; Law et al. 1998) and tracked using gas chromatography (Watson et al. 1991). Experiments duration ranged from 13 days to two months and added iron ranged from 490 to 2820 kg (Martin et al. 2013; Boyd et al. 2007; de Baar et al. 2005).

In all iron enrichment experiments, including those in the SO, a 3-step floristic shift (Gall et al. 2001) in the originally picophytoplankton-dominated community took place. First, nanophytoplankton Haptophyta [e.g., *Dicrateria* spp., *Imantonia* spp., and *Chrysochromulina* spp. in addition to the initially present and later blooming *P. antarctica* before declining (Assmy et al. 2007)] blooms were instantly observed until grazed [by ciliates in SOIREE (Gall et al. 2001; Hall & Safi 2001) and in EisenEx (Assmy et al. 2007)]. Following, blooms of weakly-silicified large diatoms took place (e.g., *Pseudo-nitzschia* spp., *Rhizosolenia* and *Thalassiothrix* spp. morphotypes, and *Chaetoceros debilis*). Finally, the originally dominating adapted to iron limitation species (e.g., *F. kerguelensis*) showed a long-term response to iron addition (Marchetti et al. 2006; Trick et al. 2010; Gall et al. 2001; Assmy et al. 2007). Haptophyta contribution was observed to be kept constant by grazing pressure after initial peaking in its DMSP productivity (de Baar et al. 2005; Gall et al. 2001; Boyd et al. 2005; Boyd et al. 2007), while diatoms blooms were suggested to be terminated due to silica depletion along with diatom species succession rather than grazing pressure (Boyd et al. 2007; Boyd et al. 2005).

To date, the effect of iron enrichment on Haptophyta might seem to be under-investigated during iron fertilization experiments (Olson et al. 2000; Boyd 2002a; Gall et al. 2001) in spite of the reported interesting findings. For instance, Assmy *et al.* (EisenEx), Hoffmann *et al.* (EIFEX), and Mazzocchi *et al.* (LOHAFEX), reported a

marked increase in colonial *P. antarctica* accompanied by a decline of solitary cells suggested to be an anti-grazing mechanism (Assmy et al. 2007; Hoffmann et al. 2006; Mazzocchi et al. 2009). Furthermore, in iron enrichment experiments, the silica-poor water was dominated by *Phaeocystis* spp. unlike the silica-rich water (Coale et al. 2004). Table 1 summarizes the reported abundances of Haptophyta (*P. antarctica* in particular) before and after iron repletion throughout the SO iron enrichment experiments.

Two ecological drawbacks have been postulated –and investigated– of anthropogenic iron supply: the increase of the production of the amnestic shellfish poisoning biotoxin, domoic acid, and the precursor of the greenhouse gas DMS, DMSP. The neurotoxin domoic acid is produced by the coastal species of the cosmopolitan genus *Pseudo-nitzschia* accumulated in shellfish that is consumed by humans. The toxicity of the water-soluble neurotoxin is due to its analogy to the neurotransmitter glutamic acid (Jeffery et al. 2004; Mos 2001). It has been suggested as an iron- and copper-scavenger in cases of limitation and toxicity, respectively (Wells et al. 2005; Mos 2001; Maldonado et al. 2002; Rue & Bruland 2001) improving its competitiveness (Prince et al. 2013). During iron fertilization experiments, *Pseudo-nitzschia* showed exponential high increase (Assmy et al. 2007), however, domoic acid levels were undetectable in the SO-endemic [on EisenEx preserved samples (Assmy et al. 2007)] and in the subarctic Pacific Ocean-endemic [*in vitro* on SERIES-related isolates (Marchetti et al. 2008)], while it reached lethal concentrations in the SO-endemic *Pseudo-nitzschia* spp. [on SOFeX preserved samples (Silver et al. 2010)]. In their *in situ*, and continuous and batch *in vitro* experiments, Trick *et al.* reported variable elevations in cellular domoic acid levels in iron-enriched subarctic Pacific-endemic *Pseudo-nitzschia* spp. giving them a competitive advantage over neighbor diatoms. Such elevation raised an alarm regarding potential copper contamination of the enrichment iron that might be used in potential large-scale geoengineering attempts (Trick et al. 2010), however, no mortalities were recorded from iron-enriched oceanic *Pseudo-nitzschia* spp. (Trick et al. 2010; Silver et al. 2010).

Regarding the longer-term hazard, DMSP production by *Phaeocystis* spp. (Turner et al. 2004; Turner et al. 1996) endemic to the subarctic Pacific has been modeled following SERIES (Le Clainche et al. 2006; Levasseur et al. 2006; Merzouk et al. 2006).

Surprisingly, in spite of its general assessment throughout SOIREE (Gall et al. 2001; Hall & Safi 2001) and EisenEx [reviewed in (Boyd et al. 2007)], and specific assessment throughout LOHAFEX (Smetacek 2009), similar modeling is yet to be developed for the other polar HNLC region, the SO (Smetacek & Naqvi 2008) to assess a hazardous consequence of iron fertilization and a potential replacement of one greenhouse gas (i.e., CO_2) with another (i.e., DMS) (Boyd et al. 2007) that might decrease global temperature by 1-2 °C (Smetacek & Naqvi 2008). Furthermore, DMS release from DMSP due to DMPSlyase activity has been suggested as a result of grazing pressure (Hall & Safi 2001; Wolfe et al. 1997; Smetacek 2009). Thus, studying *P. antarctica* at a molecular level is crucial to gain insights into the intrinsic factors influencing its dynamics, response to iron depletion and repletion, as well as to grazing and DMSP production.

1.2.3. Phytoplankton Iron Utilization and Adaptation to Iron Limitation

Iron has been proved to be essential for redox-based reactions, which applies to photosynthesis, respiration, and nitrate and sulfur utilization through reduction in phytoplankton [reviewed in (Raven 2013)]. Consequently, iron is limiting for phytoplankton growth. As a result of the Great Oxygenation Event, the soluble biologically active ferrous ion [Fe(II)] became less abundant in the ocean, thus phytoplankton needed to: (1) develop strong iron-binding organic ligands (i.e., siderophores) of high affinity to the more abundant ferric ions [Fe(III)], (2) develop transmembrane Fe(III) and Fe(II) import systems comprising ferrireductases, ferroxidases and permeases, (3) reduce the overall iron demand, and (4) develop non-reactive-oxygen-species-generating iron storage systems in the already iron-limited environments [reviewed in (Raven 2013; Morrissey & Bowler 2012)].

Iron uptake and utilization have been studied, modeled and compared in various algal groups using various methods in single organisms and on mesoscale (Raven 2013; Sutak et al. 2012; Naito et al. 2008; Miller et al. 2014; Strzepek et al. 2011; Sunda & Huntsman 1995; Morrissey & Bowler 2012; Marchetti et al. 2009; Hartnett et al. 2012; Maldonado et al. 2001). Two strategies of iron acquisition are adopted by in land plants: strategy I for unbound iron uptake, which is Fe(III)-reductase-dependent involving

$Fe(II)/H^+$ symport uptake system, and strategy II, which is sidrophores-dependent involving a nonreductive uptake system of Fe(III). Strategy I is more prominent in all algal groups (Hartnett et al. 2012; Raven 2013; Sutak et al. 2012). Iron utilization in diatoms based on genomic and experimental data of the two diatom model organisms (*Thalassiosira pseudonana* and *Phaeodactylum tricornutum*) and others has been modeled (Morrissey & Bowler 2012) to be a reductive utilization of sidrophores-bound Fe(III) or an oxidative utilization of Fe(II), which are later transported into the cell to be either used or stored as Fe(III) in ferritin (as in *Phaeodactylum tricornutum*) or as Fe(II) vacuoles (as in *Thalassiosira pseudonana*) (Morrissey & Bowler 2012).

In iron-limited environments, diatoms have developed various mechanisms to efficiently use the available iron. Here a few are listed. Diatoms have been observed to down-regulate the iron-dependent photosystem I cytochromes and rely on the iron-economic photosystem II copper-containing plastocyanins, or the least efficient photorhodopsins (Raven 2013). All diatoms, except the coastal *Thalassiosira pseudonana*, have copper-zine superoxide dismutases as an alternative to the iron-based reactive oxygen species defense. They also possess flavodoxin instead of the iron-containing ferredoxin in their electron transport system. In addition, diatoms increase their iron uptake through acquiring genes to utilize the bacterial siderophores-bound iron and to store it (Strzepek et al. 2011). Finally, diatoms decrease their overall iron requirements by reducing their cell size, increasing their pigment content rather than photosynthetic units [reviewed in (Raven 2013) and explained in (Morrissey & Bowler 2012)]. In order to infer similar iron limitation adaptive mechanisms in haptophytes, light will be shed on their biology and evolution.

1.3 Haptophyta

"Haptophyta" refers to the nonoflagellates mostly haptonema-containing partially calcium carbonate-mineralized (i.e., "coccolithophores") monophyletic group belongs to the supergroup, Chromalveolata. Secondary endosymbiosis gave rise to chromalveolates through the engulfment of a red alga by their common ancestor that has been illustrated to be predated by an ancient green algal endosymbiotic event (Moustafa et al. 2009).

Supporting the hypothesis on their origin, *P. antarctica* and *P. globosa* plastid and mitochondrial genomes analysis showed that the red-algal-derived plastid-bearing *Phaeocystis* has mutation rates consistent with the red alga *Porphyra* that, unlike land plants, has been reported to have high mitochondrial mutation rates (Smith et al. 2014).

Stramenopiles, cryptophytes and haptophytes are the chromists of Chromalveolata [reviewed in (Hackett et al. 2007; Moustafa 2009)] diverged 1,300 million years ago (Yoon et al. 2004), cryptophytes and haptophytes were thought to be split early (Moustafa et al. 2009; Hackett et al. 2007), however, plastid multi-gene phylogenomic analyses suggested a stramenopiles-haptophytes split 1,047 million years ago, which took place after cryptophytes split 1,255 million years ago (Yoon et al. 2004). Phylogenomic analyses have been conducted to infer whether the closest relative to the chromists lineage, cryptophytes, is haptophytes (Hackett et al. 2007) or the plastid-lacking lineage katablepharids (Burki et al. 2012). Haptophytes have been suggested to be involved in tertiary endosymbiotic events early throughout dinoflagellates plastid evolution (i.e., plastids containing 19'-hexanoyloxyfucoxanthin/fucoxanthin and chlorophyll $c_1 + c_2$) (Yoon et al. 2002) consistent with haptophytes-specific pigment composition.

Haptophyta phylogenetics has been revised based on the 18S rDNA analysis by Edvardsen *et al.* (Edvardsen et al. 2000) and later by Sáez *et al.* (Sáez et al. 2004) leading to the division of Haptophyta into two classes: Prymnesiophyceae and Pavlovophyceae. The 18S rDNA analysis also subdivided Prymnesiophyceae into four orders: the non-calcified Phaeocystales and Prymnesiales, and the calcified Coccolithales and Isochrysidales (e.g., *Emiliania huxleyi*). Prymnesiales were subdivided into two families: the harmful algal blooms Chrysochromulinaceae and Prymnesiaceae, and later Chrysocampanula (Edvardsen et al. 2000; National Center for Biotechnology Information n.d.; Beszteri 2011). Fossil coccolith (i.e., a calcified scale of coccolithophores) records are paleo-temperature (Ternois et al. 1997) and -acidification (Stoll et al. 2009) proxies [see also (Henderiks 2008)].

Haptophyta include plastid-containing and plastid-lacking mostly unicellular mostly marine members (Billard & Inouye 2004; Andersen 2004; Medlin 2009). Haptophytes morphology is characterized by the presence of a flexible microtubular

structure named haptonema that is suggested to be involved in feeding and attachment and lies between two flagella and is lost in few haptophytes. In addition, haptophytes are characterized by the presence of 1-2 four-membrane surrounded golden brown plastids containing chlorophyll *a+c* and lack girdle lamella (Andersen 2004; Medlin 2009; Sáez et al. 2004; Billard & Inouye 2004) (Figure 3A). Most haptophytes have an unmineralized flagellated haploid-mineralized nonmotile diploid alternative generation life cycle (Billard & Inouye 2004; Sáez et al. 2004).

1.3.1. *Phaeocystis* Systematics, Life Stages, and Genomics

Molecular timing of haptophytes using 18S rDNA and rDNA internal transcribed spacer 1 (ITS1) revealed, respectively, that Phaeocystales diverged ~480 million years ago from Prymnesiophyceae and later (~30 million years ago) the cold water *Phaeocystis* spp. (the Antarctic *P. antarctica* and the Arctic *P. pouchetti*) diverged from the warm water *P. globosa* (Medlin, 2009; Medlin and Zingone, 2007). The latter split followed the evolution of ACC that insulated Antarctica from the warm water input and the formation of Drake Passage in the SO, both led to the bipolar distribution of *Phaeocystis* across the equatorial water facilitated by the cold climate period at the time (Darling et al. 2000; Medlin & Zingone 2007).

The cell structure of the genus *Phaeocystis* is similar to those of sister Haptophyta members. Flagellates [haploid motile scaled cells escape the colony under growth limitation conditions; variable sizes (Figure 3B)], macroflagellates and/or attached aggregates (diploid nonmotile attached to diatoms spines at low turbulence; ~4.2-9.8 µm in diameter), and colonial cells (diploid nonmotile; 4-6 µm colonial cell diameter in colonies up to 2000 µm in diameter) are the identified life cycle stages of *P. antarctica* (Gaebler-Schwarz et al. 2010; Peperzak & Gäbler-Schwarz 2012; Zingone et al. 2011). Pigment composition of *Phaeocystis* has been also determined to include diatoxanthin, diadinoxanthin, 19'-hexanoyloxyfucoxanthin and fucoxanthin in addition to chlorophyll *a+c* consistently among different *Phaeocystis* species (Vaulot et al. 1994; Andersen 2004).

The molecular information on *Phaeocystis* is particularly limited, however, ongoing collaborative efforts are being conducted to sequence *P. antarctica* nuclear genome. *P. antarctica* genome was estimated by flow cytometry to be of size of 160 million bases (Mb) (Vaulot et al. 1994; Phaeocystis.org n.d.), larger than the recently released genome of the sister species, *E. huxleyi*, which has been reported to be of size of 141.7 Mb in the haploid cells as well (Read et al. 2013). Interestingly, flow cytometric analysis of *Phaeocystis* species has revealed a GC content of 54% of its genome (Vaulot et al. 1994; Phaeocystis.org n.d.), lower than that of *E. huxleyi* (65%) (Read et al. 2013). At the time of conducting this study, the plastid (~105 kilo bases; Kb) and mitochondrial (~30 Kb) genomes sequences of *P. antarctica* and *P. globosa* have been released (Smith et al. 2014).

Phaeocystis is a cosmopolitan genus and dominates the nanophytoplankton fraction (Schoemann et al. 2005), however, six species have been fully characterized to date, three of which form colonies (*P. antarctica*, *P. pouchetti* and *P. globosa*) (Peperzak & Gäbler-Schwarz 2012) and consequently blooms (Schoemann et al. 2005) (Figure 3C), along with the Mediterranean *P. jahnii* and *P. cordata* (Zingone et al. 1999) and *Phaeocystis* spp. in the coastal toxic assemblage (Zingone et al. 2006) and the Australian *P. scrobiculata* (Medlin & Zingone 2007). *P. pouchetti* and *P. globosa* have been reported to be toxin-producing and ecologically harmful [reviewed in (Edvardsen & Imai 2006; Andersen 2004; Medlin & Zingone 2007; Schoemann et al. 2005)].

1.3.2. *Phaeocystis antarctica* Blooms and DMS Production

Blooms of *Phaeocystis* are exceptionally unique, as they give colonial *Phaeocystis* a competitive growth advantage [storing macro- and micro-nutrients (Liss et al. 1994)]. They are also of a global biogeochemical beneficial (due to the high productivity and high contents of C associated with the blooms polysaccharide matrix to be recycled), harmful [due to the contribution of the globally distributed blooms to atmospheric sulfur (S)], and ecological impacts (due to its well-document hindrance of fisheries industries) (Schoemann et al. 2005). Typically, *P. antarctica* blooms in the SO, *P. pouchetti* blooms in the Arctic Ocean and the southern coasts of China, while *P.*

globosa blooms in the North Sea (van Duyl et al. 1998) (Figure 3C) (Schoemann et al. 2005). Here *P. antarctica* blooms formation factors and fate are reviewed in details.

Blooms, by definition, indicate for a cell density exceeding 10^6/L (Schoemann et al. 2005) which is consistent with the largest *P. antarctica* reported bloom in the Ross Sea in December time. Light (due to adaptation to low light intensity in deep mixed layers), temperature (optimally from -2 to +2 and not exceeding 10 °C), and macro- (by converting non-bioavailable HCO_3^- to CO_2 using extracellular carbonic anhydrase, and due to the abundance of N and P in the SO) and micro- (by colonies mucus complexation of iron and manganese) nutrients scavenging are among the success factors of *P. antarctica* bloom-formation (Schoemann et al. 2005; van Hilst & Smith 2002). On the other hand, grazing [in a size-dependent fashion or by release of acrylic acid from its precursor DMSP (Liss et al. 1994) in *E. huxleyi* (Verity et al. 2007)], viral infections, and sinking (that is minimal at colonial stage) are considered among *Phaeocystis* bloom-termination factors. The skin of *P. globosa* protects it against viral attacks (Schoemann et al. 2005) probably of the bloom-terminating effect of the megavirus PgV-16T, *P. globosa*, which genome has been revealed. The virus has been classified as a sister to two Antarctic megaviruses which their hosts are yet to be revealed (Santini et al. 2013) and it has not been reported to be *P. antarctica*. Ungrazed blooms undergo bacterial degradation and thus do not contribute in deep C export (Schoemann et al. 2005; Verity et al. 2007).

DMS production has been linked to *Phaeocystis* bloom occurrence (Stefels et al. 2007; Verity et al. 2007), nevertheless, dinoflagellate species produce DMSP as well (Caruana et al. 2012). The SO *P. antarctica* production of DMSP, for instance, has been calculated to contribute 70% of atmospheric DMS contents in the Southern Hemisphere and 5-10% globally (Schoemann et al. 2005; Kettle & Andreae 2000). DMSP production has been postulated as a defense mechanism against low temperature, reactive oxygen species damage, osmotic pressure changes, and grazing reviewed in (Gaebler-Schwarz 2009; Schoemann et al. 2005; Verity et al. 2007)]. Oceanic DMS contribution has been modeled (Kettle & Andreae 2000) and a database of DMS metadata has been assembled

(Kettle et al. 1999). Transcriptomic assay of the genes involved in DMSP biosynthesis (Summers et al. 1998) in bloom time is yet to be achieved.

1.3.3. Iron Utilization in Haptophytes

Iron uptake (Sutak et al., 2012), transport and storage (Hartnett et al., 2012) have been studied in *E. huxleyi*, a close relative to *Phaeocystis*. *Emiliana*, the most abundant haptophyte has been assumed an important role in the global carbon cycle but it has never been reported to be influenced by iron enrichment experiments. The iron-binding capacity of this species was observed to be directly proportional to iron concentration in contrary to typical siderophore-dependent mechanisms [(Boye & van den Berg 2000) reviewed in (Sutak et al. 2012; Hartnett et al. 2012)]. In their comprehensive comparative study, Sutak *et al.* reported that *E. huxleyi* neither stores iron nor its growth is dependent on iron concentration. The study also showed that *E. huxleyi* has no spectrofluorimetrically detectable reductase activity, transmembrane electron transport activity, or preference for Fe(III) over Fe(II) suggesting a nonreductive, non-siderophore-dependent uptake mechanism, however (Sutak et al. 2012). Contrastingly, cell surface reductase assay confirmed that *E. huxleyi* takes up Fe(III) in a concentration- and temperature-dependent yet form-independent manner (Hartnett et al. 2012). Upon uptake, *E. huxleyi* was not observed to reduce Fe(III), but to store it in a plant-ferric-phosphate-like/animal-ferrihydrite-like ferritin, though neither reductase- nor to ferritin homologs were found in the *E. huxleyi* genome (Hartnett et al. 2012). Another coccolithophore, *Cricosphaera roscoffensis*, showed preference to bound iron utilizing it in a reductive uptake manner (Naito et al. 2008).

On the contrary to diatoms and *E. huxleyi*, iron acquisition, transport, storage and adaptive mechanisms in the low-iron-adapted haptophyte, *P. antarctica*, are yet to be fully revealed specially in the lack of genomic data (Morrissey & Bowler 2012; Hartnett et al. 2012). *P. antarctica* behavior in iron enrichment experiments strongly suggests its adaptation to the SO iron-limited environment (reviewed in 1.2.3). *In vitro* studies supported the iron enrichment observation and reported that *P. antarctica* half-saturation

constant of iron required for growth can be as low as 0.005 nM and as high as 0.258 nM dissolved iron depending on cell size [reviewed in (Strzepek et al. 2011)].

Strzepek *et al.* have successfully characterized *P. antarctica* iron uptake mechanism by studying short-term and steady-state growth and iron uptake and utilization in *P. antarctica* applying varying iron concentrations, bound iron organic ligands, light conditions, and *P. antarctica* morphotypes (Strzepek et al. 2011). They reported a ferrireducase-based non-ligand-specific iron uptake system overexpressed in *P. antarctica* under iron limitation of high resemblance to the oceanic SO diatoms in the study (Strzepek et al. 2011). *P. antarctica* showed the ability to utilize iron bound to EDTA, aerobactin, enterobactin, ferrichrome and desferrioxamine B (DFB) in a decreasing extent (Strzepek et al. 2011), which provided the basis of the choice of the chelator used in this current study. Furthermore, iron uptake and extracellular reduction rates were reported to be significantly higher under iron limitation than enrichment conditions, uptake:flux ratios were the highest under iron limitation, and iron uptake and growth rates were positively correlated throughout the experiment. Relative to the diatom subjects of the study, *P. antarctica* showed the highest growth rates, Fe:C contents and iron uptake rates. It also showed iron-use efficiencies inversely proportional to its intracellular iron contents which were greatly affected by temperature [table 2 in (Strzepek et al. 2011)]. Overall, *P. antarctica* iron requirements were the lowest in the studied SO species (Strzepek et al. 2011).

As reported during *in situ* experiments, iron directly affected *P. antarctica* colonization (reviewed in 1.2.2). In batch culture experiments, *P. antarctica* showed no colonization under iron limitation, while colonies have been observed upon iron repletion suggesting cell size reduction as a prominent adaptive mechanism of *P. antarctica* towards iron limitation (Strzepek et al. 2011). Furthermore, the colony-forming morphotype showed lower intracellular iron contents under both enriched and limited conditions than those of the solitary morphotype (Strzepek et al. 2011). These findings suggest that colonization upon iron enrichment could be an adaptive strategy for scavenging and complexion more of the now-available iron for time of limitation.

P. antarctica has also been observed to dominate the deeply mixed layers (Wolf et al. 2013) during iron enrichment experiments, and *in vitro* iron uptake rates from the photostable DFB were reported to be 2.2-fold higher under light *vs.* dark conditions which the authors explained as a physiological rather than an photoreduction-related iron uptake response (Strzepek et al. 2011). This might suggest that the non-ligand-specific iron uptake mechanism observed in *P. antarctica* might also be irradiance-independent, a theory that requires measuring intracellular iron content as a function of irradiance change compared to irradiance-dependent iron uptake in the subarctic Pacific Ocean community (Maldonado et al. 2005).

One further indication of *P. antarctica* adaptation to iron limitation (summarized in Table 2) has been studied through studying its pigment composition and photodamage repair under changing irradiances (DiTullio et al. 2007; Arrigo et al. 2010). *P. antarctica* grows at a wide range of light intensities because of its ability to harvest light through the interconversion of its two photopigments, 19'-hexanoyloxyfucoxanthin and fucoxanthin (DiTullio et al. 2007). Upon iron enrichment, prior to spring, a photopigment shift towards the more efficient light-harvesting fucoxanthin has been observed to take place in *Phaeocystis* cells. In contrast, 19'-hexanoyloxyfucoxanthin has been observed to be of higher concentrations in iron-starved *P. antarctica* and was proposed as a photoprotective pigment for fragile iron-limited *P. antarctica* during the Antarctic summer (DiTullio et al. 2007; van Leeuwe & Stefels 1998; Schoemann et al. 2005). Cellular iron demand has been reported to increase under low light intensities due to the increase in the photosynthetic apparatus size [reviewed in (Verity et al. 2007)].

1.4 Molecular Profiling of Nutrient Limitation in Phytoplankton

The advances in spatial and temporal phytoplankton studies' methodology [e.g., pigments concentration qualitative and quantitative determination by coastal zone color scanner (CZCS) in 1978, sea-viewing wide-field of view sensor (SeaWiFS) in 1999 and high-performance liquid-chromatography (HPLC) (Wolf et al. 2013)], the introduction of the "Antarctic Paradox" of the HNLC regions concept (de Baar 1994), and the global initiative Joint Global Ocean Flux Study (JGOFS) in 1989 for studying polar waters

biogeochemistry were the historical landmarks that preceded the synthesis of iron fertilization experiments (Boyd 2002a; de Baar et al. 2005). Iron fertilization experiments indeed benefited from advances in physiological and cytochemical analyses [e.g., pump-during-probe microflowcytometry for single-cell photosynthetic fitness analysis (Olson et al. 2000)], however, with the advance in DNA and RNA sequencing technologies, it is crucial to sharpen the resolution of iron enrichment experiments through (meta-)genomic and (meta-)transcriptomic profiling of (collective) phytoplankton community response to iron addition (Allcock & Strugnell 2012). Unfortunately, older iron enrichment experiments did not benefit from such advance in methodology and relied on physiological and chemical profiling. The latest *in situ* SO iron fertilization experiment, LOHAFEX, fortunately used 454-pyrosequencing to characterize the community composition before and after iron supplementation [(Wolf et al. 2012) and Wolf *et al.* (in preparation)], but the community response towards iron addition was not assessed through functional genomics.

Transcriptomics relied for a long time on the low-throughput technique, expressed sequence tag (EST), and the short tags sequencing, serial analysis of gene expression (SAGE) coupled with quantification methods such as microarray designed based on the sequenced ESTs (Wang et al. 2009).

Species-specific *in vitro* nutrient limitation studies were even more fortunate than *in situ* experiments in terms of sequencing-based functional genomics. In spite of the lack of genomic information, nutrient limitation in model phytoplankton member organisms have been also profiled [e.g., transcriptomic profiling of P and N limitation in *E. huxleyi* using long-SAGE (Dyhrman et al. 2006), and transcriptomic and proteomic profiling of P limitation in *Thalassiosira pseudonana* using transcript tag sequencing (tag-seq) coupled with quantitative shotgun liquid-chromatography mass spectrometry (LC-MS) (Dyhrman et al. 2012)]. Nevertheless, using quantitative transcriptomic techniques, the response of a number of harmful algal bloom-forming species to macronutrient limitation have been profiled. EST sequencing has been widely used in profiling N and P limitation in the dinoflagellate *Alexandrium tamarense* coupled with massively parallel signature sequencing (MPSS) (Moustafa et al. 2010) as well as in the haptophyte *Prymnesium*

parvum (Beszteri et al. 2012) and the dinoflagellates *Alexandrium minutum* (Yang et al. 2011) and *Karenia brevis* (Morey et al. 2011) coupled with microarray hybridization in a time-series fashion.

Transcriptomic studies have recently benefited from the progressive development in sequencing technologies. RNA sequencing (RNA-Seq) has quickly gained advantage over microarray and tiling array, for instance (Wang et al. 2009), especially in studying gene expression of species whose genomes are yet to be sequenced (Haas et al. 2013). The advance in *de novo* assembly approaches [e.g., Trinity (Grabherr et al. 2011; Haas et al. 2013)] made RNA-Seq even more appealing to researchers to use for characterizing novel genes and their isoforms and alternative splicing events. RNA-Seq using Illumina, the highest throughput genome analyzer (160 million reads per run), employs massive parallel sequencing of the clustered amplified fragments of cDNA utilizing a modified method of Sanger's sequencing-by-synthesis of reversing the chain termination by the fluorescently-labelled nucleotides (Bentley et al. 2008) [reviewed in (Moorthie et al. 2011)].

The response of different phytoplankton species towards iron starvation and repletion have been profiled using functional genomic approaches as well. The responses of the well-adapted to iron limitation *Phaeodactylum tricornutum* (combining EST with microarray analysis) (Allen et al. 2008) and *Thalassiosira oceanica* (using 454-pyrosequencing) (Lommer et al. 2012) have been profiled transcriptomically and proteomically. The aforementioned studies set the current knowledge of the molecular bases of iron limitation behind the previously reported phenotypic data and provided models of the cellular response of the most important phytoplankton members towards iron addition and its adaptive mechanisms towards iron limitation.

The HNLC subarctic Pacific community response towards iron addition has been recently profiled at a mesocosm scale using comparative metatranscriptomics (Marchetti et al. 2012) to shed light on both the change in species composition and the species functional response. Being the most abundant and the fastest to respond to iron enrichment, diatoms continued –on the short course of the experiment– to express iron-economic proteins (Marchetti et al. 2012). Interestingly, the only study that assessed the

molecular responses of haptophytes towards iron repletion showed that they exhibited a typical behavior shifting to the use of iron-dependent enzymes (Marchetti et al. 2012). An experiment of longer duration might reveal more regarding the time point at which both diatoms and haptophytes change their expression patterns after "their confidence" of the existence of ample iron contents.

1.5 Study Objectives and Design

Despite of the well-proved iron starvation conditions of the largest HNLC region, the SO, *P. antarctica* recurrent blooms have been documented [reviewed in (Schoemann et al. 2005)] indicating its adaptation to iron limitation. Nevertheless, *P. antarctica* has been reported to be the first to bloom utilizing supplied iron during iron enrichment experiments (Assmy et al. 2007). The mechanism of iron utilization of *P. antarctica* has been elucidated (Strzepek et al. 2011), and the effect of iron limitation and enrichment on its colonization behavior (Assmy et al. 2007; Mazzocchi et al. 2009; Hoffmann et al. 2006), blooming and consequently the global C and S cycles has been reported, yet deeper understanding of the functional changes in *P. antarctica* following iron enrichment over time is crippled by the lack of genomic and transcriptomic data. Time-series transcriptomic and proteomic profiling investigating iron uptake and utilization mechanisms as well as the evolution of the species is required for a better understanding of the impact of iron repletion on cellular processes as well as of the course of adaptation of the species.

Taken all together, *Phaeocystis* has been proposed as an ideal model organism for understanding the role of phytoplankton in global C, N, and S biogeochemical cycles and consequently its pivotal role in Earth system and climate regulation (Schoemann et al. 2005) and even further *P. antarctica*, as the endemic species to the SO, one of the most important Earth system regulators, would fit as such an ideal model organism.

The overall aims of the project, of which this thesis is part, are to: (1) assess the effect of iron repletion on the most abundant haptophyte species *Phaeocystis antarctica* at a transcriptomic level in real-time supported by its physiological response, (2) model

iron utilization and adaptation to iron limitation in this species, and (3) infer the evolution of iron utilization mechanisms in Chromalveolata.

The phase of the project described here specifically aims at: (1) assessing the physiological, morphological and elemental changes of *P. antarctica* under iron-limited and iron-enriched condition, (2) reporting the results of the preliminary assembly and functional characterization of the first sequenced transcriptome of *P. antarctica*, and (3) inferring the statistically and biologically differentially expressed genes and their expression patterns in *P. antarctica* in a time-dependent manner before and after iron supplementation.

The study design is illustrated by Figure 4. Iron-limited stock culture was used to inoculate iron-limited and iron-enriched (control for physiology) replicates. After reaching a suitable cell concentration in the iron-limited replicates, iron was supplemented. Physiological measurements and RNA were obtained in parallel day by day before and after iron supplementation. RNA was extracted from the stock culture (Day 0) and daily from the iron-limited cultures at 4 time-points: Day 2 and Day 3-1 (before enrichment) and Day 3-2 and Day 5 (after enrichment). Total sequenced RNA was assembled, annotated and differential expression was inferred via pair-wise comparisons between days implementing a variety of clustering approaches (Figure 4).

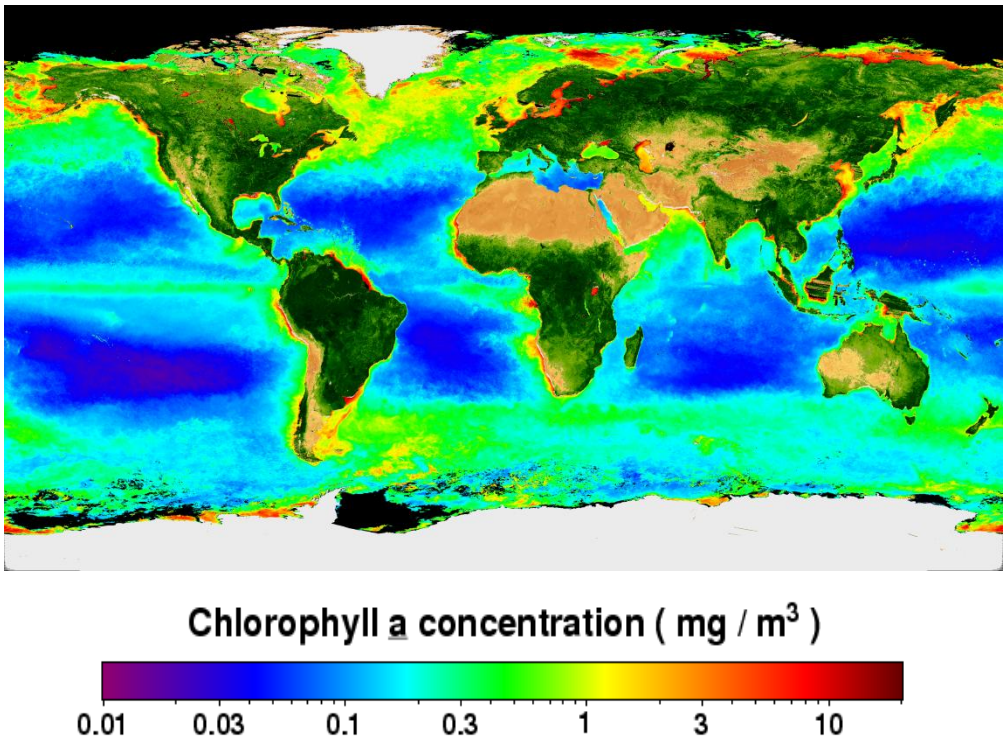


Figure 1 NASA's satellite sensor, SeaWiFS, image of global annual chlorophyll *a* concentration.

Chlorophyll *a* is an indicator comparing phytoplankton biomass to that of land plants in 2010. (http://oceancolor.gsfc.nasa.gov/cgi/l3).

A B

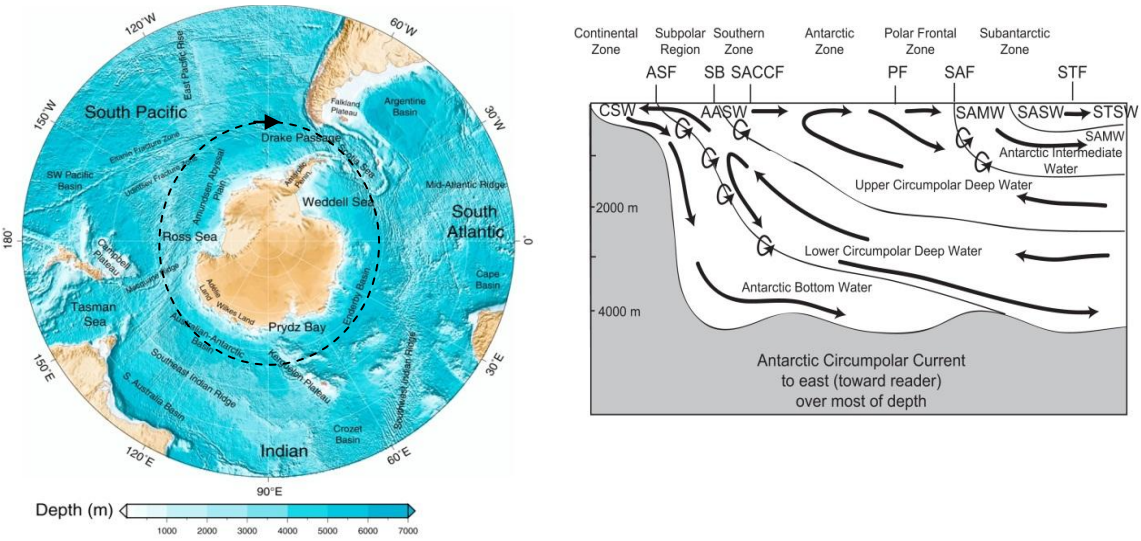
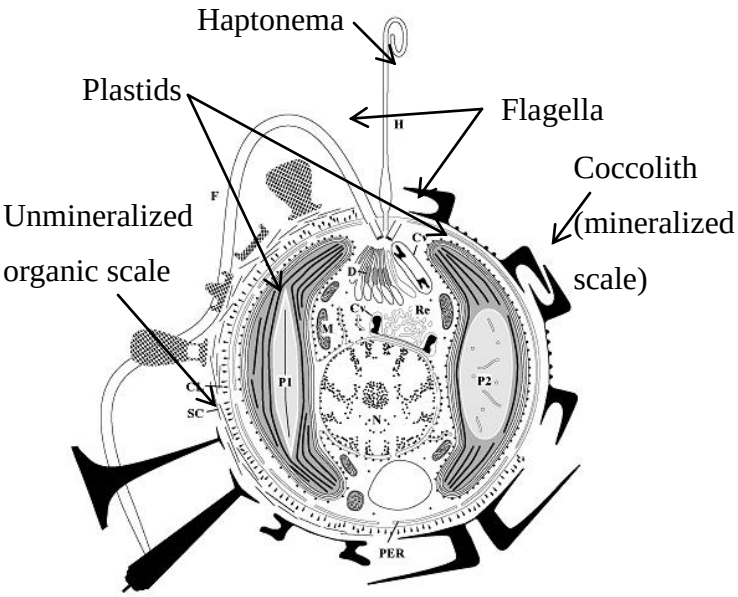


Figure 2 The Southern Ocean geography and hydrology maps.

(A) A map of the SO based on Etopo2 bathymetry data from The National Oceanic and Atmospheric Administration (NOAA) National Geophysical Data Center (NGDC) (2008) (Talley et al. 2011); the Antarctic Circumpolar Current (ACC) is represented by the dashed line and its direction is represented by the black arrow.

(B) A diagram of the Southern Ocean's circulations, fronts, and water masses (Talley et al. 2011). The fronts are from the north to the south: the Subantarctic Front, the Polar Front (the Antarctic Convergence), the Southern ACC Front (the Antarctic Divergence), and the Continental Water Boundary. The SO zones and water masses are from the north to the south: the Subantarctic Zone (the Subantarctic Surface Water; SASW), the Antarctic Intermediate Water, AAIW; the Polar Frontal Zone, the Antarctic Zone (the very cold, high-salinity, dense Circumpolar Deep Water; CDW), the cold fresh Antarctic Surface Water; ASW, and the Continental Zone (the near freezing point highest-salinity dense Antarctic Bottom Water; ABW). Subtropical Front (STF); Subantarctic Mode Water (SAMW); Subantarctic Front (SAF); Polar Front (PF); Southern ACC Front (SACCF); Southern Boundary (SB); and Antarctic Slope Front (ASF).

A



B

C

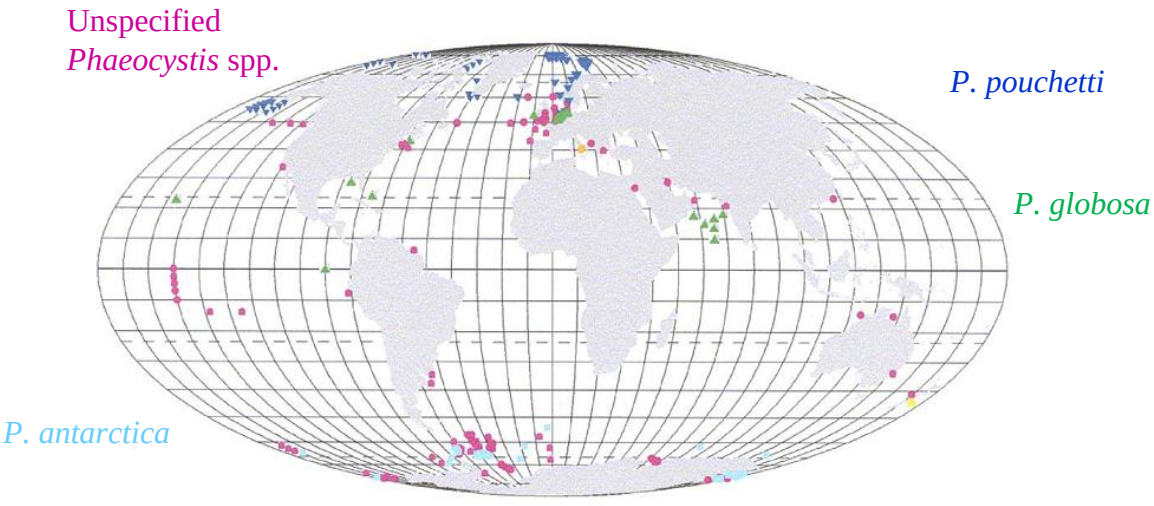


Figure 3 Haptophyta cell structure and *Phaeocystis* microscopy and ecology.

(A) Haptophyta (Coccolithophorids) cell structure [after (Billard & Inouye 2004)]. (B) Scanning electron microscopy micrograph of *P. antarctica* flagellated stage in a dorsal view; arrow points at the haptonema (Zingone et al. 2011). (C) Geographical distribution of *Phaeocystis* [after (Schoemann et al. 2005)]. Figure structure courtesy: Phytoplankton ecology lecture notes, Universität Bremen, 2013.

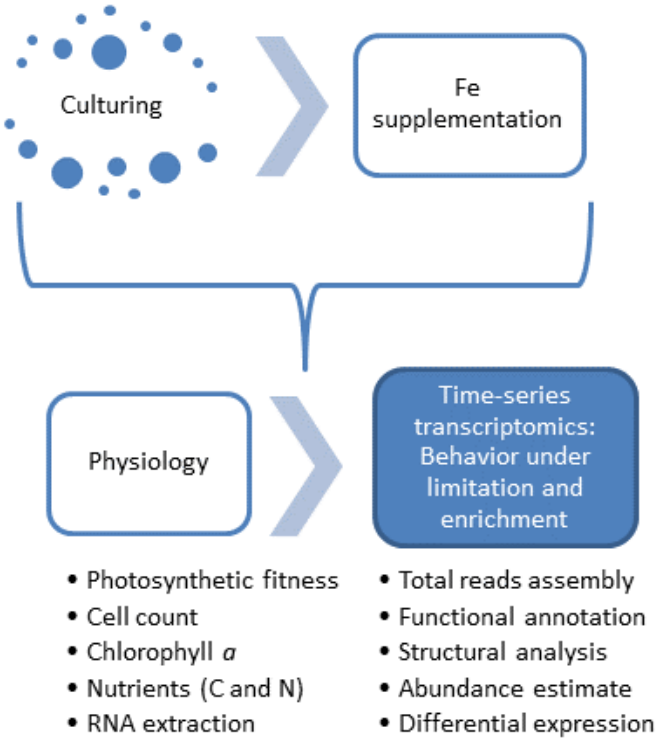


Figure 4 Flowchart of the study design.

Table 1 Haptophyta (*P. antarctica*) response during Southern Ocean iron fertilization experiments.

"Eisen" is German for iron, while "Loha" is Hindi for iron and "Ex" is the abbreviation for experiment (Smetacek and Naqvi, 2008; Thiele et al., 2012).

Experiment	**Location**	**Seasonality**	**Pigment % before/at peak**	**Carbon % before/at peak**	**Morphological and DMSP production remarks**	**References**
Southern Ocean Iron RElease Experiment (SOIREE)	The Australian Antarctic Zone	Late summer, 1999	30-40/35 (day 10)	10/60 (day 6)	Cell size tripled, and DMSP production (day 3-8) and DMS release (day 8-13) reported.	(Gall et al. 2001; Hall & Safi 2001; Boyd 2002b)
EisenEx	The Atlantic Polar Frontal Zone	Late spring, 2000	N/A	0.6/1.1 (day 11)	A sixteen-fold increase in solitary and colonial *P. antarctica*, decrease in solitary but increase in colonial cells, no effect on *Emiliania huxleyi*, and DMSP reported.	(Assmy et al. 2007)
Southern Ocean Iron Experiment (SOFeX)	Northern and southern the Polar	Summer, 2002	N/A	84/56 North, while 33/31 South (day 20)		(Coale et al. 2004; Buesseler et al. 2004)

European Iron Fertilization Experiment (EIFEX)	Front The Atlantic eddy of the ACC	Summer/fall, 2004	37.5/20 (150/250 ng/L)	N/A/< 30	Maximum solitary cells count reported (day 15), while *E. huxleyi* abundance decreased.	(Hoffmann et al. 2006; Smetacek et al. 2012; Assmy et al. 2013; Hoffmann et al. 2007)
Indo-German Iron Fertilization Experiment (LOHAFEX)	The Atlantic sector of the ACC	Summer/fall, 2009	N/A	N/A	Two morphotypes having maximum cell counts associated with highest DMSP levels (day 23) and increase in colonial form abundance reported, however detailed 454-pyrosequencing results (day 10 and day 18) are to be published.	(Martin et al. 2013; Wolf et al. 2012; Thiele et al. 2012; Smetacek 2009)

Table 2 *P. antarctica* colony-forming clone under iron –limited and –replete conditions (Strzepek et al. 2011; DiTullio et al. 2007; Schoemann et al. 2005).

Parameter	**Under iron limitation**	**Under iron repletion**
Growth rate (day^{-1})	0.28	0.52
Cell size	16.9 fL/cell	2-fold increase
Morphology	Only solitary form	Equal mixture of solitary and colonial forms
Photopigments	0.5-fold decrease in 19'-hexanoyloxyfucoxanthin:Chl *a* and 8-fold increase in fucoxanthin:Chl *a*.	
C content (mol/L cell volume)	~15.3 (1.4-fold increase)	11.1
C:N	5-8	6
Fe content (µmol/L cell volume)	31	63.9
Fe uptake	Uptake significantly increases with increasing limited conditions, and increases with increasing Fe.	
Irradiance	Uptake rates increases in light conditions.	
Non-ligand specificity	Extracellular reduction of DFB-bound Fe(III) significantly greater under iron limitation with no effect of dissolved iron concentration.	

CHAPTER 2: MATERIALS AND METHODS

2.1 Cultures and Culturing Conditions

In order for us to decide on our subject clone, we observed the difference in photosynthetic fitness of a number of *Phaeocystis antarctica* clones in different morphotypes isolated from different regions from the SO under iron –limited and –replete conditions as a proxy assessment for their adaptability to iron limitation and the recovery of photosynthetic fitness after iron repletion. A colony-forming clone (Col.) showed a recovery in photosynthetic fitness of the iron-replete culture compared to its iron-limited one and thus we chose it as the subject of the study.

Phaeocystis antarctica clone referred to as *P. antarctica* (Col.) in the colonial stage, isolated from the Ross Sea in 1992, was the test organism of this study. *P. antarctica* (Col.) stock culture was acclimated for five batch dilutions (8, 10, 7, 10, 14 days, respectively) under iron limitation in modified f/2 growth medium (pH 8.0-8.3) for phytoplankton prepared with SO seawater (Guillard & Ryther 1962) (Table 3). Prepared media were stored in the dark at 2 ± 2 °C after filter sterilization. Cultures were maintained at a photoperiod of 16:8 light:dark in the 2 ± 2 °C growth chamber using daylight fluorescent lamps at photon flux density (PFD) as indicated (Gaebler-Schwarz 2009). For acclimation of iron-limited cultures, syringe-filtered (0.22 µm Cole-Parmer® Cellulose Acetate Sterile Filters, Cole-Parmer, Montreal, Canada) iron-free trace metal mixture was used in media preparation and syringe-filtered desferrioxamine B (DFB) chelator (10 nmol/L final concentration) was added to ensure total iron chelation. Cultures were incubated in 2L Nalgene® polycarbonate bottles (Nalgene, New York, USA) detergent- and acid-cleaned (3-day 0.1% citranox-bath followed by 7-day 0.1N HCl-bath). Cleaned bottles were rinsed with ultrapure Milli-Q® water (Millipore, Darmstadt, Germany) 7 times and bagged in plastic bags. Inoculation and culture dilution were conducted in the cleanroom under a laminar airflow hood at 11 ± 1 °C.

Once a suitable quantum photosynthetic fitness (0.27) and cell density ($2.2x10^5$ cell/ml) were obtained, the iron-limited *P. antarctica* (Col.) stock culture was used for

inoculation of 4 treatment and 3 control replicate with a starting cell concentration of $2x10^4$ cell/ml each.

2.2 Trace Nutrient Supplementation

On Day 3 (after zero) of the experiment, upon obtaining satisfactory quantum photosynthetic fitness (0.36 ± 0.02) and cell concentration ($1.75x10^4 \pm 3.2x10^3$ cell/ml), iron was added (as FeCl3.6H2O dissolved in ultrapure Milli-Q® water, syringe-filtered to reach a final concentration of $5x10^{-4}$ g/L) to the four treatment replicates dissolved in ultrapure Milli-Q® water, syringe-filtered to reach a final concentration of $5x10^{-4}$ g/L.

2.3 Physiological, Growth and Biomass Assessments

Throughout the study, a set of assessments was conducted daily to support our transcriptomic data. Physiological (e.g., photosynthetic fitness and growth), morphological (e.g., single cells and colonies size and count), whereas chemical (e.g., particulate organic carbon (POC) and nitrogen (PON), and chlorophyll *a* contents) analyses were carried out on the control and treatment replicates before and after iron supplementation. Here the protocols' principles, procedures and calculation formulas are stated.

2.3.1 Pulse Amplitude Modulation (PAM) Fluorometry

Chlorophyll fluorometry is the measurement of the efficiency of the photosystem in emitting the absorbed light mirroring the efficiency of the photosystem in using the light absorbed in photosynthesis. Modulated fluorometers are modified to specifically measure the fluorescence excited by the device light reducing background noise (Maxwell & Johnson 2000). Minimum (F_o) and maximum (F_m) chlorophyll *a* fluorescence readings (in mV) were obtained and quantum efficiency of photosystem II (PSII) assuming that all reaction centers are open (F_v/F_m) was calculated according to (Maxwell & Johnson 2000) as follows: $Fv/Fm = (Fm - Fo)/Fm$.

Photosynthetic fitness was obtained through fluorometic measurement of chlorophyll *a* using Xenon Pulse Amplitude Modulation photosynthetic yield analyzer

XE-PAM® (Walz, Effeltrich, Germany). XE-PAM® is coupled with a FL-100 control unit operated by Fluorwin Software (version 3.5) (Photon System Instruments, Drasov, Czech Republic). In case of very high cell density, emission filters were used to avoid jeopardizing the sensitivity of the detector.

Cultures were gently mixed by rotation and subsampled (10 ml) into dark glass vials and allowed to stand in the dark at 2 ± 2 °C for 20 minutes for dark adaptation before measurements were undertaken in the dark. Measurement time of the day was kept constant throughout the study.

2.3.2 Epifluorescence Microscopy

Cultures were subsampled (5 ml) for viable cell count in glass scintillation vials, immediately fixed with formaldehyde neutralized with hexamethylenetetramine (2% final concentration) and preserved in the dark at 4 °C. The cells were to be stained with 1 μM 4',6-diamidino-2-phenylindole (DAPI), gently mixed by rotation and stored in the dark at 4 °C prior to microscopic examination. DAPI is a fluorescing stain that binds double stranded-DNA (i.e., nuclear, mitochondrial and chloroplast DNA). The excited (at 360 nm wavelength) DNA-DAPI complex emits a characteristic blue fluorescence at > 390 nm wavelength allowing its visualization using epifluorescence microscopy (Porter & Feig 1980).

Fixed, stained, cold subsamples were gently mixed by rotation to obtain a homogenous distribution of the cells, poured into Utermöhl cell-sedimentation chambers (approx. 2.7-3.14 ml volume and 2.6 cm diameter) and allowed to settle for 24 hours in the dark at 4 °C. Sampling time of day as well as sedimentation chambers used were kept constant throughout the experiment to minimize variability. Viable single cells and colony-forming aggregates in the settled samples were counted using Zeiss Axiovert® 135 inverted microscope (Carl Zeiss AG, Göttingen, Germany) coupled with incident-light fluorescence illuminator at a total magnification of 400x. Images were captured using Zeiss AxioCam® HRc 14-bit color CCD camera coupled to Zeiss Axiovert® 200 inverted fluorescence microscope and processed by Zeiss AxioVision® software (v3.1) at a total magnification of 200x.

Settled viable single (2 μm in diameter) and aggregated (5-10 μm in diameter) cells were counted in 4 stripes; longitudinally, transversely and diagonally across the chambers (Utermöhl 1958). A general formula to obtain the number of phytoplankton cells per ml is: $No.\,of\ cells/ml = \frac{C}{L\ x\ D\ x\ W\ x\ S}$, where *C* is the total cell (single or colony) count, *L* is stripe length, *D* is the stripe depth, *W* is the stripe width and *S* is the number of counted stripes. A conversion factor for the area of the counted stripe at 400x magnification corresponding to the total area and divided by the total volume of the chamber was calculated. First, the area of Utermöhl chamber (*A*) was calculated to be 5.31 cm^2 as follows: $A = \pi r^2$, where *r* is the radius of the chamber. The area of 100 units/stripe at 400x magnification (A_s) was calculated as follows: $As = L$ x W, where the length of the stripe (*L*) is the chamber diameter and the stripe width at 400x magnification (*W*) corresponds to 0.026 cm. Finally, the conversion factor for the volume in total was calculated to be equal to $\frac{A}{As\ x\ V}$, where *V* is the chamber volume, and single and aggregated cell counts were multiplied by the conversion factor and divided by the number of counted stripes to obtain the cell concentrations per ml. To obtain total cell concentrations, the number of single cells in a colony-forming aggregates were estimated to be equal to 3. Specific growth rates (μ) were calculated using the following formula: $\mu\ (per\ day) = (\ln(C2) - \ln(C1))/(t2 - t1)$, where C2 is the solitary cell concentration (cell/ml) at t_2, and C_1 is solitary cell concentration at t_1 (John et al. 2010; Assmy et al. 2007).

2.3.3 Fluorometry

Pigment analysis has been used as a proxy for biomass build-up as well as primary production *in vitro* and *in vivo* either spectrophotometrically or fluorometrically [compared in (Lorenzen & Jeffrey 1980)]. Cultures were examined fluorometrically for the determination of chlorophyll *a* and its corresponding phaeophytin contents according to [(Evans et al. 1987) cited in (Riegger & Robinson 1998)]. Chlorophyll *a* (Chl *a*) is excited at a wavelength of 430 nm emitting light at a wavelength of 665-670 nm (for PSII), while phaeophytin *a* (Phaeo *a*) emits light at a wavelength of 663 nm.

Cultures were subsampled (40-60 ml) to be concentrated on Whatman® GF/C glass-fiber filters (1.2 µm; Omnilab, Bremen, Germany) using vacuum filtration. Filters were preserved in cryovials, immediately frozen in liquid N_2 and stored at -20 °C to avoid degradation. Chlorophyll was extracted by adding 90% acetone to the filters that were afterwards sonicated for 20 seconds in an unltrasonic ice bath with Sonoplus® HD70 (Bandelin Electronic, Berlin, Germany) to disrupt cell and chloroplast membranes to ensure total pigment elution. Homogenized samples were incubated in dark at 4 °C for 2 hours. Samples were centrifuged for 10 minutes at 4500 revolutions per minute (rpm) at 5 °C. Chlorophyll *a* fluorescence was measured in the supernatant using TD-700™ fluorometer (Turner Designs Inc., California, USA) to obtain fluorescence readings in florescence units (fsu) corresponding to total chlorophyll *a* and phaeopigment (F_0). Acidification of the supernatant using 0.1N HCl was done to convert all the chlorophyll *a* to phaeophytin *a*, and their fluorescence readings (F_a) were obtained.

Chlorophyll *a* and phaeophytin *a* concentrations were calculated according to (Lorenzen 1967) [equations adjusted by E. M. Nöthig and C. Lorenzen, unpublished chlorophyll *a* measurement protocol, AWI and explained in (Newton 2002)] as follows:

$$Chl\ a\ \left(\frac{\mu g}{L}\right) = \frac{\frac{F0/Fa\ max}{F0/Fa\ max-1} * (F0-Fa) * Kx * Ve}{Vf}$$ and

$$Phaeo\ a\ \left(\frac{\mu g}{L}\right) = \frac{\frac{F0/Fa\ max}{F0/Fa\ max-1} * ((F0/Fa\ max * Fa)-F0) * Kx * Ve}{Vf}$$, where V_e is the volume extracted (ml), V_f is the volume filtered (ml), $F_0/F_a\ max$ is the ratio of F_0 to F_a of the standard chlorophyll (2.14 fsu) and K_x is the calibration factor of the fluorometer (calculated as 0.7556). Fluorescence reading of blank 90% acetone was subtracted from F_0 and F_a.

2.3.4 Combustion-Gas Chromatography

Correlating with chlorophyll *a* contents, particulate organic carbon (POC) and nitrogen (PON) content determination is another proxy for the biomass of the primary producers. Comparing observed POC:PON ratio to expected POC:PON ratio by Redfield is an indicator for phytoplankton growth under nutrient limitation (Vaillancourt et al. 2003).

In order to determine POC and PON contents in the cultures before and after iron supplementation, elemental analysis was conducted. The elemental analysis is based on combustion-gas chromatography converting all C and N contents into CO_2 and N_2, thus pre-combusted filters are used to avoid added C content. PON is transformed into its equivalent ammonium salts by acidification prior to oxidation by combustion (1000 °C) to their corresponding gaseous products to be measured chromatographically. POC is oxidized to CO_2 while all nitrogen oxides produced through the oxidation step are reduced (600 °C) to N_2.

Cultures were subsampled (50-80 ml) and concentrated on pre-combusted [at 490 °C for 2 hours (Hickel 1984)] Whatman® GF/C glass-fiber filters (1.2 µm; Omnilab, Bremen, Germany) using vacuum filtration to be stored at -20 °C. Filters were oven-dried at 60 °C overnight following addition of 0.1N HCl. Filters were encapsulated into chloroform-washed tin containers and compressed to be analyzed for POC:PON using EA3000 ElementalAnalyzer® (EuroVector, Milan, Italy) (Beszteri et al. 2012). POC and PON contents (µg) in the samples, standard (acetanilide) and blanks were calculated from area readings from the elemental analyzer upon the protocol's equations (C. Lorenzen, elemental analyzer protocol, AWI). POC and PON contents per cell were obtained as follows: $Nutrient\ content\ (\mu g/cell) = \frac{weight\ in\ sample - weight\ in\ blank}{no.of\ cells\ filtered}$.

2.3.5 Student's *t*-Test and Principal Component Analysis (PCA)

All experimental assessment were subjected individually to statistical testing using two-sided unpaired *t*-test at a confidence level of 0.95 and repeated measure analysis of variance (ANOVA) when indicated. Furthermore, principal component analysis (PCA) was conducted to reveal the hidden patterns and elucidate the correlation between the different components of the study. Statistical analyses were conducted using R statistical programming language with default parameters (R Core Team 2013).

2.4 Molecular Assessments

2.4.1 Cell Harvesting and RNA Extraction

In addition to stock culture (T0), treatment cultures were subsampled (100 ± 10 ml per filter) daily and on Day 3 of the experiment before and after iron supplementation. Following concentration on MF-Millipore™ membrane filters (1.2 µm; Merck KGaA, Darmstadt, Germany) using vacuum filtration, the cells were resuspended in 500 µl β-mercaptoethanol/RLT buffer and (Qiagen, Hilden, Germany) preserved in liquid N_2 and stored at -80 °C prior to RNA extraction.

Total RNA was extracted using RNeasy® Plant Mini Kit (cat. nos. 74903 and 74904; Qiagen, Hilden, Germany) according to manufacturer's instructions (Beszteri et al. 2012), modifications and specifications are included in this section. After addition of glass beads, the cells were broken open using a vortex mixer (30 seconds at 50,000 rpm). The homogenized lysates were purified using the following purification steps: (1) lysate transfer to QIAshredder spin column and centrifugation (5 seconds at 13,000 rpm) to remove cell debris, (2) 0.5x volume ethanol addition to the flow through, and (3) transfer to RNAeasy spin columns and DNA digestion by incubation with RW1 buffer (5 minutes), DNase mixture (15 minutes), RW1 buffer (5 minutes), and RPE buffer (5 minutes) for two times. Each step was followed by centrifugation (1 minute at 13,000 rpm). Purified RNA was eluted with 30 µl RNA-free water and preserved at -80 °C.

2.4.2 RNA Quantitative and Qualitative Assays

The concentration of RNA (ng/µl) was estimated using NanoDrop® ND-1000 Spectrophotometer (Peqlab, Erlangen, Germany) according to manufacturer's instructions. RNA purity was estimated by 260/280 and 260/230 absorbance ratios. The qualitative assay of RNA was verified using 2100 Bioanalyzer coupled with 2100 Expert Software (Agilent Technologies Inc., Böblingen, Germany) which is based on microfluidics instead of the traditional sample-intensive gel electrophoresis technique according to manufacturer's instructions. RNA integrity was estimated by RNA integrity numbers (RIN).

2.4.3 RNA Sequencing

Extracted RNA of acceptable quality (260/280 > 1.6 and RIN > 5) from the stock culture and all treatment replicates (Days 2, 3-1, 3-2 and 5 except Day 3-1 replicate 2) was processed by The European Molecular Biology Laboratory (EMBL) Genomic Core Facilities (GeneCore, EMBL Heidelberg, Germany) for complementary DNA (cDNA) library construction and strand-specific paired-end total RNA sequencing using 5x Illumina HiSeq™ sequencer (Illumina Inc., California, USA).

2.5 Computational Analysis

Trinity *de novo* transcriptome assembler (Grabherr et al. 2011; Haas et al. 2013) running on the Data Intensive Academic Grid (DIAG) (http://diagcomputing.org/) was used for RNA-Seq data assembly and partially for downstream analysis. Trinity's workflow using the programs default parameters was followed for reads assembly, transcripts extraction, abundance estimation as well as automated curation of transcripts (i.e., isoforms) functional annotations. Differential gene expression analysis of the genes (i.e., components) was performed using the R/Bioconductor package DESeq (v1.14.0) (Anders & Huber 2010).

2.5.1 *De novo* Transcriptome Assembly, Abundance Estimation and Protein-Coding Transcripts Extraction

In order to increase the depth of the transcripts assembly and decrease the number of isoforms generated by Trinity to a realistic number, assembly of paired-end reads from all replicates before and after iron supplementation altogether was performed using `Trinity.pl` (release 2013-02-25) and assembly statistics were obtained using `TrinityStats.pl` (r2013-11-10).

Abundance of the generated transcripts has been estimated by the Trinity's utility utilizing RNA-Seq by Expectation Maximization (RSEM) (Li & Dewey 2011) `run_RSEM_align_n_estimate.pl` (r2013-02-25) aligning paired-end reads of each replicate of each sample individually with the generated transcripts, while extraction of potential coding transcripts from the entire assembled transcriptome was performed

using TransDecoder's `transcripts_to_best_scoring_ORFs.pl` (r2012-08-15).

2.5.2 Transcripts Functional Analysis and Automated and Manual Curation

Transcripts were analyzed using the tools and databases (summarized in Table 4) using their default parameters unless stated. Similarity search [using BLAST+ suite (Altschul et al. 1990; Camacho et al. 2009)], protein domain search [using HMMER package (Eddy 1998; Eddy 2011)], cellular localization prediction [using SignalP (Petersen et al. 2011), TMHMM (Krogh et al. 2001), WoLF PSORT (Horton et al. 2006)], and ribosomal RNA genes detection [using RNAmmer (Lagesen et al. 2007)] were among the analyses conducted. For the automated curation and cross-association of the different annotations, as well as association with evolutionary genealogy of genes: Non-supervised Orthologous Groups (eggNOG) (v3.0) (Powell et al. 2012) and Gene Ontology (GO) (Ashburner et al. 2000) orthologous gene groups, `Trinotate` (r2013-11-10) was used. The filtration criterion of the reported curated associated annotations was set to be *e*-value ≤ 1e-3.

2.5.3 Differential Gene Expression Analysis and Clustering

Raw fragment counts (number of mapped reads per feature) and trimmed mean of *M*-values (TMM)-normalized gene expression values [fragments per feature Kb per million reads mapped (FPKM)] of the genes generated by `merge_RSEM_frag_counts_single_table.pl` (r 2013-11-10) and `abundance_estimates_to_matrix.pl` (r beta s2014-03-16), respectively, were used for the differential expression analysis at gene-level. Log values of the raw counts were clustered based on correlation using pvclust R library (bootstrap of 100). Raw counts were fed to DESeq for normalization to effective library size, dispersion estimation (specifying method of dispersion as "blind", and sharing mode as "fit-only"), and binomial statistical testing between every two conditions biological replicates.

Sequential filtration steps of the genes were undertaken; firstly genes with difference between maximum and minimum rounded counts ≥10 were used for

downstream analysis. Secondly, genes with negative binomial Exact test Benjamini-Hochberg adjusted *p*-value (padj) for false discovery rate ≤ 0.001 and absolute log fold-change ≥ 5 were considered as biologically and statistically significant differentially expressed across all days of experiments. Lastly, the expression values normalized to effective library size of the genes fulfilled the previous criteria were extracted and subjected to variance stabilizing transformation.

Clustering of differentially expressed genes using variance stabilization transformed counts was done through hierarchical clustering and *k*-means clustering. Heatmap generation of the genes with the 50 highest mean counts using the transformed values was done to visualize the hierarchical clustering. Sample-to-sample Euclidean distances and PCA were done using total transformed expression values specifying the number of genes to be used as the number of those significantly differentially expressed.

Table 3 Constituents of f/2 phytoplankton growth medium modified from (Guillard & Ryther 1962) and their final concentrations

Constituent	**Final concentration (g/L seawater)**
$NaNO_3$	0.075 g/L
$Na_2HPO_4.2H_2O$	0.00625 g/L
Metal mix:	
$ZnSO_4.H_2O$	$1.5x10^{-5}$ g/L
$CuSO_4.5H_2O$	$1x10^{-5}$ g/L
$CoSO_4.7H_2O$	$1.2x10^{-5}$ g/L
$MnSO_4.H_2O$	$2x10^{-4}$ g/L
$FeCl_3.6H_2O$	$5x10^{-4}$ g/L
$Na_2MoO_4.2H_2O$	$6.5x10^{-6}$ g/L
$Na_2EDTA.2H_2O$	$5x10^{-3}$ g/L
Vitamin mix:	
Cyanocobalamin	$1x10^{-6}$ g/L
Biotin	$1x10^{-6}$ g/L
Thiamin	$2x10^{-4}$ g/L

Table 4 Functional annotation tools and databases used for *P. antarctica* assembled transcriptome analysis.

Analysis tool	**Sequences analyzed**	**Database and/or parameter**
blastx (v2.2.25+)	Transcripts	UniProtKB/Swiss-Prot (r2014-01-22) (The UniProt Consortium 2014)
blastp (v2.2.25+)	Extracted ORFs	UniProtKB/Swiss-Prot (r2014-01-22)
hmmscan (v3.0)	Extracted ORFs	Pfam-A (r27.0) (Punta et al. 2012)
signalp (v4.1)	Extracted ORFs	N/A
tmhmm (v2.0c)	Extracted ORFs	N/A
WoLF PSORT (v0.2)	Extracted ORFs	Organism type: "plant"
rnammer (v1.2)	Transcripts	Using Trinotate's `RnammerTranscriptome.pl` to create a transcriptome superscaffold specifying organism type as "eukaryote"

CHAPTER 3: RESULTS AND DISCUSSION

3.1 Physiological, Growth and Biomass Assessments

Throughout the present study, quantum efficiency of PSII, growth rates and cell concentrations, as well as cellular Chl *a*, POC and PON contents have been measured either daily or at defined time-points in control (PA_plusFe) and treatment (PA_minusFe) replicates before and after iron supplementation of the treatment cultures, which took place on Day 3.

3.1.1 *P. antarctica* Recovers Photosynthetic Fitness Following Iron Supplementation

As an assessment of the efficiency of the photosystem in utilizing the light absorbed in photosynthesis, chlorophyll fluorometry was employed to determine the maximum yield of PSII in the replete control and treatment cultures on daily basis (Table 5 and Figure 5B). Quantum photosynthetic fitness of the stock cultures were used as a proxy for negative control (Figure 5A). A significant difference in photosynthetic fitness of PSII (F_v/F_m) between control and treatment cultures (p-value < 0.05) has been observed at Days 1 (p = 4.9e-5), 2 (p = 9.6e-4) and 3 (p = 4.0e-4). F_v/F_m values ranging between 0.2 and 0.36 at the three time-points prior to iron addition indicate that the cultures were under iron limitation stress. Day 3 F_v/F_m values are yet at the border of reported healthy values of the replete cultures (ranging between 0.35 and 0.49) (de Baar et al. 2005; Gaebler-Schwarz 2009; Marchetti et al. 2009; Marchetti et al. 2012; Boyd & Denman 2008). Interestingly, one day upon iron supplementation, no significant difference has been observed in quantum photosynthetic yield between the control and the treatment cultures at Days 4 (p = 0.05) and 5 (p = 0.04) both reaching ~0.5, the fitness ratio that has been previously reported for healthy *P. antarctica* (Gaebler-Schwarz 2009). The relatively low initial F_v/F_m of the treatment cultures indicates the cultures recovery from the iron-limited growth condition of the stock culture.

3.1.2 Iron-Replete *P. antarctica* Shift Towards Colony-Formation

Microscopic examination of treatment cultures at the time-points prior to iron addition revealed that the cultures constituted mostly of solitary cells, an observation that was also supported by the examination of the stock culture (T0) where a small number of small colonies were observed (Figure 7A). On the other hand, after iron addition, from Day 4 and on, minusFe showed a peak in number of colony-forming aggregates (Figure 6B) side by side with the increase in solitary cells (Figure 6A). Furthermore, the colonies in the enriched control exhibited a pattern of increase in size (Figure 7D-F), however, no skin has been observed to surround the colonies (Assmy et al. 2007) perhaps due to long preservation time.

Throughout the experiments, control and treatment cultures exhibited a pattern of exponential increase in total cell concentrations. The increase in cell concentration was more prominent in the enriched control. Test statistics have not revealed significant difference between the two conditions across the days of the experiment. The total abundance of *P. antarctica* increased from $1.75x10^4 \pm 3.2x10^3$ cell/ml in the minusFe to $2.43x10^4 \pm 1.26x10^4$ cell/ml at Day 4 and $1.44x10^5 \pm 2.08x10^4$ cell/ml by the end of the experiment following iron addition comparable to total abundances of the plusFe $3.35x10^4 \pm 1.36\ x10^4$ cell/ml and $2.73x10^5 \pm 2.89\ x10^4$ cell/ml, respectively. The solitary cells concentrations in Days 4, 5 and 8 (Table 5) reached comparable values in both the control and the treatment cultures following iron addition. Likewise, colony-forming aggregates numbers in the treatment cultures followed an exponentially increasing trend after the near constancy trend before iron addition.

Assmy *et al.* calculated the percentage of contribution of the colonial cells prior to iron addition to be 2% (Assmy et al. 2007). Likewise, in the present study, average abundance of colony-forming aggregates increased from 4% in T0 (Day 0) and 5% in Day 1 to 12% from Day 4 until the end of the experiment. The colonial abundance in the minusFe after iron addition was comparable to its abundance in plusFe (11% on Day 2), consistent with the alternative hypothesis that *P. antarctica* tends to shift towards colonization upon iron enrichment, however, the decrease in solitary cells abundance observed during EisenEx was explained by grazing by ciliates (Assmy et al. 2007).

Due to the reported and modeled difference in the growth rate (μ) between the solitary and colonial morphotypes (Peperzak et al. 2000), in the present study growth rate of *P. antarctica* was calculated for the single cells. Overall mean growth rate of *P. antarctica* was calculated as 0.2 day^{-1} in the minusFe cultures and 0.22 day^{-1} in the plusFe cultures. The mean growth rate of the three time-points before iron addition were calculated to be 0.1 day^{-1} in both plusFe and minusFe, while the mean growth rates increased to reach 0.44 and 0.46 day^{-1} in plusFe and minusFe, respectively, in the following three time-points upon iron enrichment. Although previously reported growth rates of *P. antarctica* under variable light intensities ranged from 0.04 to 0.34 day^{-1} (Moisan & Mitchell 1999), iron enrichment effect on growth rate of non-colony-forming and colony-forming cells determined from Chl *a* fluorescence has been reported to increase the growth rate from 0.25 to 0.55 and 0.52 day^{-1}, respectively (Strzepek et al. 2011) complying with the data from the present study.

The disagreeing statistics of pairwise *t*-test and repeated measures ANOVA, insignificant *vs.* significant difference between treatment and control in all time-points and along all counts, respectively, and the large standard deviation (SD) (Table 5) both can be explained with microscopic counting errors as well as the use of constant estimates of the number of cells per colony-forming aggregate for all experiment days and cultures.

3.1.3 Chlorophyll *a* Production Increase upon Iron Enrichment

Chlorophyll *a* is a proxy for photosynthetic fitness as well as biomass and its measurement is usually coupled with and compared to its grazing-driven degradation product, phaeophytin *a*. The lower the Phaeo *a*:Chl *a* ratio, the healthier the phytoplankton community is predicted to be (Bracher et al. 1999). In DMS flux assessments, corrected Chl *a*/Phaeo *a* ratio have been calculated to estimate the daily flux of DMSP (Stefels et al. 2007).

In the present study, Chl *a* contents have been measured at two middle time-points before and after iron addition to the treatment cultures (Day 2 and Day 5) showing a significant ($p < 0.05$) difference in Phaeo *a*:Chl *a* between the control and treatment at

Day 2 indicating limited photosynthetic efficiency, and due to iron addition, insignificant difference was observed at Day 5 between plusFe and minusFe (p = 0.38). Chl *a* contents of Day 2 were 3.5 ± 0.7 and 3.6 ± 1 µg/L in minusFe and plusFe, respectively, while a sharp increase at Day 5 that constituted for 6.9 ± 0.2 and 13.3 ± 2.6 µg/L in minusFe and plusFe, respectively, has been observed (Table 5 and Figure 8). The increase is suggested to be due to the increase in Chl *a* production in response to iron addition (Bertrand et al. 2011) and the increase in cell density due to the increase in colonial form abundance (Smith et al. 1998). The close figures of Chl *a* concentrations at Day 2 between plusFe and minusFe can be investigated further using HPLC to measure the 19'-hexanoyloxyfucoxanthin and fucoxanthin for resolving the complete pigments profile under iron-limited and –replete conditions (DiTullio et al. 2007).

3.1.4 Cellular Nitrogen Contents Increase Following Iron Enrichment

Reported C:N ratios of *Phaeocystis* under non-growth-limiting conditions exceed that of Redfield's (160:16). Even more specific for *P. antarctica*, it has been reported that healthy colonies have a characteristic C:N ratio of 6, slightly deviating from Redfield ratio due to the high C content accounting for the colonies' polysaccharide matrix. Iron-limited *P. antarctica*, however, has been observed to have different C:N ratios ranging from 5 to 8 (Schoemann et al. 2005).

Interestingly, minusFe *P. antarctica* C:N observed to decline from 7.6 ± 2.8 (6.23 excluding an outlying value) at Day 2 to 5.57 ± 0.2 at Day 5, comparable to those ratios reported for *P. antarctica* cells under iron limitation and to colonies at the exponential phase, respectively (Schoemann et al. 2005). On the other hand, plusFe C:N ratios were calculated to be 3.3 ± 0.1 at Day 2 and 4.7 ± 0.2 at Day 5. Such low C:N at Day 2 in the replete control was observed to be due to a surprisingly high N content (6.30×10^{-6} ± 1.58×10^{-6}) (Table 5) which fact might be explained by a rapid N assimilation upon iron enrichment that is not compensated for with a matching increase in C contents due to late colony formation. Pairwise *t*-test revealed a significant difference between C:N in control and treatment at Day 5 (p = 0.0015) and the difference at Day 2 can be considered significant as well (p = 0.055). Furthermore, the difference in PON content between

plusFe and minusFe at Day 2 is significant ($p = 0.04$) (Figure 9A). In summary, C:N, prior to iron addition, exceeded Redfield ratio in the treatment cultures indicating iron stress conditions limiting growth, while at Day 5 both the control and the treatment reached a ratio typical for exponentially growing *P. antarctica* colonies (Schoemann et al. 2005) (Figure 9B).

Chl *a*:C ratio (light saturation index or photoacclimation indicator) has been considered as a more realistic estimate of algal biomass than Chl *a* solely due to the many sources of variability in Chl *a* such as photoadaptation and species composition (Dugdale & Wilkerson 1991). Iron also increases the pigment content and normalization by C content is required for modeling phytoplankton biomass and growth rates (Wang et al. 2009). Chl *a*:C (mg:g) of healthy *Phaeocystis* has been reported to be 30 in solitary cells and 18 in colonial cells, a decrease that also can be explained by the carbon contribution of the mucus matrix. However, Chl *a*:C (mg:g) of iron-limited *P. antarctica* was reported to be as low as 10 [table 6 in (Schoemann et al. 2005)]. In the present study, Chl *a* contents per cell:C per cell (mg:g) has been observed to reach 9.6 ± 3 in minusFe at Day 2, a value typical for iron deficient cultures (Schoemann et al. 2005). Chl *a*:C ratio declined at Day 5 to 7.4 ± 0.4, a value comparable to that of plusFe 6.7 ± 0.9 at both Day 2 and Day 5, however, such lower values are reported to be associated with N deficiency (Schoemann et al. 2005; Wang et al. 2009) and can be explained by measurement errors. Likewise, elevation in C:Chl *a* (g:g) indicates for photoadaptation (Smith et al. 1998), and in the present study, C:Chl *a* (g:g) increased in minusFe at Day 5 from 111.5 (± 34) to 135.06 (± 7.5), while it was constant (~150) in the replete control throughout the experiment.

3.1.5 Iron State is Possibly the Principal Component that Affects Photosynthetic Fitness

Principal component analysis identified nine principal components, an equivalent number to that of the study's observations. Three of these components constituted for the largest proportion of variance (of cumulative proportion of variance = 90%), PC1 (66.7%), PC2 (15.4%), and PC3 (8.8%), and were used to visualize the variable coefficients. Being sensitive to relative data scaling, normalization of the data (per liter

cell volume, for instance) and removal of iron state (0 or 1) might help resolving the pattern.

Plotting the variables coefficients against PC1 *vs.* PC2 clustered cell counts with pigments concentrations and time, while iron state correlated with quantum photosynthetic fitness (PAM) and POC (Figure 10A). This suggests iron as the governing component of F_v/F_m and C content, however, it affects colonial counts and PON to lower extents. Concordant with Figure 10A, PC1 *vs.* PC3 separated cell counts, however, it correlated iron state with solitary cell counts, time and quantum photosynthetic fitness, and confirmed the correlation between colony-forming aggregates counts with pigments and nutrients contents (Figure 10C). Finally, scatterplot of PC2 *vs.* PC3 separated cell counts correlating solitary cell counts with time, and iron state with photosynthetic fitness, while clustering colony-forming aggregates counts with pigments, while nutrients contents had no correlation with the other observations (Figure 10B).

3.2 *P. antarctica* Transcriptome

Sixteen cDNA libraries have been sequenced using Illumina RNA-Seq technology (Table 6) to both characterize the genome of *P. antarctica* and to capture the changes in gene expression patterns of the iron-limited *P. antarctica* towards iron enrichment in a time-series fashion. The transcriptomes of the stock culture (T0) as well as minusFe replicates of Days 2, 3 prior to and following iron enrichment and Day 5 were sequenced. Due to the unavailability of *Phaeocystis* genome sequence, *de novo* assembly and functional annotation of the first reported transcriptome of Ross Sea *P. antarctica* have been conducted using Trinity pipeline (Grabherr et al. 2011; Haas et al. 2013).

3.2.1 Transcriptome Statistics

A total of 389,846,414 paired-end reads (19,882,167,114 b) have been sequenced of which 136,501,307 b have been assembled into 162,436 transcripts (i.e., isoforms) of 88,630 genes (i.e., components) with a contig N50 of 1190 b (Table 7) and estimated GC content of 63.36%. The transcriptome GC content might indicate for a genome GC content comparable to that of *E. huxleyi* (Read et al. 2013) and higher than flow

cytometry estimates (Vaulot et al. 1994), however, accurate identification of the genome AT-rich regions is required to test this hypothesis.

The frequency distribution of isoforms mapping (Figure 11) showed that the largest fraction of genes (77.6%) had only one isoform mapped to each, indicating assembly accuracy (Haas et al. 2013). The maximum number of isoforms mapped to a single gene was 241 isoforms mapped to comp95704_c11. comp95704_c11 has showed a constitutive non-differential expression pattern throughout the experiment regardless of the growth conditions and its functional annotation suggests that it codes for an ubiquitous zinc-dependent quinone oxidoreductase (blastp hit sp|A7RK30; 2e-46) and/or a zinc-dependent RNA-directed DNA polymerase (blastp hit sp|P08548; 3e-05). Given that all isoforms have similar expression patter, paralogs might be a more accurate term to describe comp95704_c11 isoforms.

The high estimate of isoforms and genes in *P. antarctica*, compared to 30,569 predicted number of genes in *E. huxleyi* (Smith et al. 2014), can be explained. The high overlapping threshold used in Trinity in both identifying genes (Chrysalis group overlapped Inchworm contigs based on *k*-1-mer) and resolving recent gene duplication events and alternative splicing events (Butterfly) (Anon n.d.; Grabherr et al. 2011; Haas et al. 2013) are possible reasons. The "highly similar" identified genes can further be grouped as recent duplication events (i.e., paralogs), while many of the isoforms associated with these genes can be considered as polymorphic sites between the population individuals, especially in diploid cells (Grabherr et al. 2011), given the long time between the isolation of the test clone and the time of conducting the study. Being a paired-end library of an organism of a compact genome, the assembly would largely benefit from the use of `--jaccard_clip` parameter to better resolve fused transcripts (Haas et al. 2013). Furthermore, the use of PasaFly and CuffFly algorithms that have been implemented in the latest versions of Trinity would certainly reduce the reported estimate (Haas n.d.).

3.2.2 Transcripts Functional Annotation

Open reading frames (113,563 ORFs) were extracted and translated from the identified transcripts by Trinity's TransDecoder based one predicted exons scoring of GeneID's coding DNA Markov Model (Parra et al. 2000). Annotations of transcripts associated with blastp and blastx hits of *e*-value < 1e-3 were used, a cutoff that allows for gene discovery from relatively distant homologs. Because a novel transcriptome is investigated, the use of non-redundant protein database for similarity search was omitted and replaced with the use of UniProt and Pfam databases for functional annotation through sequence similarity and domain search, respectively.

At gene-level, out of the total 88,630 genes/components identified, 2,923 (33%) were similar to UniProt entries, while 64,822 (73.1%) were not found similar to UniProt proteins below the specified cutoff. In regard to functional characterization and category assignment, 25,836 (29.2%) genes were assigned to GO categories, 17,729 (20%; 2,932 unique groups) were assigned to eggNOG orthologous groups, while 23,809 (26.9%) were assigned to Pfam families.

Being the most comprehensive (Powell et al. 2012) and of the most specific annotations, *P. antarctica* functional genomics based on eggNOG gene clusters is reported in detail at transcript-level. *P. antarctica* transcriptome has been assigned to 2,575 unique eggNOG orthologous groups, 44 of which are of transcript size ≥ 100 (Table 8). The largest gene family (COG0515) constituted of 1864 transcripts, of which 42 transcripts (31 genes) were significantly differentially expressed throughout the experiment. The genes encode for ubiquitous Calcium-dependent protein kinases involved in stress response, nitrogen utilization and myosin filament assembly.

As a proxy for gene duplication events in the favor for increasing protein synthesis capacity mechanisms in chromalveolates, the abundance of transcripts encoding for ribosomal proteins in *P. antarctica* transcriptome have been assessed. A large cluster of families of size of 634 transcripts (538 components) has been annotated by UniProt (and GO (GO:0006412); 473 isoforms) as a potential ribosomal proteins gene cluster, the largest member of which was comp93790_c4 family (30 isoforms) encoding for the large subunit ribosomal protein RL32 (blastx hit Q962T1; 1e-46). However, in dinoflagellates, RL27A family consisted of 5-isoform as well as –component members compared to

Alexandrium tamarense (74 members) (Moustafa et al. 2010; Moustafa 2009) which might indicate a lineage-specific duplication pattern of ribosomal proteins in chromalveolates. The lack of dinoflagellate genomic and RNA-Seq information on total ribosomal protein genes abundance limited the use of the proposed proxy.

Interestingly, 43 transcripts (33 components) have been identified as expressed potential plastid-specific ribosomal protein genes of the recently reported genes in *P. antarctica* (70) and *P. globosa* (66 plastid-specific and 4 mitochondrion-specific) organelles genomes (Smith et al. 2014). Variations in numbers of plastid-specific ribosomal protein genes that have been reported in diatoms (44) (Oudot-Le Secq et al. 2007) and *E. huxleyi* (33) (Sánchez Puerta et al. 2005) can be attributed to plastid genome size. For a higher resolution examination, phylogenetic analysis of the compartment-specific (typically plastid) ribosomal proteins would be undertaken across chromalveolates (Li et al. 2006; Oudot-Le Secq et al. 2007) and to other plastid-containing yet plastid-genome-lacking green algal groups (Smith & Lee 2014).

Metabolic potential of *P. antarctica* under iron –limited and –replete conditions was mapped based on eggNOG orthologous groups to which the transcriptome isoforms were assigned and visualized by iPath (v2.0) (Yamada et al. 2011). Out of 2,575 unique eggNOG orthologous groups, 1,289 groups were mapped to metabolic pathways (Figure 12), while 572 and 378 groups were mapped to regulatory and secondary metabolites biosynthesis pathways, respectively.

Furthermore, out of 114,420 *P. antarctica* GO terms (6,720 of which are unique), 14,230 were summarized and mapped to ancestral terms using CateGOrizer (Zhi-Liang et al. 2008). Figure 13 shows the percentage of GO terms (>1%) mapped to parent categories. The vast majority of the category to which the transcripts were assigned was metabolism (1,462; 10.27%) followed by catalytic activity (1,186; 8.33%). Metabolic process included carbohydrate (121; 0.85%), nucleic acid (459; 3.23%), lipid (161; 1.13%), protein (330; 2.32%) amino acid (1) and tricarboxylic acid cycle (2). 23 (0.16%; 32 putative genes) terms were mapped to plastid compartment and 12 terms to thylakoid (0.08%; 95 putative genes). Within the transcriptome, 126 genes were predicted to be

involved in catalytic activity. The aforementioned estimates are not accurate because of the GO terms lost in the mapping.

Of special interest was DMSP pathway (Summers et al. 1998). Two GO terms involved in methyltransfer were mapped, furthermore, within the transcriptome, 21 genes were predicted to comprise methyltransferase activity and 13 of which putatively encode for S-adenosyl-L-methionine-dependent methyltransferase. Three putative 2-oxoglutarate 5-aminotransferase genes (5e-138 to 2e-05) were present as well.

3.2.3 Nuclear-Encoded Plastid-Targeted Protein-Encoding Genes

Out of total 113,563 translated ORFs, WoLF PSORT (concordant with SignalP and TMHMM results) predicted nuclear-encoded genes encoding for 2,456 plastid-targeted and 8 mitochondrial-targeted proteins [based on maximum weight (= 14)]. The number of the identified plastid-targeted proteins is lower than that reported in the two sequenced diatoms (3,696 and 3,468 in *Thalassiosira pseudonana* and *Phaeodactylum tricornutum*, respectively) (Moustafa et al. 2009). 971 ORFs (967 transcript isoforms; 731 components) were predicted to be of known function based on similarity to UniProt proteins. The taxonomic distribution of the predicted nuclear-encoded plastid-targeted ORFs revealed that the majority of the UniProt hits were of green algal origin (Streptophyta 8%) followed by Metazoa (*Phaeocystis* grazers 4%) (Figure 14). Bacterial [cyanobacterial (36) and chlamydial (3) (Moustafa et al. 2008)], algal (3), excavate (1) and chromalveolate (10) contributions were smaller in the *P. antarctica* potential nuclear-encoded plastid-targeted ORFs. The biased taxonomic representation in UniProt is the most likely cause of the observed high mammalian yet low algal and chromalveolate distribution. In addition, in order to elucidate the potential horizontal gene transfer events from *Phaeocystis* to its metazoan predators (e.g., salps), a further phylogenetic analysis of a more comprehensive similarity search excluding chromalveolates is needed. It is worth mentioning that unassigned ORFs (156; 6%) and those assigned to proteins of unknown taxonomic classification (19; 1%) or assigned to proteins of mammalian origin (201; 8%) were excluded (Figure 14).

3.2.4 Non-Coding RNA: rRNA and tRNA

RNAmmer identified 4 genes (0.0045%; 6 transcripts) coding for five 28S rRNA, three 18S rRNA and one 8S RNA genes (Table 9A). The observed fusion of 28S rRNA and 18S rRNA genes would indicate for fusion incidences in transcripts/isoforms generation due to the genome compactness (3.2.1). Functional annotation, on the other hand, revealed 210 transcripts potentially encoding for 210 genes involved in tRNA aminoacylation and 125 genes belong to tRNA aminotransferases. Because they gave a higher resolution, the occurrence of the genes predicted to be involved in tRNA aminoacylation has been assessed as a proxy for *P. antarctica* codon preference (Table 9B). The tRNA gene frequencies might be, in particular, concordant with what is known about *Phaeocystis* physiology.

The highest occurrences were for the sulfur-containing amino acid cysteine (54; 25.6%) followed by the essential amino acid threonine (18; 8.5%). The reported occurrence of cysteine (25.6%) *vs.* the occurrence of DMSP precursor methionine (6%) (Gage et al. 1997; Summers et al. 1998) raises amino acid synthesis questions regarding the potential of *Phaeocystis* to convert cysteine into methionine. In addition, the reported occurrences questions the ability of *Phaeocystis* to synthesize threonine through the glycine-serine-threonine pathway supported by functional annotation that assigned 20 transcripts to COG0111 and/or KOG0068 the D-3-phosphoglycerate dehydrogenases group involved in threonine biosynthesis.

3.3 Iron Enrichment is Coupled with Significant Metabolic and Floristic Shifts in *P. antarctica*

3.3.1 Hierarchical and *K*-means Clustering

Hierarchical clustering of raw counts per component across replicates in all conditions (i.e., time-points before and after iron enrichment) revealed that expression pattern of the replicates is correlated with significance in a time-dependent manner (Figure 15). Furthermore, pair-wise statistical comparisons of each days replicates blind dispersion estimates were conducted and the statistically (adjusted p-value ≤ 0.001) and biologically (absolute log fold-change ≥ 5) significantly differentially expressed genes were obtained (Figure 16).

Each comparison yielded a number of differentially expressed (DE) genes (Figure 16), and overall 2,367 genes (2.67% of total assembled components) were considered DE genes across all conditions. The identified DE genes are potentially transcribed into 4,209 unique isomers comprising 4,489 possible annotations of 3,659 identified ORFs. Only 1,323 isoforms (31.4% of DE genes) were assigned to GO orthologous groups (902 groups).

Cluster correlation of the expression of the most highly expressed 50 DE genes expressed across all experiment days (Figure 17) revealed a distinct pattern at Day 3-1 expression. A sharp decline in expression in general at Day 3-1 can be explained as a response towards severe iron limitation. Acclimation of cultures in fresh sea water containing traces of iron would explain the long-term stress response at Day 3-1 rather than Day 2 and would also explain the low RNA content and quality that led to the exclusion of one replicate from sequencing suggesting severe iron stress. To further investigate the expression pattern across all days, *k*-means clustering was done and 4 clusters were obtained (Figure 18). Cluster 1 contained 432 genes/components and cluster 2 contained 1087, while clusters 3 and 4 comprised 664 and 188 genes, respectively. The clustered genes biological contribution is reported (Figure 19) for future cluster-specific analysis. Here the top 50 differentially expressed genes and pair-wise comparison between Days 2 and 5 are discussed in details.

The overall expression pattern suggests that *P. antarctica* is indeed adapted to iron limitation and can maintain its cellular processes with minimal amounts of iron, however, long-term or severe state of iron limitation might have impaired *P. antarctica*'s cellular functions. *P. antarctica* –similar to the previously reported behavior of subarctic Pacific haptophytes (Marchetti et al. 2012)—regained full functionality shortly (Assmy et al. 2007) after iron addition.

Hierarchical clustering, similar to the *k*-means clustering, was based on correlation and subdivided the expression pattern of the top DE genes divided them into 4 clusters (Figure 17). Cluster 1 shows high expression levels at Days 2, 3-2 and 5 (e.g., tricarboxylic acid (TCA) cycle enzymes). Cluster 2 shows high expression levels only at Days 0 and 2 and continued decrease after iron enrichment (e.g., mucins). Cluster 3 is

overexpressed at Day 3-1 and its expression started to decline after iron enrichment (e.g., Calvin cycle), while cluster 4 shows a near constancy expression during all experiment days (e.g., cold shock protein and cytochrome c peroxidase).

DNA replication which was overexpressed at Days 0 and 2 (Ribonucleoside-diphosphate reductase and Histone H1), significantly dropped at Day 3-1 (prior to iron addition) yet showed a ~5-fold increase (Ribonucleoside-diphosphate reductase) after iron enrichment at Day 3-2. DNA replication was significantly restored at Day 5 to levels comparable to Days 0 and 2. Starvation response genes (e.g., the initial fatty acid metabolism enzyme medium-chain acyl-CoA dehydrogenase) showed a 4.5- and 5-fold increase at Days 3-2 and 5 (significantly) after iron addition after a 4-fold decline from Day 2 to Day 3-1. Acyl-CoA dehydrogenase has been reported to be overexpressed in diatoms following the onset of nitrogen starvation to catalyze amino acid catabolism (Hockin et al. 2012). Likewise, TCA enzymes (a probable 2-oxoglutarate dehydrogenase, *e*-value = 0) showed significant increases at Day 3-2 and Day 5 relative to the significant decline from Day 2 to Day 3-1. However, an overall insignificant slight decrease in TCA enzymes from Day 2 to Day 5 (0.2-fold) was observed and might indicate that iron addition effect on TCA in *P. antarctica* is minimal.

The iron-sulfur cluster containing aconitate hydratase 2 showed a constantly declining pattern until its abrupt significant 6.5-fold increase instantly after iron addition. Likewise, a similar expression pattern was observed by carotenoid biosynthesis [probably the all-*trans* fucoxanthin carotenoid (Dambek et al. 2012)] and regulation of DNA replication enzymes yet it maintained its significant 5-fold and ~6-fold increases, respectively, at Day 5 or increased at Day 5 as DNA polymerase after a significant 5-fold down-regulation at Day 3-1.

An excellent candidate marker for iron limitation and enrichment response assessment in *P. antarctica* is the transcription regulator, light-repressed protein A, which was significantly differentially expressed day by day. Light-repressed protein A is known to be underrepresented under iron limitation (Nodop et al. 2008). Here it showed a 5-fold and 6-fold drop in expression at Day 0 and Day 3-1, respectively, only to show an instant

7-fold increase after iron addition at Day 3-2 and the increase was maintained at Day 5 (1.5-fold insignificantly).

Moving to the gene clade overexpressed at Day 3-1, a stress response protein (plastid pyrophosphatase) has showed a 6.3-fold increase at Day 3-1 and continually decreased after iron enrichment to reach ~1-fold. Pyrophosphates (PPi) have been proved to enhance the release of Fe(III) bound to transferrin, increase iron availability, however, in vertebrates [reviewed in (Heinonen 2001)] and, if present, transferring it to desferrioxamine (Pollack et al. 1977). It can be suggested that the up-regulation of pyrophosphatases was to drive the release of iron from an iron storage protein in *P. antarctica* to increase its availability in plastid similar to its role in endosomes (Heinonen 2001).

3.3.2 Signaling, Oxidative Stress and Electron Transport Gene Expression Reflects Cell Iron State

Here, for consistency with the pigments and POC and PON contents analysis, the comparison between the two middle points, Day 2 and Day 5, is selected to be presented. As mentioned, Day 2 represents a state of low iron rather than severe iron starvation (Day 3-1). Between Day 2 and Day 5, 160 components (1.8% of the transcriptome) have been differentially expressed, 52 (32.5% of the DE genes) of which are annotated by similarity to UniProt entries. 112 components were found to be up-regulated at Day 2, while 48 were up-regulated at Day 5, given the minimization of metabolic and DNA replication processes in phytoplankton under iron starvation.

Generally, genes encoding for helicases, histone H4, and proteins involved in nucleic acid phosphodiester bond hydrolysis, nucleophagy and protein autophosphorylation, and cell signaling (calcium-dependent proteins, possibly response to nitrogen compounds, and acetylcholine receptor subunits) were 6-8.4-fold up-regulated at Day 2. Up-regulation of calcium-driven signaling- (Allen et al. 2008) and apoptosis-related genes (including autophagy of nuclei) (Thamatrakoln et al. 2012) were observed under iron limitation in diatoms.

In addition, genes involved in oxidative stress were 5-fold overexpressed at Day 2 (e.g., glutathione peroxidase and a possible 2-oxoglutarate Fe(II)-dependent oxygenase). Glutathione peroxidase can be suggested as a defense against the increase reactive oxygen species activity under iron limitation replacing iron-dependent peroxidases (Allen et al. 2008). In diatoms, iron limitation was coupled with an increase in cellular glutamate which formation 2-oxoglutarate Fe(II)-dependent oxygenase catalyzes (Allen et al. 2008). An iron-stress response protein was found to be highly overexpressed under iron limitation (PF07692.6; 3.2e-06; 8.4-fold) resembling that reported in green algae (Rubinelli et al. 2002).

Upon iron enrichment, a potential red algal vanadium-dependent bromoperoxidase has been found to be highly up-regulated (1e-38; 7.2). Bromoform has been detected in the Atlantic and Arctic Oceans yet not in the SO (Wever et al. 1993) produced by the red algal genus *Corallina* (Itoh & Shinya 1994). Interestingly, it has been reported that vanadium-dependent bromoperoxidase functions, coupled with superoxide dismutase, as cellular defense reactive oxygen species. Furthermore, an unexplained iron content in the red algal bromoperoxidase has been reported (Ohsawa et al. 2001). Here it is argued that *P. antarctica* might possess an iron-dependent bromoperoxidase-like peroxidase inherited from an ancient red algal ancestor. The evolved peroxidase would function as an iron storage/scavenging metabolite and/or (Marchetti et al. 2012; Ohsawa et al. 2001) a peroxidase expressed to defend against the increase in reactive oxygen species in the iron limitation period (Allen et al. 2008).

Aerobic respiration genes [e.g., pyruvate carrier 3 (1e-08; 5.5) and cytochrome c oxidase subunits 1 (0; 6.9), 2 (2e-82; 5) and 3 (2e-110; 6.7)] were found to be up-regulated, at Day 5, typical for a shift towards pyruvate metabolism and iron-rich aerobic respiration enzymes following iron addition (Marchetti et al. 2012; Allen et al. 2008).

3.3.3 Structural C Reallocation under Iron Limitation and Increased N Biosynthesis Following Enrichment

Genes involved in polysaccharide biosynthesis, mucin-5B (blastx *e*-value = 0.0005 and 0.0004; 9.3- and 6.5-fold) and mucin-2 (5e-07; 7.6) have been observed to be

highly overexpressed by three genes at Day 2. These results were surprisingly non-concordant with the microscopic observations and the reported evidence of colony-formation following iron enrichment. However, mucin indeed was found to be up-regulated in iron-limited diatoms as part of iron-limitation-driven structural C reallocation (Allen et al. 2008). In addition, a chitinase-like protein was overexpressed under iron limitation (0.001; 5). *P. antarctica* haploid cells star-forming filaments are long-known to be formed of chitin (Chretiennot-Dinet et al. 1997; Rousseau et al. 2007), and the transcriptomic data is concordant with the reports from iron-limited diatoms (Durkin et al. 2012).

In regard to N metabolism, a potential marker of iron and silica starvation in diatoms has been hypothesized which is polyamine biosynthesis (Hamana & Matsuzaki 1982; Marchetti et al. 2012; Durkin et al. 2012). Under iron limitation, a significant up-regulation of potential genes involved in biogenic amine biosynthesis (sperimidine synthase; 2e-06; 8) was observed. The up-regulation of polyamine biosynthesis suggests iron-limitation-driven structural remodeling in *P. antarctica* that indicates for N starvation. In the present study, N contents increased at Day 5, however, transcriptomics of N metabolism and utilization in *P. antarctica* under changing iron availability requires further investigation.

Fructose-1,6-bisphosphatase (9e-22 and 7e-60; 6 and 5.7) were overrepresented at Day 2. Furthermore, the key carbon fixation enzymes in Calvin cycle, fructose-bisphosphate aldolase (4e-116; 8.6) coupled with glyceraldehyde-3-phosphate dehydrogenase (8e-116; 6.6) and the metalloenzyme ribulose-phosphate 3-epimerase (2e-87; 6.9), were overexpressed at Day 5. An opposite behavior has been observed (Allen et al. 2008; Allen et al. 2012) in diatoms which overexpress Calvin cycle genes under iron limitation. Various bisphosphatases have been members of cluster 1 and 2, while aldolases were members of also cluster 1 and 4 (Figure 18and Figure 19). However, Ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) gene was not among the differentially expressed genes. The transcriptomic data suggests that *P. antarctica* depends on Calvin cycle in carbon fixation under both iron –limited and –replete conditions. Another difference with diatoms, *P. antarctica* has been reported to export C

utilizing the available NO_3 reserves at faster rates than that of diatoms (Arrigo 1999), and the transcriptomic data might be concordant with the latter observations.

More interestingly, two putative carbonic anhydrase isoforms (9e-1 and 6e-22) were found to be severely down-regulated from Day 2 to Day 3-1 (9.8-fold) and sharply up-regulated instantly after iron supplementation (Day 3-2; 9.5-fold and Day 5; 8.6-fold). The latter finding supports the suggested shift from complete-dependence on Calvin cycle under iron limitation to beta-carboxylation under enrichment (Smetacek et al. 1997). This is valid due to the reported activity of extracellular carbonic anhydrases in *Phaeocystis* (Schoemann et al. 2005).

Furthermore in mitochondria, a putative gene encoding for a mitochondrial iron ion transport, mitoferrin-1, was overrepresented under iron limitation (1e-20; 5.6) perhaps to ensure iron supply to mitochondria. Under iron enrichment, a putative gene encoding for an iron-binding prolyl 4-hydroxylase subunit alpha (0.0002; 5.5) involved in collagen synthesis was up-regulated. Its low expression under iron limitation can be explained by low proline supply due to the impaired amino acid synthesis (Allen et al. 2008).

3.3.4 Photosynthesis and Photopigments

In consistency with the previously observed floristic shift and down-regulation of the iron-economic PSII expression (Allen et al. 2008) following iron addition (Gall et al. 2001), transcriptomic data revealed that putative gene encoding for the more photosynthetically efficient pigment *Isochrysis*-like fucoxanthin (2e-30; 7.3) was severely up-regulated in the iron-limited samples. However, another component that has been overrepresented under iron enrichment has predicted to be a *Phaeodactylum*-like fucoxanthin (0.001; 5) with a lower *e*-value. These results might indicate for an alternative splicing event in the transcription of the genes encoding for fucoxanthin to express two isoforms (currently components), a 313-amino-acid-long form expressed under iron limitation and a shorter 123-amino acid-long isoform expressed under enrichment. However, pair-wise alignment using blastp has not shown significant similarity. Nevertheless, the results significantly show that fucoxanthin is overexpressed

following iron addition (van Leeuwe & Stefels 1998) and a fucoxanthin-like pigment has been overexpressed under iron limitation. Furthermore, the photoreceptor phototropin-2 (4e-09; 7.3) was highly up-regulated in the iron-limited samples, while a under iron enrichment, a PSI proton gradient regulation 5 protein (2e-10; 7.8) was found up-regulated.

Typically, upon iron enrichment, the iron-rich photosystem I reaction center subunit XI, was found to be overexpressed (7e-44; 5) coupled with a putative ferredoxin-2 (1e-35; 5) and/or ferredoxin (4e-36; 5), and a potential flavanone 4-reductase (3e-30; 5.7) were overexpressed. Nevertheless, flavodoxin was found to be differentially expressed showing ~6-7-fold increase at Day 0 and Day 2 *vs.* a sharp decrease at Day 3-1 and a slight elevation at Day 3-2. It was not differentially expressed at Day 5. These results suggest that *P. antarctica* expresses the iron-economic flavodoxin under iron limitation while upon enrichment it shifts to ferredoxin that is expressed in parallel with flavodoxin. Elevation in ferredoxin levels at Day 3-1 can be explained by the ability of *P. antarctica* to scavenge iron from DFB.

3.3.5 Ferric Reductase Expression under Iron Limitation and Enrichment

A putative ferric reduction oxidase (5e-18) gene showed significant decrease in expression between Day 2 *vs.* Day 3-2 (8.8) and Day 5 (5.3) supporting that *P. antarctica* employs a Fe(III) reductive uptake mechanism that is up-regulated under iron limitation (Strzepek et al. 2011). It is worth mentioning that none of DMSP biosynthesis enzymes (Summers et al. 1998) has been considered differentially expressed in *P. antarctica* at Day2 and Day5 nor across other Days of the experiment.

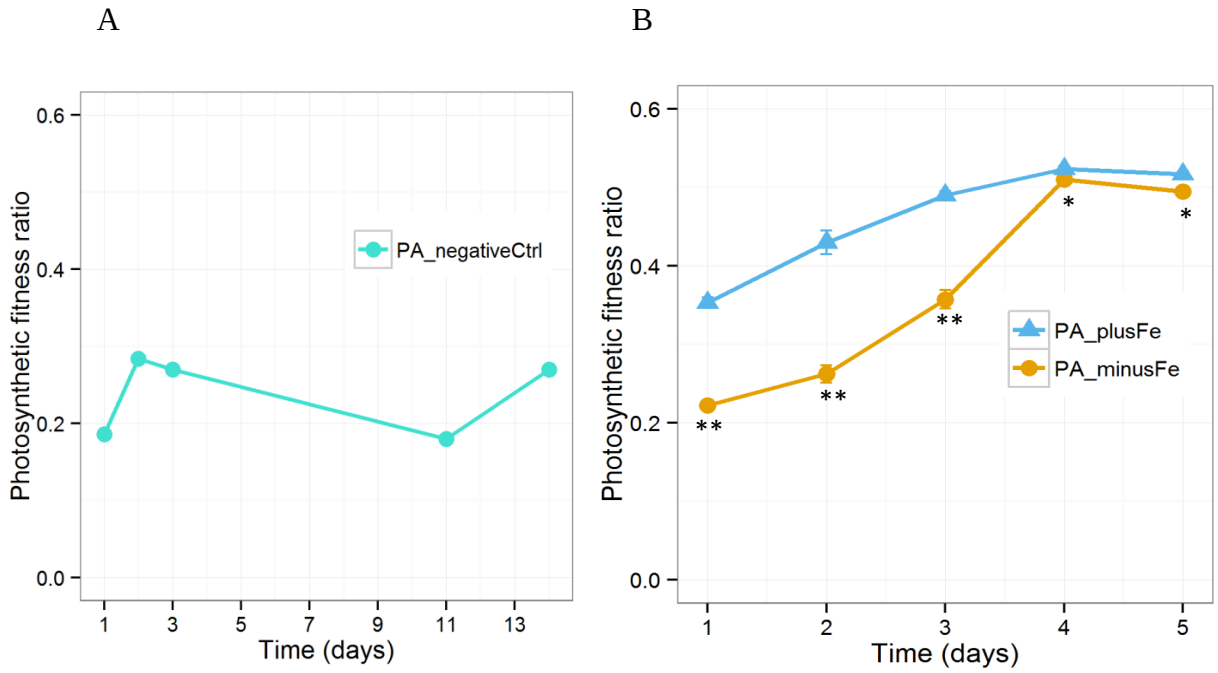


Figure 5 Quantum photosynthetic fitness of PSII (*Fv/Fm*) over time in control and treatment cultures (mean ± SE).

(A) Negative control (stock culture) photosynthetic ratios. (B) Replete control (n = 3) and treatment (n = 4) cultures photosynthetic ratios (mean ± SE). At Day 3, measurements were taken before iron supplementation. **significant difference; *no significant difference following iron supplementation tested by unpaired *t*-test (*p*-value < 0.05).

A B

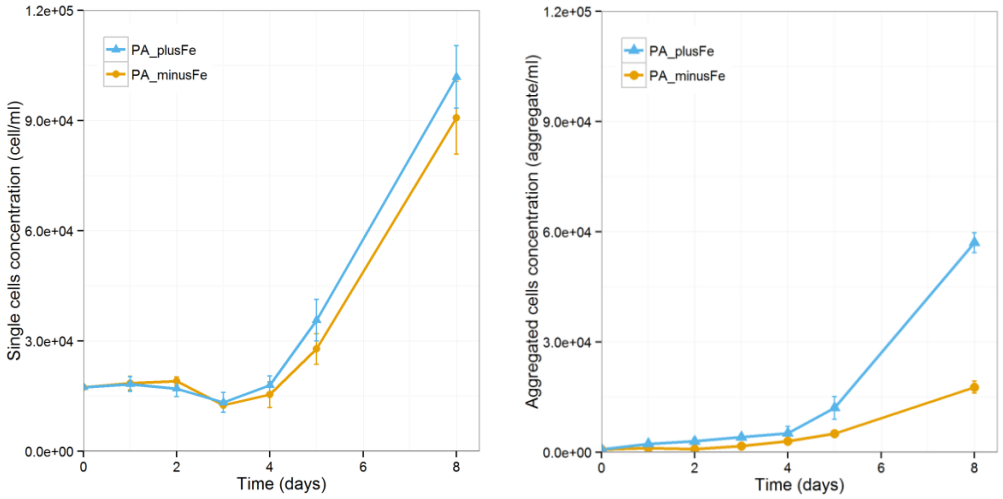


Figure 6 *P. antarctica* cell concentrations in control (n = 3) and treatment (n = 4) cultures (mean ± SE).

(A) Single cells concentrations. (B) Colony-forming aggregates concentrations. Note: At Day 3, counts were taken before iron supplementation.

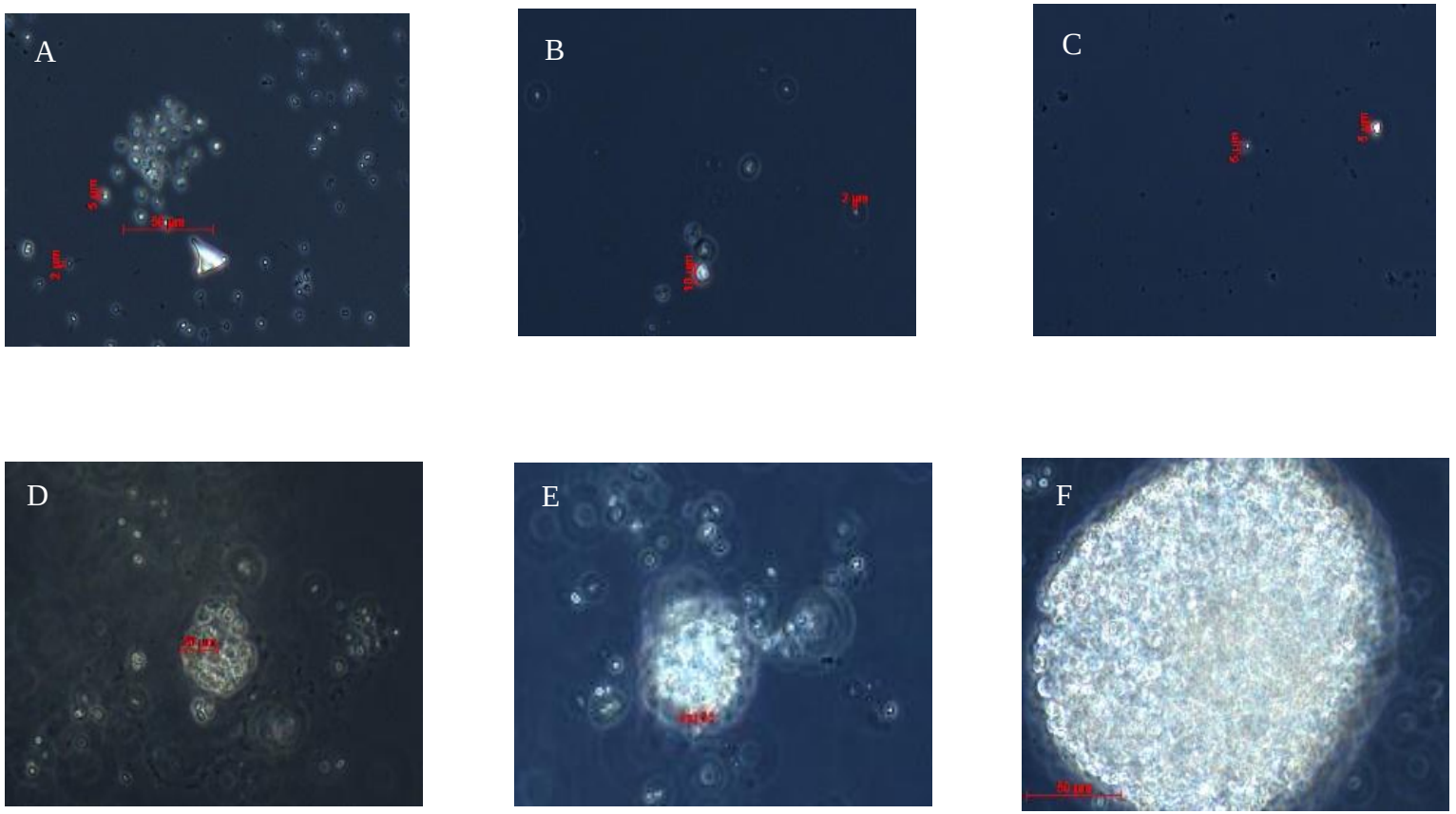


Figure 7 *P. antarctica* (Col.) replete control changes over time; stained with DAPI using epifluorescence microscopy (200x).

(A) Single cells (2 µm) and colony-forming aggregates (5 µm) in diameter forming a small thin colony (50 µm) of the iron-limited T0 (Day 0). (B) Single cells (2 µm) and colony-forming aggregates (10 µm) at Day 1. (C) Single cells and colony-forming aggregates (5 µm) at Day 3. (D) Colony (20 µm) at Day 5. (E) Colony (50 µm) at day 8. (F) Colony (50 µm) at Day 8. B-F are the replete control samples.

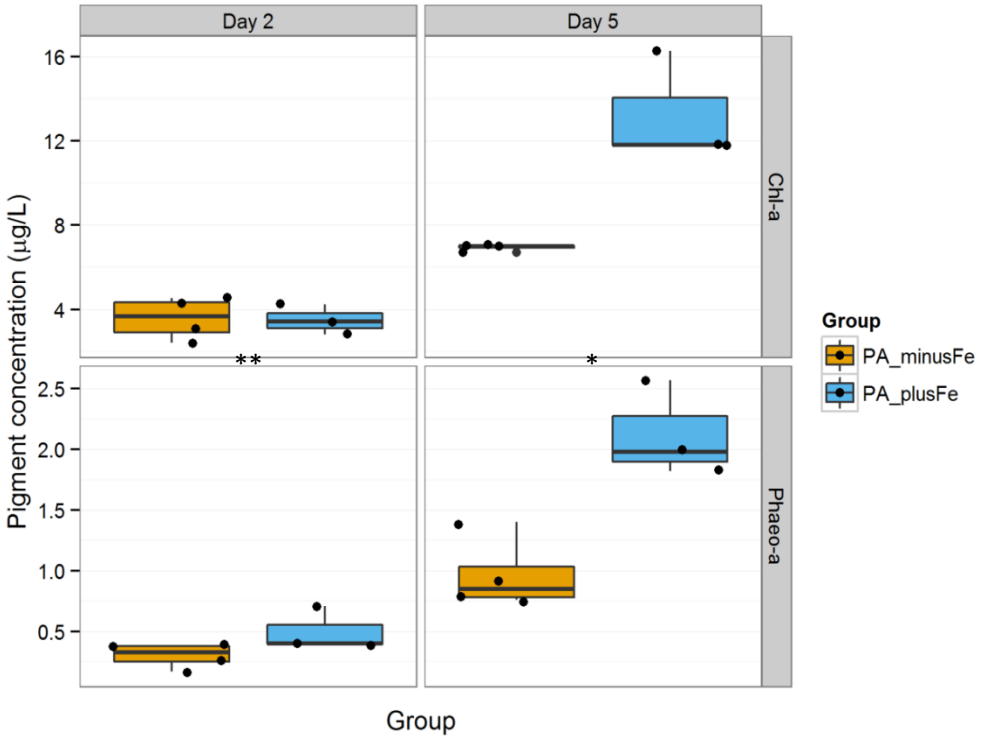


Figure 8 Pigments concentrations in control (n = 3) and treatment (n = 4) culture groups (µg/L).

**significant difference; *no significant difference between Phaeo *a*:Chl *a* ratios (*t*-test *p*-value < 0.05).

A B

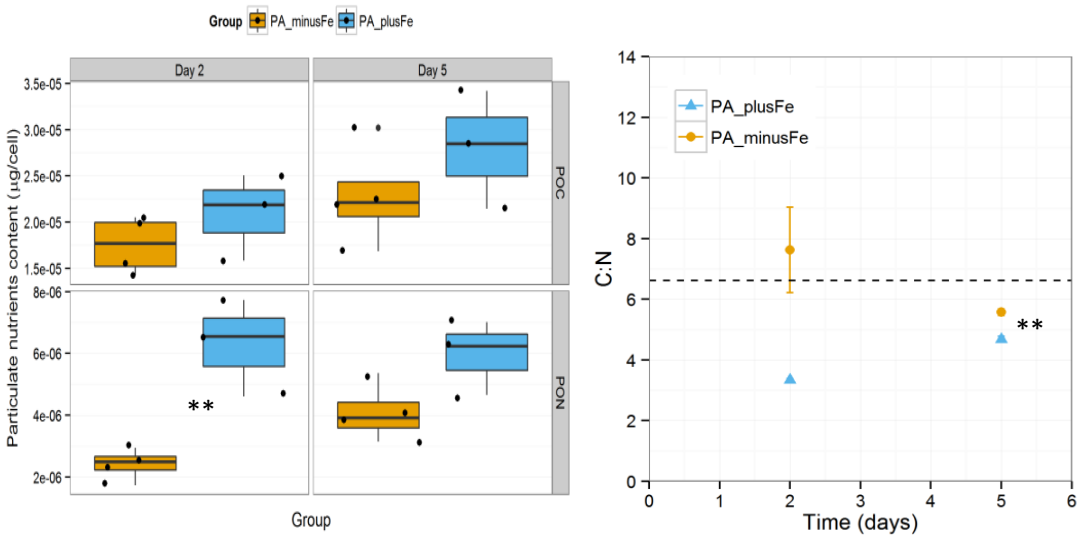
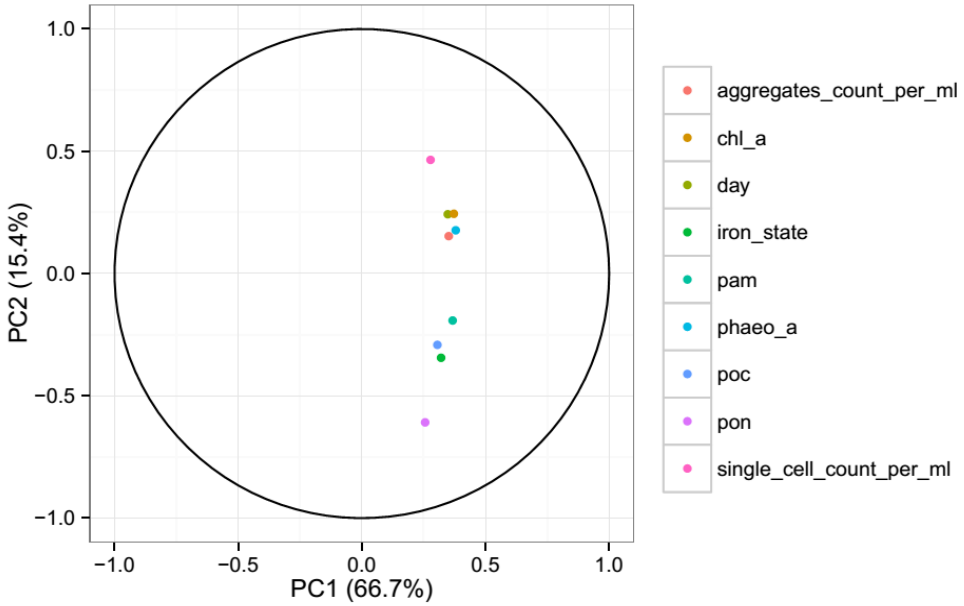


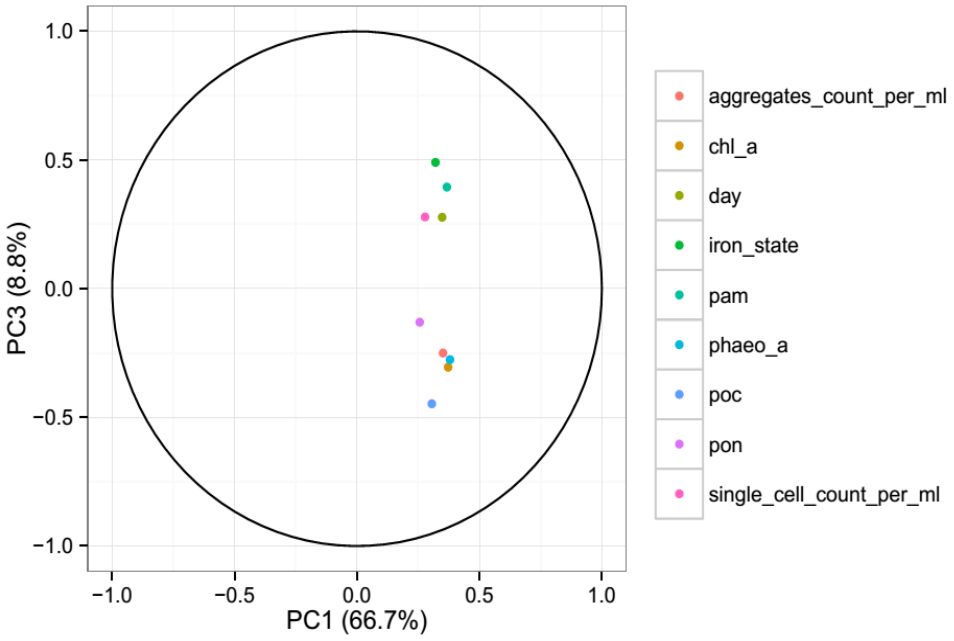
Figure 9 *P. antarctica* POC and PON contents in control (n = 3) and treatment cultures (n = 4) before and after iron supplementation at Day 3.

(A) POC and PON (µg/cell). (B) C:N ratio in control and treatment cultures (mean ± SE), the dotted line represents Redfield ratio. **significant difference (*t*-test *p*-value < 0.05).

A


1.0
0.5
0.0
−0.5
−1.0
PC2 (15.4%)
−1.0
−0.5
0.0
0.5
1.0
PC1 (66.7%)
aggregates_count_per_ml
chl_a
day
iron_state
pam
phaeo_a
poc
pon
single_cell_count_per_ml

B


1.0
0.5
0.0
−0.5
−1.0
PC3 (8.8%)
−1.0
−0.5
0.0
0.5
1.0
PC1 (66.7%)
aggregates_count_per_ml
chl_a
day
iron_state
pam
phaeo_a
poc
pon
single_cell_count_per_ml

C

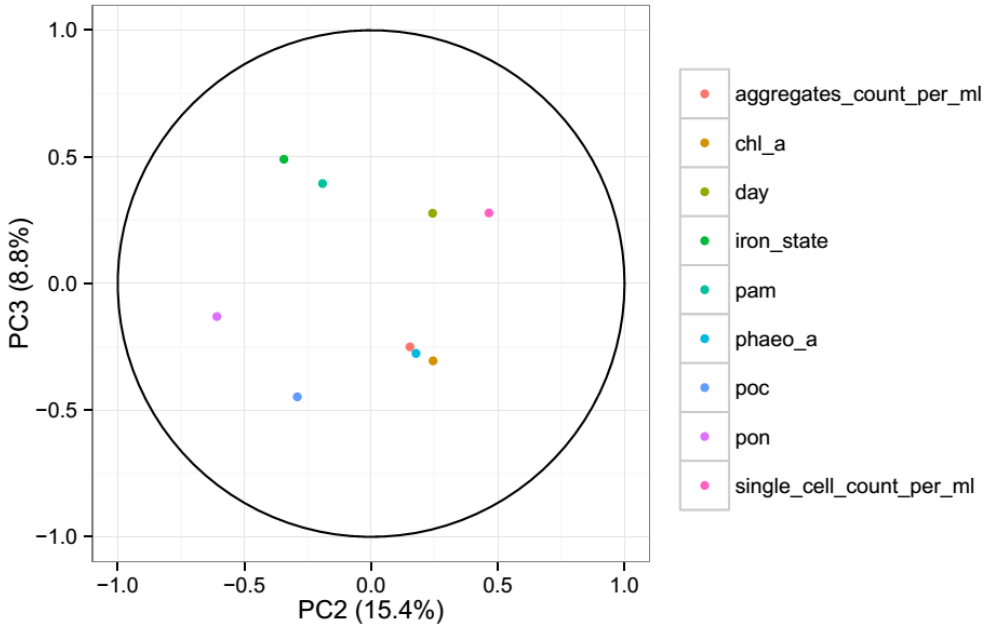


Figure 10 Principal component analysis of *P. antarctica* parameters.

Out of the nine identified principal components, the three of the largest proportion of variance (cumulative proportion of variance = 90%), PC1 (66.7%), PC2 (15.4%), and PC3 (8.8%) were used to visualize the variable coefficients (Martins n.d.) (Code: https://gist.github.com/thigm85/7689508). (A) PC1 *vs.* PC2. (B) PC1 *vs*. PC3. (C) PC2 *vs.* PC3.

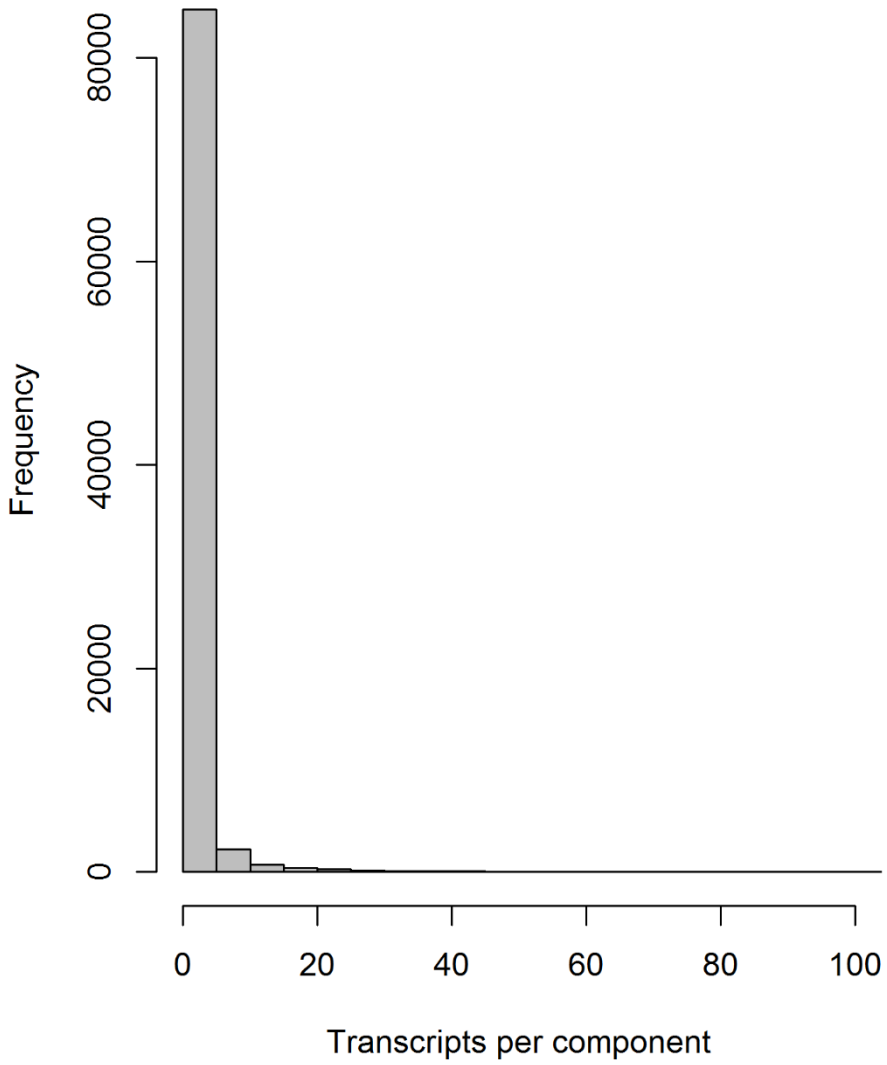


Figure 11 Frequency distribution of *P. antarctica* isoforms (i.e., transcripts) counts per component.

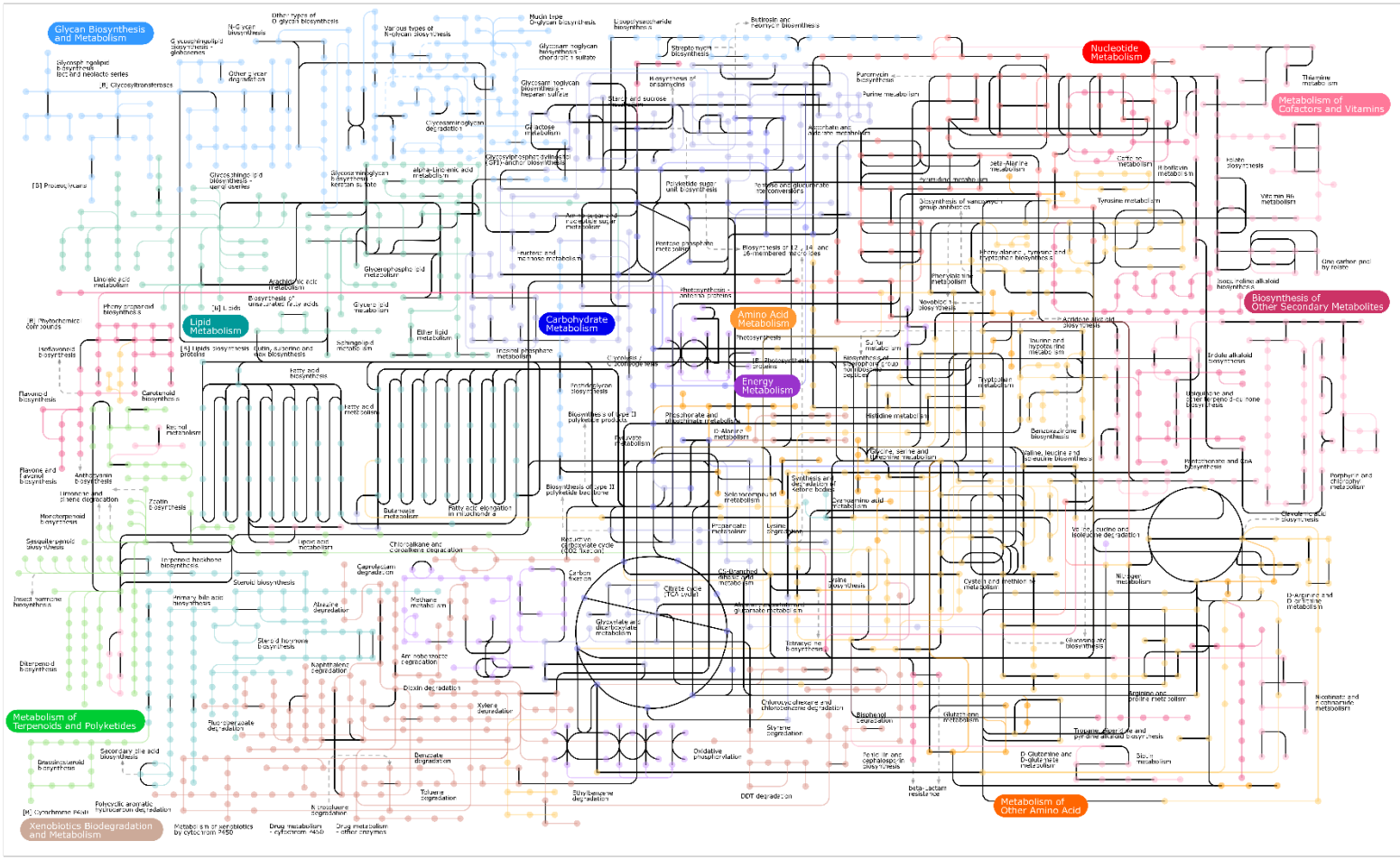


Figure 12 Metabolic potential map of *P. antarctica* under iron –limited and –replete conditions.

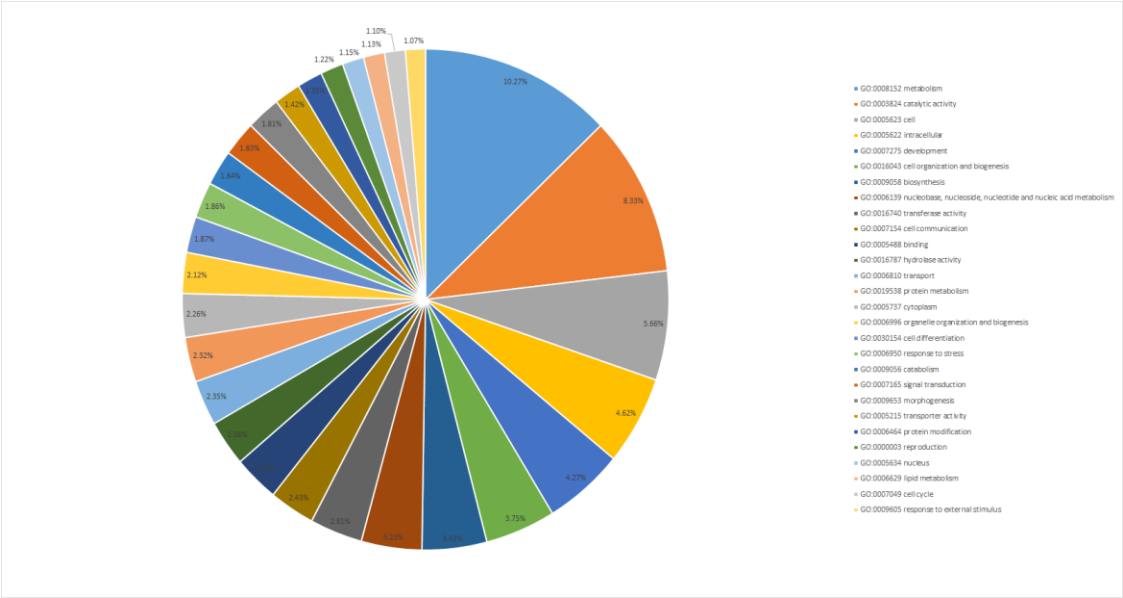


Figure 13 Putative functional categories of P. antarctica transcriptome based on GO terms grouping by CateGOrizer (Zhi-Liang et al. 2008) showing fractions > 1%.

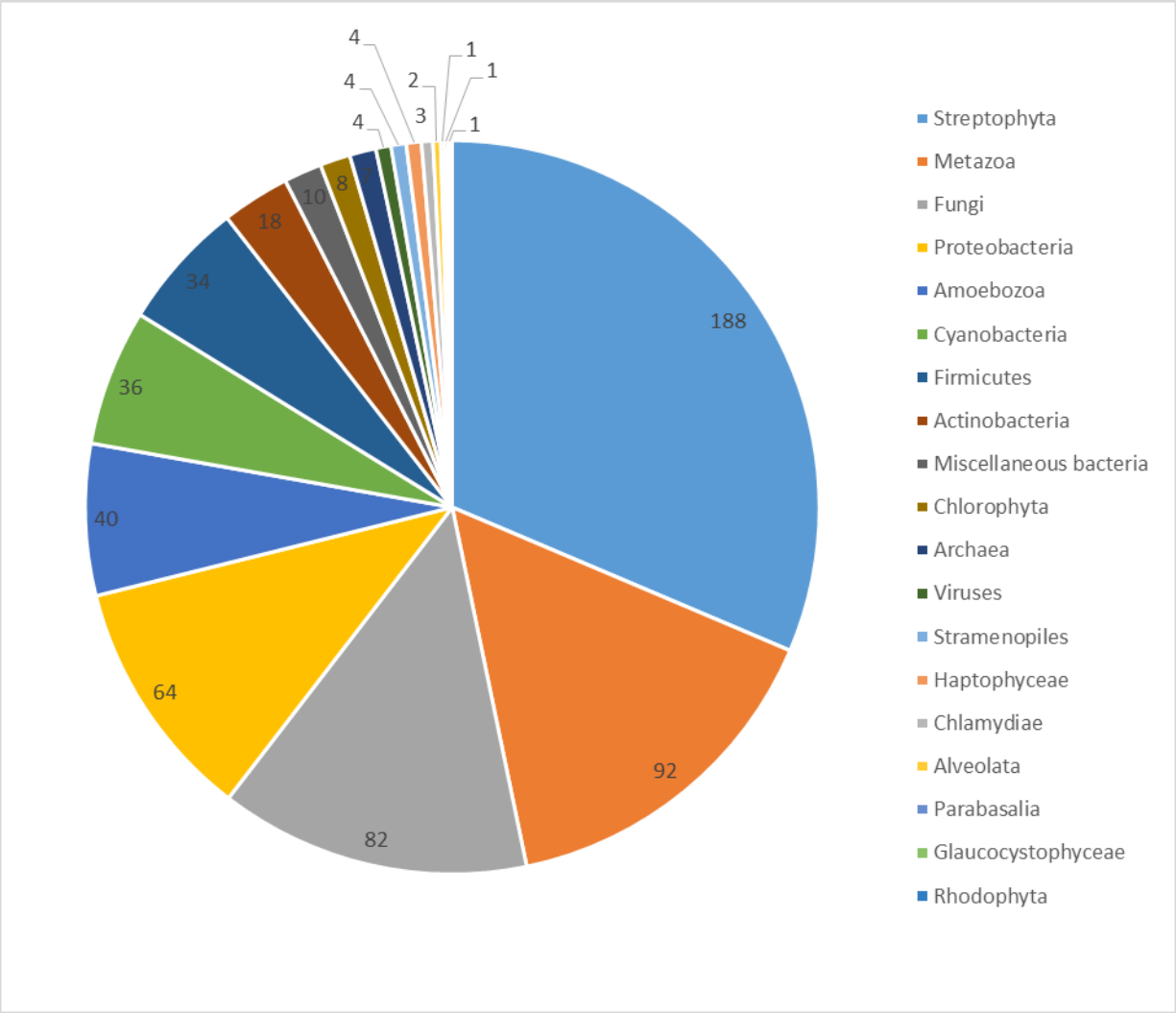


Figure 14 The taxonomic distribution of *P. antarctica* nuclear-encoded plastid-targeted predicted ORFs.

The numbers refer to the counts of ORFs predicted for transcripts (isoforms) based on blastp UniProt hits excluding ORFs of unknown function or assigned to proteins of unknown taxa or to proteins of mammalian origin.

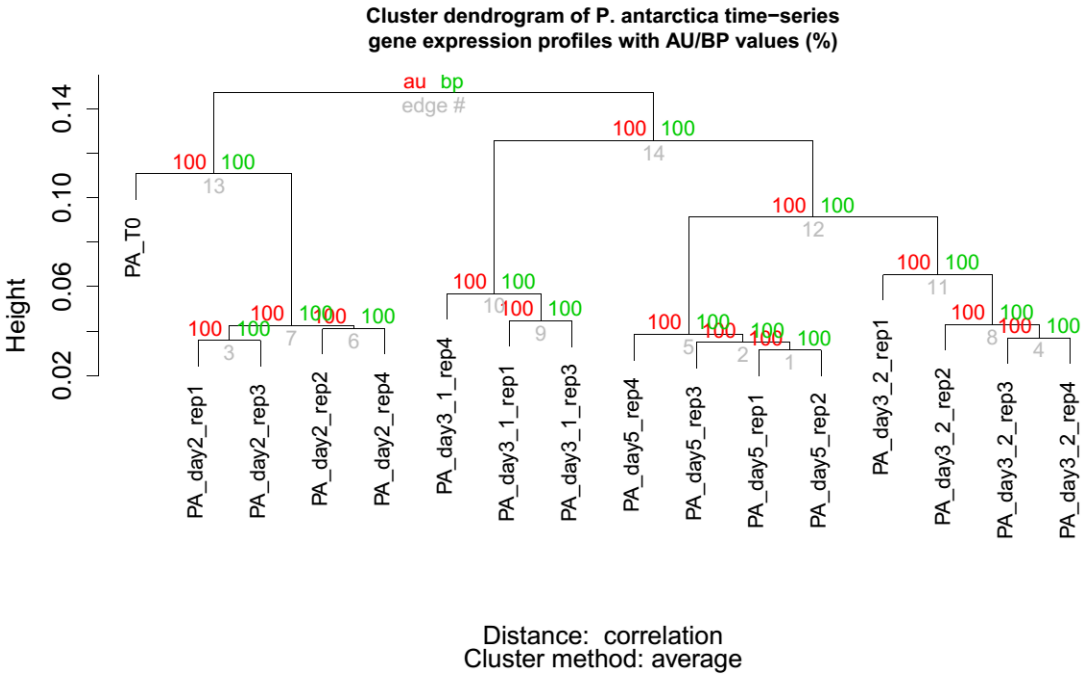


Figure 15 Hierarchical clustering of raw read counts per component across replicates and days (conditions) (bootstrap = 100).

AU = Approximately Unbiased; BP = Bootstrap Probability.

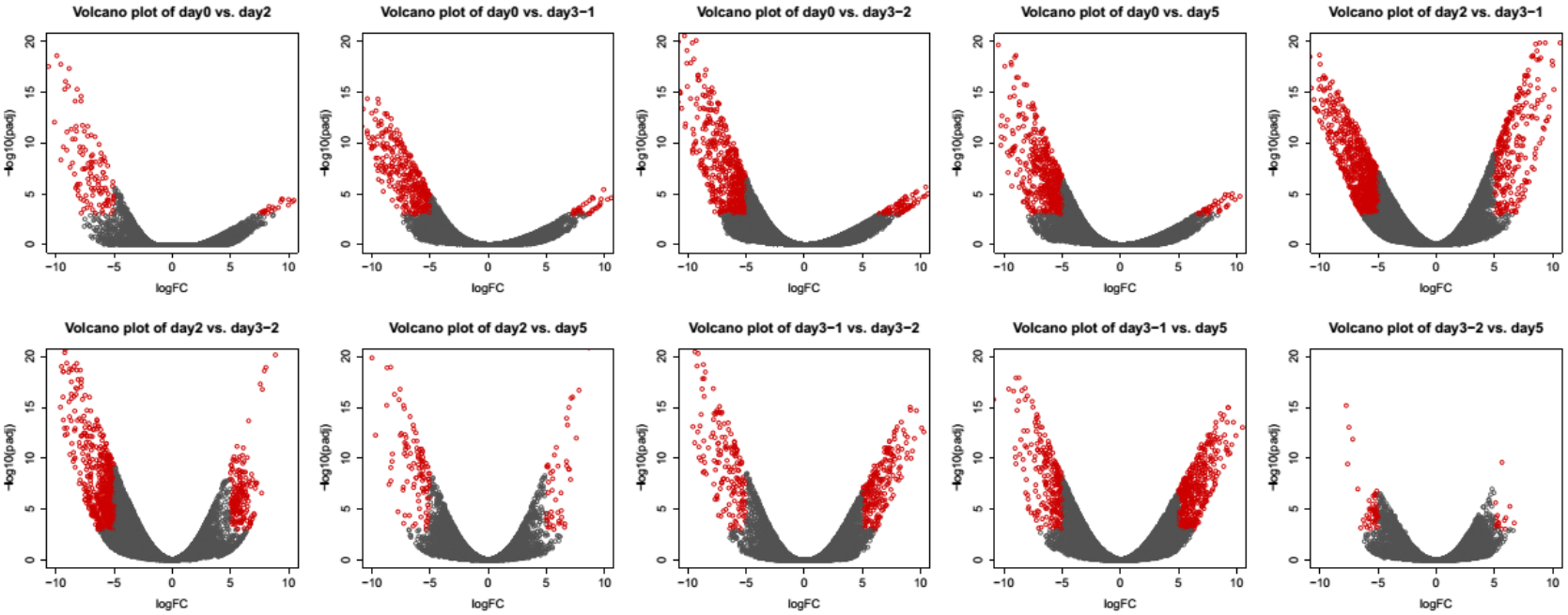


Figure 16 Volcano plot of the binomial test statistics, log fold-change (logFC) *vs.* -log10 adjusted *p*-value [-log10(padj)], between each two days of the experiment estimated dispersions of each expressed component.

The significantly differentially expressed genes are colored in red.

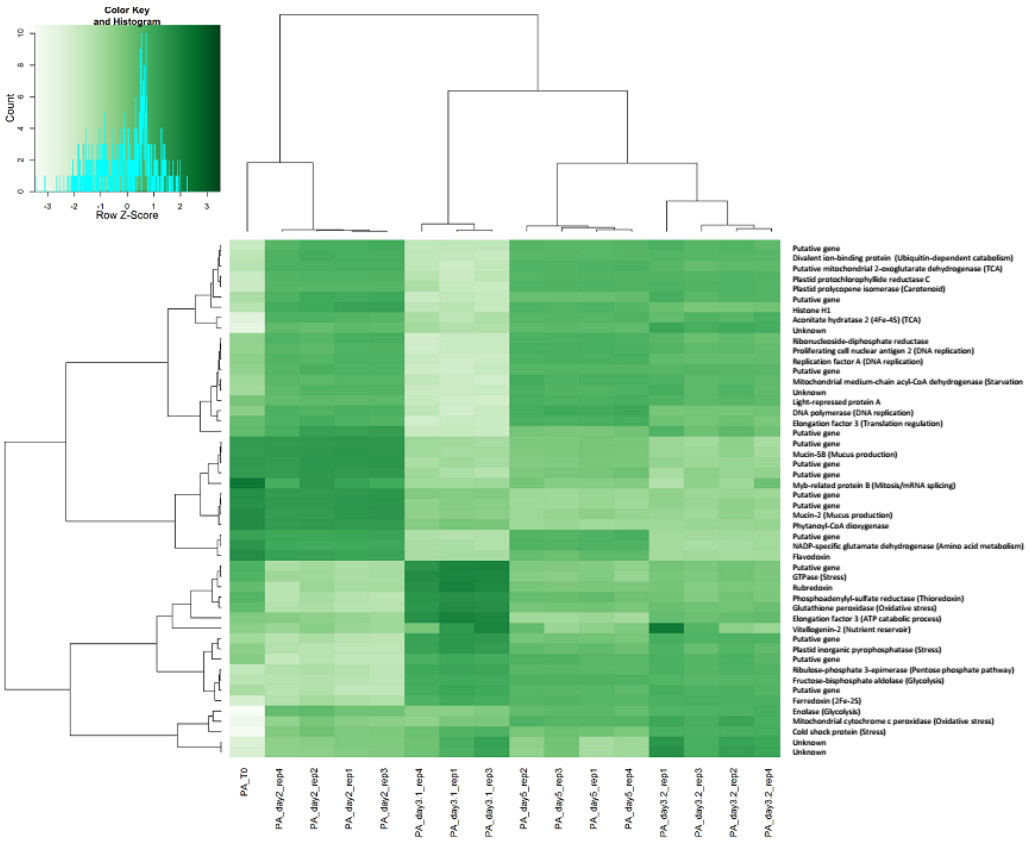


Figure 17 Heatmap of the variance stabilization transformed count data of the 50 most highly differentially expressed genes clustered by correlation.

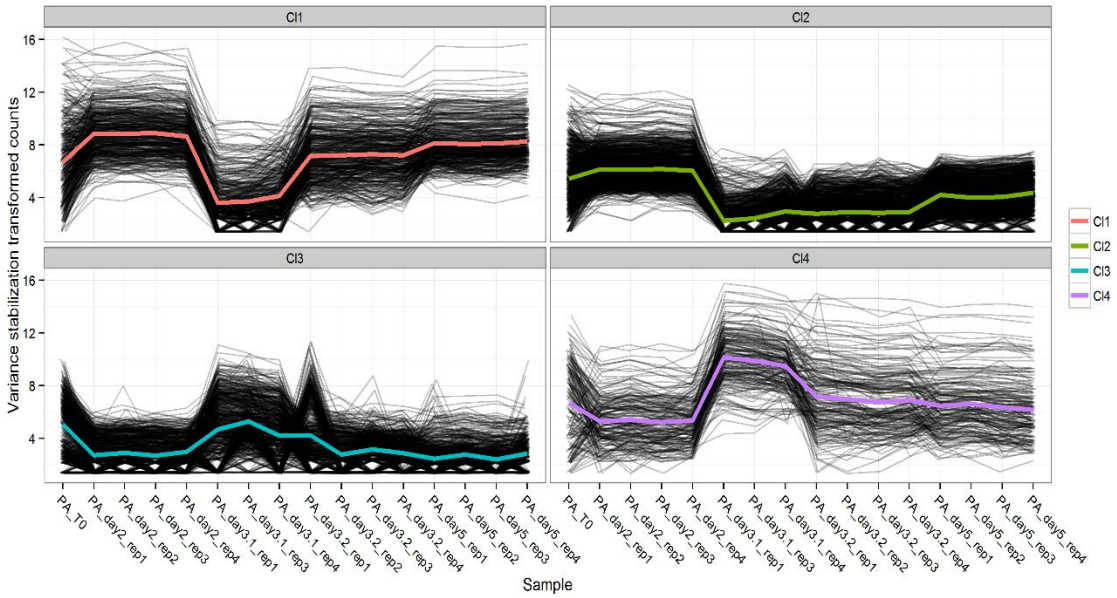
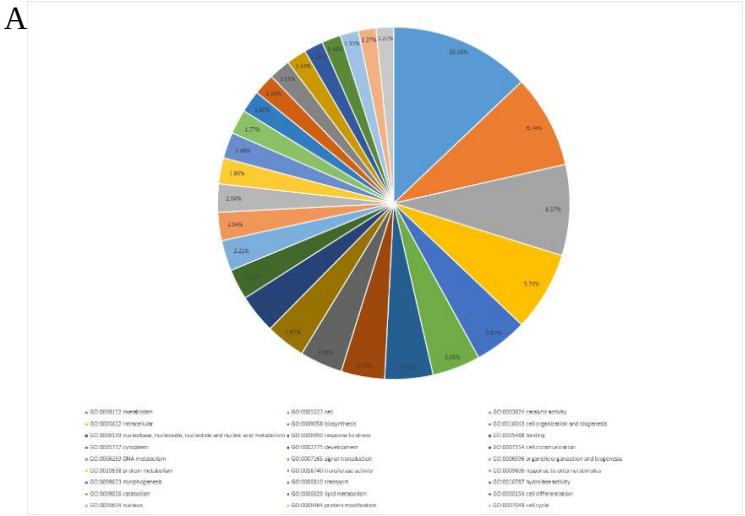


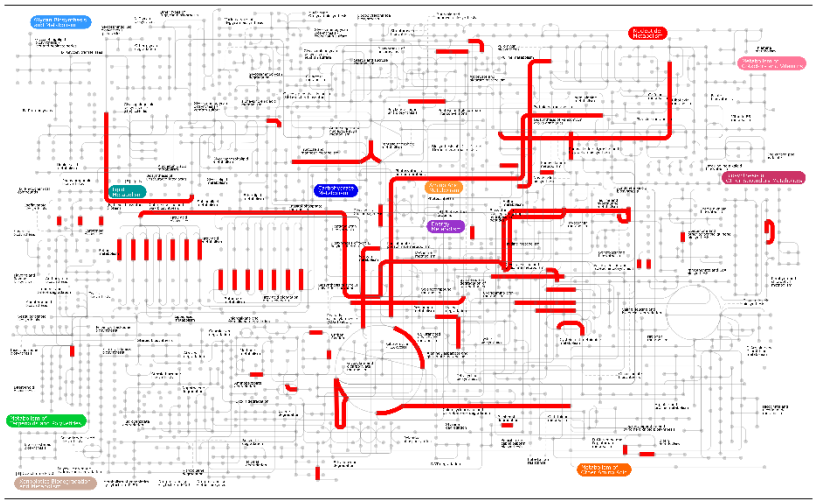
Figure 18 *K*-means clustering of total differentially expressed genes across all samples using variance stabilization transformed counts.

Centers = 4 and algorithm = "Hartigan-Wong"; code adopted from Philip Parker (http://adventuresinr.wordpress.com/2010/10/22/displaying-k-means-results/) (Parker n.d.)

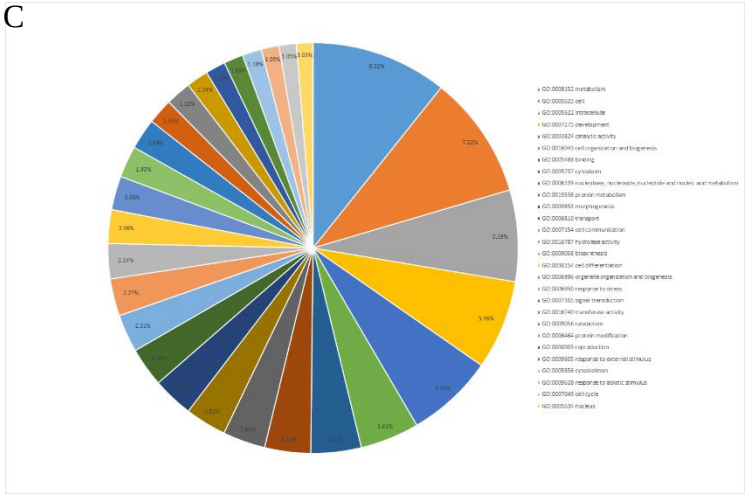
A



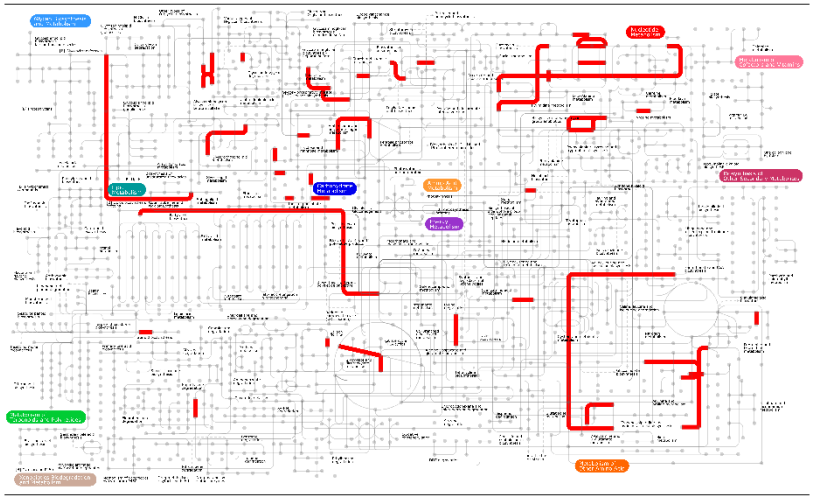
B



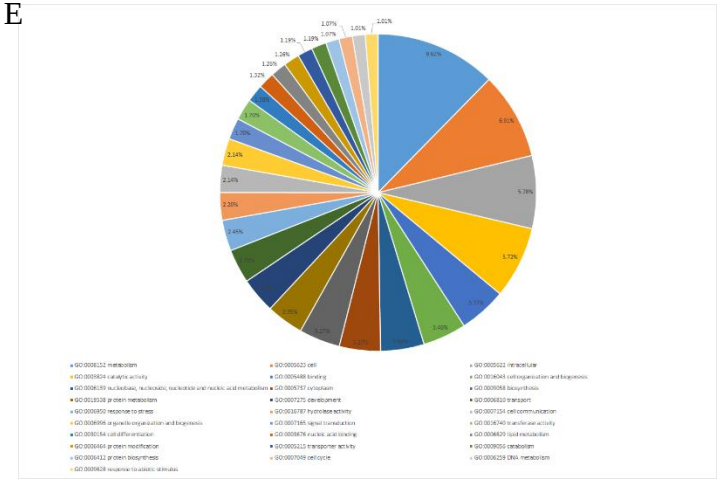
C



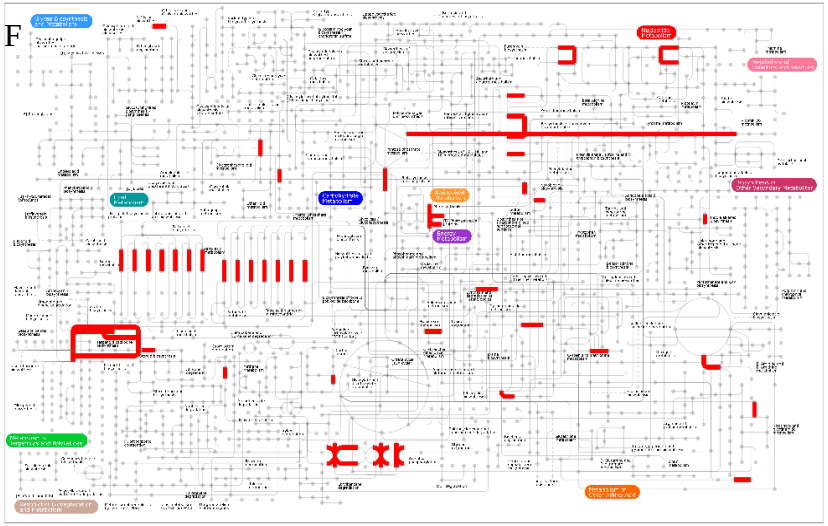
D



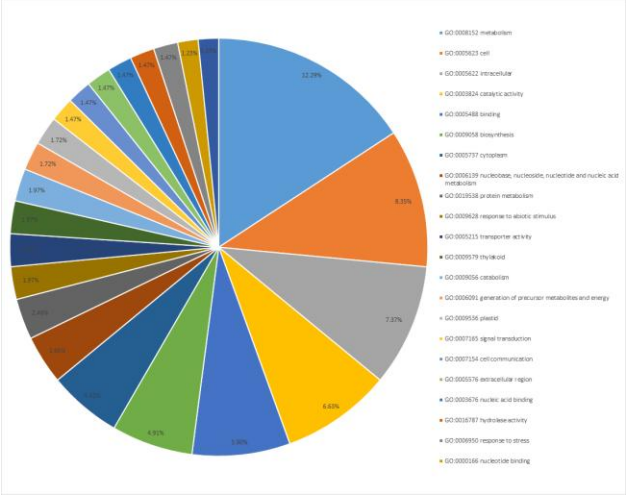
E



F



G



H

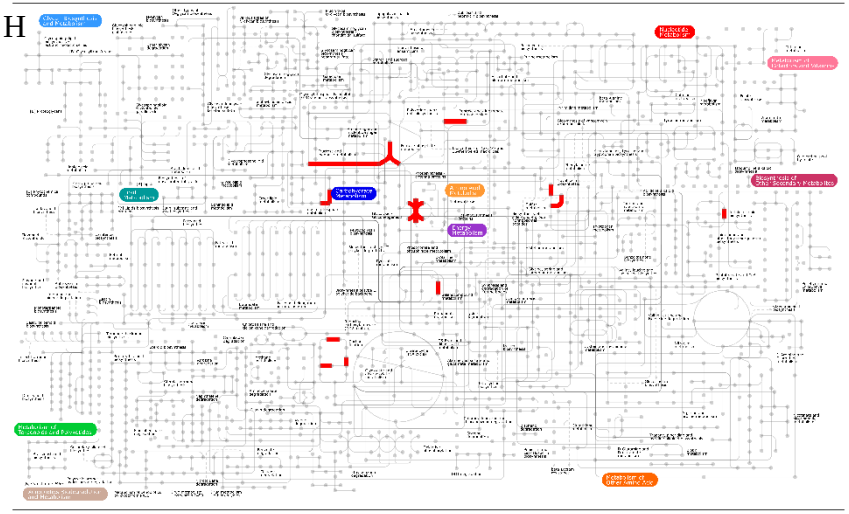


Figure 19 Functional analysis of *k*-means clusters.

Percentages of child GO term occurrences mapped to ancestor terms using CateGOrizer (Zhi-Liang et al. 2008) based on GO_slim2 mapping showing fractions > 1% (left), and pathway visualization using iPath2.0 (Yamada et al. 2011) (right) of the gene members of the five defined clusters by *k*-means clustering. (A-B) cluster 1, (C-D) cluster 2, (E-F) cluster 3, and (G-H) cluster 4.

A B

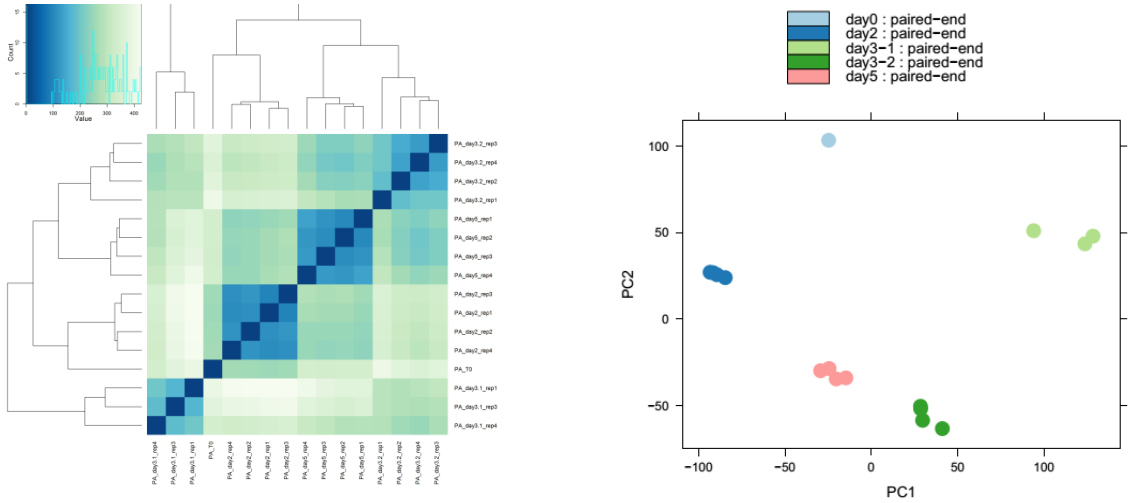


Figure 20 Visualization of samples variance stabilization transformed expression values.

(A) Heatmap of Euclidean sample-to-sample distances. (B) PCA of the samples using 2367 genes based on the two principal components, condition and library type.

Table 5 *P. antarctica* parameters under iron limitation and enrichment (mean ± SD).

Parameter	**Day 1**	**Day 2**	**Day 3**	**Day 4**	**Day 5**	**Day 8**
F_v/F_m						
PA_plusFe	0.35 ± 0.01	0.43 ± 0.03	0.49 ± 0.01	0.52 ± 0.01	0.52 ± 0.01	N/A
PA_minusFe	0.22 ± 0.01	0.26 ± 0.02	0.36 ± 0.02	0.51 ± 0.01	0.50 ± 0.01	N/A
Single cell/ml						
PA_plusFe	1.83E+04 ± 3.47E+03	1.71E+04 ± 3.86E+03	1.32E+04 ± 4.67E+03	1.80E+04 ± 4.27E+03	3.57E+04 ± 9.78E+03	1.02E+05 ± 1.48E+04
PA_minusFe	1.85E+04 ± 3.73E+03	1.90E+04 ± 2.30E+03	1.25E+04 ± 1.41E+03	1.54E+04 ± 7.03E+03	2.79E+04 ± 8.33E+03	9.08E+04 ± 1.98E+04
Colony-forming aggregate/ml						
PA_plusFe	2.26E+03 ± 9.13E+02	2.95E+03 ± 1.40E+03	4.08E+03 ± 1.31E+03	5.19E+03 ± 3.13E+03	1.21E+04 ± 5.34E+03	5.71E+04 ± 4.68E+03
PA_minusFe	1.11E+03 ± 3.21E+02	8.54E+02 ± 1.64E+02	1.68E+03 ± 7.99E+02	2.97E+03 ± 1.87E+03	5.01E+03 ± 6.95E+02	1.77E+04 ± 3.26E+03
Chl *a* (µg/L)						
PA_plusFe	N/A	3.5 ± 0.7	N/A	N/A	13.3 ± 2.6	N/A
PA_minusFe	N/A	3.6 ± 1	N/A	N/A	6.9 ± 0.2	N/A
POC (µg/cell)						
PA_plusFe	N/A	2.09E-05 ±	N/A	N/A	2.80E-05 ±	N/A

		4.68E-06			6.37E-06	
PA_minusFe	N/A	1.75E-05 ± 3.11E-06	N/A	N/A	2.28E-05 ± 5.51E-06	N/A
PON (µg/cell)						
PA_plusFe	N/A	6.30E-06 ± 1.58E-06	N/A	N/A	5.97E-06 ± 1.20E-06	N/A
PA_minusFe	N/A	2.41E-06 ± 5.05E-07	N/A	N/A	4.09E-06 ± 9.36E-07	N/A

Table 6 *P. antarctica* replicates RNA extraction and cDNA library sequence data.

	T0	**Day2_rep1**	**Day2_rep2**	**Day2_rep3**	**Day2_rep4**	**Day3-1_rep1**	**Day3-1_rep3**	**Day3-1_rep4**	**Day3-2_rep1**	**Day3-2_rep2**	**Day3-2_rep3**	**Day3-2_rep4**	**Day5_rep1**	**Day5_rep2**	**Day5_rep3**	**Day5_rep4**
RNA concentration (ng/μL)	36.67	23.97	28.22	26.29	27.53	26.22	30.45	27.86	13.39	19.9	33.4	15.01	60.68	130.49	37.53	54.95
260/280	N/A	2.06	1.92	2.05	1.99	1.66	1.57	1.84	1.61	1.64	1.75	1.63	N/A	N/A	N/A	N/A
RIN	5.10	5.50	6	5.50	4.60	3.70	4	N/A	N/A	N/A	N/A	N/A	5.20	5	5.30	2.80
No. of paired-end reads	27,181,526	21,254,194	23,097,368	23,956,268	23,061,370	21,945,028	24,219,258	22,380,348	23,484,084	19,627,806	18,475,506	22,404,362	38,996,488	32,508,688	26,337,532	20,916,588
Total no. of bases	1,386,257,826	1,083,963,894	1,177,965,768	1,221,769,668	1,176,129,870	1,119,196,428	1,235,182,158	1,141,397,748	1,197,688,284	1,001,918,106	942,250,806	1,142,622,462	1,988,820,888	1,657,943,088	1,343,214,132	1,066,745,988

Table 7 *P. antarctica* transcriptome assembly statistics.

	Number	**Mean (b)**	**Median (b)**	**N50 (b)**	**Total assembled bases**
All transcript contigs	162,436 (transcripts)	840.34	604	1190	136,501,307
Longest isoform per component	88,630 (components)	725.96	495	1041	64,341,503

Table 8 Transcript families identified using eggNOG (≥100 members).

eggNOG ID	Category description	Transcript family size
COG0515	Serine/threonine protein kinase	1864
COG0666	FOG: Ankyrin repeat	1476
COG2319	FOG: WD40 repeat	691
COG5126	Ca2+-binding protein (EF-Hand superfamily)	447
NOG12793	Calcium ion binding protein	417
COG0457	FOG: TPR repeat	379
COG5245	Dynein, heavy chain	352
COG4886	Leucine-rich repeat (LRR) protein	349
COG0513	Superfamily II DNA and RNA helicases	298
COG0553	Superfamily II DNA/RNA helicases, SNF2 family	286
COG5059	Kinesin-like protein	265
COG0477	Permeases of the major facilitator superfamily	251
COG1226	Kef-type K+ transport systems, predicted NAD-binding component	233
COG0697	Permeases of the drug/metabolite transporter (DMT) superfamily	215
COG1028	Dehydrogenases with different specificities (related to short-chain alcohol dehydrogenases)	214
COG0484	DnaJ-class molecular chaperone with C-terminal Zn finger domain	212
COG5021	Ubiquitin-protein ligase	202
COG2940	Proteins containing SET domain	197
COG1132	ABC-type multidrug transport system, ATPase and permease components	193
COG3119	Arylsulfatase A and related enzymes	185
COG0500	SAM-dependent methyltransferases	182
COG4642	Uncharacterized protein conserved in bacteria	179

COG1100	GTPase SAR1 and related small G proteins	179
COG0664	cAMP-binding proteins - catabolite gene activator and regulatory subunit of cAMP-dependent protein kinases	173
COG0526	Thiol-disulfide isomerase and thioredoxins	169
COG0545	FKBP-type peptidyl-prolyl cis-trans isomerases 1	167
COG0642	Signal transduction histidine kinase	164
COG5022	Myosin heavy chain	163
COG0631	Serine/threonine protein phosphatase	154
COG5147	Myb superfamily proteins, including transcription factors and mRNA splicing factors	148
COG0724	RNA-binding proteins (RRM domain)	140
COG0488	ATPase components of ABC transporters with duplicated ATPase domains	136
COG1131	ABC-type multidrug transport system, ATPase component	134
NOG69209	Protein involved in I-kappaB kinase/NF-kappaB cascade	133
NOG241162	Junctional adhesion molecule 3	133
COG1112	Superfamily I DNA and RNA helicases and helicase subunits	123
COG1643	HrpA-like helicases	120
COG0790	FOG: TPR repeat, SEL1 subfamily	119
COG3781	Predicted membrane protein	117
COG0465	ATP-dependent Zn proteases	115
COG5032	Phosphatidylinositol kinase and protein kinases of the PI-3 kinase family	111
COG5038	Ca2+-dependent lipid-binding protein, contains C2 domain	110
COG0652	Peptidyl-prolyl cis-trans isomerase (rotamase) - cyclophilin family	110
COG0464	ATPases of the AAA+ class	100

Table 9 Potential non-coding RNA in *P. antarctica* by transcriptome functional annotation.

(A) rRNA genes identified by RNAmmer. (B) tRNA genes distribution assigned to GO groups. *COG isoleucyl-tRNA synthetase.

A

Transcript id	**rRNA gene**	**From-to**	**Length (bp)**
comp82182_c0_seq1	28S rRNA	57-3839	3783
comp82182_c0_seq1	18S rRNA	4095-5897	1803
comp90860_c5_seq1	8S rRNA	371-486	170
comp92609_c0_seq2	28S rRNA	1106-1619	513
comp92610_c7_seq1	28S rRNA	1-2909	2909
comp92610_c7_seq1	18S rRNA	3044-4518	1475
comp92610_c7_seq4	28S rRNA	15-2909	2895
comp92610_c7_seq6	28S rRNA	15-2909	2895
comp92610_c7_seq6	18S rRNA	3190-4664	1475

B

GO id	GO term	Occurrence
GO:0006423	cysteinyl-tRNA aminoacylation	54
GO:0006435	threonyl-tRNA aminoacylation	18
GO:0006431	methionyl-tRNA aminoacylation	13
GO:0006432	phenylalanyl-tRNA aminoacylation	13
GO:0006419	alanyl-tRNA aminoacylation	12
GO:0006428	isoleucyl-tRNA aminoacylation	11*
GO:0006433	prolyl-tRNA aminoacylation	10
GO:0006434	seryl-tRNA aminoacylation	9
GO:0006422	aspartyl-tRNA aminoacylation	9
GO:0006438	valyl-tRNA aminoacylation	8
GO:0006430	lysyl-tRNA aminoacylation	7
GO:0006425	glutaminyl-tRNA aminoacylation	7
GO:0006421	asparaginyl-tRNA aminoacylation	7
GO:0006427	histidyl-tRNA aminoacylation	6
GO:0006426	glycyl-tRNA aminoacylation	6
GO:0006436	tryptophanyl-tRNA aminoacylation	6
GO:0006424	glutamyl-tRNA aminoacylation	5
GO:0006420	arginyl-tRNA aminoacylation	4
GO:0006437	tyrosyl-tRNA aminoacylation	3
GO:0006429	leucyl-tRNA aminoacylation	3

CHAPTER 4: CONCLUSIONS AND FUTURE DIRECTIONS

4.1 Conclusions

Phaeocystis antarctica is one of the most ecologically important species endemic to the Southern Ocean due to its contribution to global sulfur gases contents. The growth and productivity of *P. antarctica* in a high-nitrate low-chlorophyll (HNLC) region such as the SO have been demonstrated to be limited by iron availability. The global influence of the SO as a major Earth system climate regulator due to its unique hydrography consequently places its phytoplankton assemblage as drivers of the global carbon cycle. Iron fertilization experiments were conducted in the SO to understand the response of its photosynthetic assemblage to iron addition, and the nanoflagellate *P. antarctica* has been reported the first to bloom and alters sulfur gases contents. Being an ideal ecologically important model organism, *P. antarctica* is studied at a transcriptomic level in a time-series manner to understand phytoplankton adaptation to iron limitation and functional changes following iron addition.

Here the first reported transcriptome of *P. antarctica* revealed that it constitutes of 88,630 putative genes (162,436 transcripts). The unexpectedly high number of potential genes is due to low cutoffs used to capture all genes and isoforms, on one hand. On the other hand, a suggested high number of polymorphic sites among the population led to the high reported number of putative transcripts. The estimates will be corrected in a further assembly. The vast majority of the genes are of unknown function (64,822; 73.1%), while 17,729 (20%; 2,932 unique groups) were assigned to eggNOG orthologous groups upon which the functional annotation of *P. antarctica* was based. Analysis of *P. antarctica* also revealed 2,456 nuclear-encoded plastid-targeted ORFs, the majority of which are of green algal origin. Differential expression at each time-point before and after iron addition was inferred at an adjusted p-value ≤ 0.001 and a log fold-change ≥ 5, revealing a total of 2,367 differentially expressed genes.

At the physiological examination level, iron-limited *P. antarctica* recovered its photosynthetic fitness, colony-forming ability, and chlorophyll *a*, particulate organic carbon and nitrogen contents shortly after iron addition comparable to the replete control

and the reported values of healthy cells. Transcriptomic data supported the observed photosynthetic fitness recovery. It revealed a floristic shift to the more efficient photopigment fucoxanthin production, and also the utilization of the more efficient iron-dependent ferredoxin after iron enrichment. In addition, a shift in gene expression from iron-economic reactive oxygen species defense and photosystem II to iron-dependent alternatives has been observed following iron enrichment. Transcriptomic data supports the previous studies that *P. antarctica* is successful in utilizing bound iron in a reductive non-ligand-dependent mechanism (Strzepek et al. 2011). At the metabolic level, the physiology data was supported by the transcriptomic data suggesting N and C reallocations under iron limitation. In addition, Calvin cycle enzymes were overexpressed under iron enrichment unlike the reports from diatoms (Allen et al. 2008).

In summary, iron levels limits *P. antarctica* growth. It arrests its growth and bloom-forming ability. *P. antarctica* shortly recovers its growth and productivity following iron supplementation (Marchetti et al. 2012). However, iron requirements of *P. antarctica* are the lowest among all phytoplankton species and its adaptation to iron limitation is well-established. Further investigation is required in order to infer the complete metabolic pathways expression at iron limitation and repletion.

4.2 Future Directions

4.2.1 Transcriptomics

4.2.1.1 Corrected Protein-Encoding Genes Estimation and Data Availability

The transcriptome assembly will be revisited to carefully distinguish between alleles, isoforms, and duplicates (i.e., paralogs). Trinity also will be used implementing `--jaccard_clip` parameter to better resolve fused transcripts (Haas et al. 2013) and PasaFly and CuffFly algorithms with stringency to reduce the reported transcriptome estimate (Haas n.d.). In addition, the now available organelles genomes (Smith et al. 2014) will be used in functional annotation. In regard to differential gene expression analysis, TMM-normalized expression values will be used to infer differential expression at both gene-level and transcript-level. Enrichment analysis of GO terms of the

predefined significantly differentially expressed genes will be conducted using topGO (v2.16.0) (Alexa, Rahnenführer and Lengauer, 2006). Genes-to-GO mapping will be based on the multi-hierarchical GO mapping output of Trinotate.

The raw sequencing reads will be made available at the Sequence Read Archive (SRA), while the corrected annotated transcriptome sequence data will be made publically available at the Transcriptome Shotgun Assembly (TSA) archive. In addition, the time-series gene expression data will be submitted to the Gene Expression Omnibus (GEO) database.

4.2.1.2 Permanently Overexpressed and Missing Genes Analysis and Small RNAs (sRNA) Possible Targets in *P. antarctica*

Fucoxanthin-chlorophyll a-c binding protein B and the mitochondrial gene expression regulation protein TAR1 (~23,250 and ~10,649 mean TMM-normalized expression, respectively), for instance, have showed a constitutive high non-differenrial expression pattern. A high resolution analysis of such genes unaffected by the iron state of *P. antarctica* is needed to understand their role in *P. antarctica* cellular processes. Functional domain search, phylogenetic analysis and pathway enrichment analysis of the constitutively expressed genes will provide us with insight into the core metabolism of *P. antarctica*. On the other hand, central metabolic pathways will be mined for predicting missing key genes using comparative genomics (Osterman & Overbeek 2003) against *E. huxleyi* genome.

The role of small RNAs (sRNA) in post-transcriptional gene regulation in land plants and fungi is well-established (Finnegan & Matzke 2003). In addition, recent evidence of sRNA expression in the adapted to iron limitation diatom *Thalassiosira pseudonana* has been reported (Norden-Krichmar et al. 2011). sRNA are known to be involved in regulating bacterial response towards iron stress (Murphy & Payne 2007; Jacques et al. 2006; Massé & Gottesman 2002). Having in hand the genome of *P. antarctica* sister *E. huxleyi* (Read et al. 2013) and sRNA analysis tools such as SoMART (Li et al. 2012) and microRNA plant target prediction tool (Moxon et al. 2008), perhaps

sRNA target prediction using sRNA libraries constructed from *E. huxleyi* can find targets in the transcriptome of *P. antarctica*.

4.2.1.3 Validation of Transcriptomic Data

Real-time quantitative reverse transcription-PCR (qRT-PCR) will be used for validation of the transcriptomic data. The expression of a subset of the highest differentially expressed genes across all days of the experiment (e.g., light-repressed protein A) will be assessed using a high-throughput sequencing-independent expression profiling approach. Having the corrected sequences in hand, gene-specific primers will be designed to track the expression of these genes quantitatively in real-time as well as the expression of the identified housekeeping genes as control.

4.2.2 Comparison to Other Algal Classes from Other HNLC Regions

The long-term success of diatoms in the open ocean over haptophytes (Assmy et al. 2007; Marchetti et al. 2012) raises a question of the fitness characteristics to which such a success is attributed. In spite of having the machinery needed for N utilization, haptophytes were outnumbered by diatoms in *in situ* (Assmy et al. 2007) and *in vitro* mesocosm (Marchetti et al. 2012) iron fertilization experiments after their initial peak, typical to spring *P. antarctica* fate (Smith et al. 1998). In the present study, *P. antarctica* clearly showed significant increase in N utilization. In the absence of grazing pressure, a comparative transcriptomic study between the two algal classes endemic to the SO is pivotal to elucidate whether the success of diatoms attributed to their N utilization ability (Smith et al. 1998) or to secondary metabolic signals produced by the able-to-recover haptophytes (Tang et al. 2009).

4.2.3 Proteomics of Iron Limitation in *P. antarctica* and Haptophyta Iron Utilization Model

The ultimate aim of the study is to model iron utilization in Haptophyta. To fulfill this aim, firstly, *P. antarctica* would extremely benefit from a proteomic study of cellular iron-bound proteins similar to that of *E. huxleyi* (Sutak et al. 2012) to complement the

current study. Nevertheless, metabolomics will enrich the current knowledge of *P. antarctica* iron limitation adaptive mechanisms assessing DMSP- and iron scavenging-related proteins. Combining both transcriptomic and proteomic data, as previously in diatoms (Allen et al. 2008; Dyhrman et al. 2012), will better elucidate the cellular pathways and processes in the iron-limitation-adapted haptophyte, *P. antarctica*.

REFERENCES

Alderkamp, A.-C. et al., 2012. Iron from melting glaciers fuels phytoplankton blooms in the Amundsen Sea (Southern Ocean): Phytoplankton characteristics and productivity. *Deep Sea Research Part II: Topical Studies in Oceanography*, 71-76, pp.32–48.

Allcock, A.L. & Strugnell, J.M., 2012. Southern Ocean diversity: new paradigms from molecular ecology. *Trends in ecology & evolution*, 27(9), pp.520–8.

Allen, A.E. et al., 2012. Evolution and functional diversification of fructose bisphosphate aldolase genes in photosynthetic marine diatoms. *Molecular biology and evolution*, 29(1), pp.367–79.

Allen, A.E. et al., 2008. Whole-cell response of the pennate diatom Phaeodactylum tricornutum to iron starvation. *Proceedings of the National Academy of Sciences of the United States of America*, 105(30), pp.10438–43.

Altschul, S.F. et al., 1990. Basic local alignment search tool. *Journal of molecular biology*, 215(3), pp.403–10.

Anders, S. & Huber, W., 2010. Differential expression analysis for sequence count data. *Genome biology*, 11(10), p.R106.

Andersen, R.A., 2004. Biology and systematics of heterokont and haptophyte algae. *American journal of botany*, 91(10), pp.1508–22.

Anderson, R., Ali, S. & Bradtmiller, L., 2009. Wind-driven upwelling in the Southern Ocean and the deglacial rise in atmospheric CO2. *Science*, 323(5920), pp.1443–1448.

Anon, Advanced Guide to Trinity. Available at: http://trinityrnaseq.sourceforge.net/advanced_trinity_guide.html [Accessed May 12, 2014].

Arrigo, K.R. et al., 2010. Photophysiology in two major southern ocean phytoplankton taxa: photosynthesis and growth of Phaeocystis antarctica and Fragilariopsis cylindrus under different irradiance levels. *Integrative and comparative biology*, 50(6), pp.950–66.

Arrigo, K.R., 1999. Phytoplankton Community Structure and the Drawdown of Nutrients and CO2 in the Southern Ocean. *Science*, 283(5400), pp.365–367.

Ashburner, M. et al., 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature genetics*, 25(1), pp.25–9.

Assmy, P. et al., 2007. Mechanisms determining species dominance in a phytoplankton bloom induced by the iron fertilization experiment EisenEx in the Southern Ocean. *Deep Sea Research Part I: Oceanographic Research Papers*, 54(3), pp.340–362.

Assmy, P. et al., 2013. Thick-shelled, grazer-protected diatoms decouple ocean carbon and silicon cycles in the iron-limited Antarctic Circumpolar Current. *Proceedings of the National Academy of Sciences of the United States of America*, 110(51), pp.20633–8.

Azam, F. & Malfatti, F., 2007. Microbial structuring of marine ecosystems. *Nature reviews. Microbiology*, 5(10), pp.782–91.

De Baar, H.J.W. et al., 2005. Synthesis of iron fertilization experiments: From the Iron Age in the Age of Enlightenment. *Journal of Geophysical Research*, 110(C9), p.C09S16.

De Baar, H.J.W., 1994. von Liebig's law of the minimum and plankton ecology (1899–1991). *Progress in Oceanography*, 33(4), pp.347–386.

Bakker, D.C.E. et al., 2005. Iron and mixing affect biological carbon uptake in SOIREE and EisenEx, two Southern Ocean iron fertilisation experiments. *Deep Sea Research Part I: Oceanographic Research Papers*, 52(6), pp.1001–1019.

Bentley, D.R. et al., 2008. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*, 456(7218), pp.53–9.

Bertrand, E.M. et al., 2011. Iron limitation of a springtime bacterial and phytoplankton community in the ross sea: implications for vitamin b(12) nutrition. *Frontiers in microbiology*, 2, p.160.

Beszteri, S., 2011. *Functional genomic insights into cellular processes related to harmful bloom formation in ichthyotoxic prymnesiophytes*. Universitaet Bremen.

Beszteri, S. et al., 2012. Transcriptomic response of the toxic prymnesiophyte Prymnesium parvum (N. Carter) to phosphorus and nitrogen starvation. *Harmful Algae*, 18, pp.1–15.

Billard, C. & Inouye, I., 2004. What is new in coccolithophore biology? In H. R. Thierstein & J. R. Young, eds. *Coccolithophores: from molecular processes to global impact*. Springer Berlin Heidelberg.

Boyd, 2002a. Environmental factors controlling phytoplankton processes in the southern ocean. *Journal of Phycology*, 38(5), pp.844–861.

Boyd, 2002b. The role of iron in the biogeochemistry of the Southern Ocean and equatorial Pacific: a comparison of in situ iron enrichments. *Deep Sea Research Part II: Topical Studies in Oceanography*, 49(9-10), pp.1803–1821.

Boyd, P. et al., 2005. The evolution and termination of an iron-induced mesoscale bloom in the northeast subarctic Pacific. *Limnology and Oceanography; (United States)*.

Boyd, P., Bakker, D. & Chandler, C., 2012. A new database to explore the findings from large-scale ocean iron enrichment experiments. *Oceanography*, 25(4), pp.64–71.

Boyd, P. & Denman, K., 2008. Implications of large-scale iron fertilization of the oceans. *Marine Ecology Progress Series*, 364, pp.213–309.

Boyd, P.W. et al., 2007. Mesoscale iron enrichment experiments 1993-2005: synthesis and future directions. *Science (New York, N.Y.)*, 315(5812), pp.612–7.

Boyd, P.W. & Doney, S.C., 2002. Modelling regional responses by marine pelagic ecosystems to global climate change. *Geophysical Research Letters*, 29(16), pp.53–1–53–4.

Boyd, P.W. & Ellwood, M.J., 2010. The biogeochemical cycle of iron in the ocean. *Nature Geoscience*, 3(10), pp.675–682.

Boye, M. & van den Berg, C.M.G., 2000. Iron availability and the release of iron-complexing ligands by Emiliania huxleyi. *Marine Chemistry*, 70(4), pp.277–287.

Bracher, A., Kroon, B. & Lucas, M., 1999. Primary production, physiological state and composition of phytoplankton in the Atlantic sector ot the Southern Ocean. *Marine Ecology Progress Series*.

Broecker, W.S. & Henderson, G.M., 1998. The sequence of events surrounding Termination II and their implications for the cause of glacial-interglacial CO 2 changes. *Paleoceanography*, 13(4), pp.352–364.

Buesseler, K.O. et al., 2004. The effects of iron fertilization on carbon sequestration in the Southern Ocean. *Science (New York, N.Y.)*, 304(5669), pp.414–7.

Burki, F. et al., 2012. The evolutionary history of haptophytes and cryptophytes: phylogenomic evidence for separate origins. *Proceedings. Biological sciences / The Royal Society*, 279(1736), pp.2246–54.

Camacho, C. et al., 2009. BLAST+: architecture and applications. *BMC bioinformatics*, 10(1), p.421.

Caruana, A.M.N. et al., 2012. Concentrations of dimethylsulphoniopropionate and activities of dimethylsulphide-producing enzymes in batch cultures of nine dinoflagellate species. *Biogeochemistry*, 110(1-3), pp.87–107.

Charles, C.D. et al., 1991. Biogenic opal in Southern Ocean sediments over the last 450,000 years: Implications for surface water chemistry and circulation. *Paleoceanography*, 6(6), pp.697–728.

Chretiennot-Dinet, M.J. et al., 1997. The chitinous nature of filaments ejected by Phaeocystis (Prymnesiophyceae). *Journal Of Phycology*, 33(4), pp.666–672.

Le Clainche, Y. et al., 2006. Modeling analysis of the effect of iron enrichment on dimethyl sulfide dynamics in the NE Pacific (SERIES experiment). *Journal of Geophysical Research*.

Coale, K.H. et al., 2004. Southern Ocean iron enrichment experiment: carbon cycling in high- and low-Si waters. *Science (New York, N.Y.)*, 304(5669), pp.408–14.

Dambek, M. et al., 2012. Biosynthesis of fucoxanthin and diadinoxanthin and function of initial pathway genes in Phaeodactylum tricornutum. *Journal of experimental botany*, 63(15), pp.5607–12.

Darling, K.F. et al., 2000. Molecular evidence for genetic mixing of Arctic and Antarctic subpolar populations of planktonic foraminifers. *Nature*, 405(6782), pp.43–7.

Delwiche, C.F. & Palmer, J.D., 1997. The origin of plastids and their spread via secondary symbiosis. *Plant Systematics and Evolution*, pp.53–86.

DiTullio, G.R. et al., 2007. Effects of iron concentration on pigment composition in Phaeocystis antarctica grown at low irradiance. In M. A. Leeuwe et al., eds. *Phaeocystis, major link in the biogeochemical cycling of climate-relevant elements SE - 7*. Springer Netherlands, pp. 71–81.

Dugdale, R.C. & Wilkerson, F.P., 1991. Low specific nitrate uptake rate: A common feature of high-nutrient, low-chlorophyll marine ecosystems. *Limnology and Oceanography; (United States)*, 36:8.

Durkin, C. et al., 2012. Frustule-related gene transcription and the influence of diatom community composition on silica precipitation in an iron-limited environment. *Limnol. Oceanogr*, 57(6), pp.1619–1633.

Van Duyl, F.C. et al., 1998. Biological control of short-term variations in the concentration of DMSP and DMS during a Phaeocystis spring bloom. *Journal of Sea Research*, 40(3-4), pp.221–231.

Dyhrman, S.T. et al., 2006. Long serial analysis of gene expression for gene discovery and transcriptome profiling in the widespread marine coccolithophore Emiliania huxleyi. *Applied and environmental microbiology*, 72(1), pp.252–60.

Dyhrman, S.T. et al., 2012. The transcriptome and proteome of the diatom Thalassiosira pseudonana reveal a diverse phosphorus stress response. *PloS one*, 7(3), p.e33768.

Eddy, S.R., 2011. Accelerated Profile HMM Searches. W. R. Pearson, ed. *PLoS computational biology*, 7(10), p.e1002195.

Eddy, S.R., 1998. Profile hidden Markov models. *Bioinformatics*, 14(9), pp.755–763.

Edvardsen, B. et al., 2000. Phylogenetic reconstructions of the Haptophyta inferred from 18S ribosomal DNA sequences and available morphological data. *Phycologia*, 39(1), pp.19–35.

Edvardsen, B. & Imai, I., 2006. The Ecology of Harmful Flagellates Within Prymnesiophyceae and Raphidophyceae. In E. Granéli & J. T. Turner, eds. *Ecology of Harmful Algae*. Ecological Studies. Springer Berlin Heidelberg.

Evans, C.A., O'Reilly, J.E. & Thomas, J.P., 1987. *A handbook for the measurement of chlorophyll a and primary production*, Biomass 8, Texas A&M University, College Station.

Field, C.B. et al., 1998. Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components. *Science*, 281(5374), pp.237–240.

Finnegan, E.J. & Matzke, M.A., 2003. The small RNA world. *Journal of cell science*, 116(Pt 23), pp.4689–93.

Gaebler-Schwarz, S. et al., 2010. A new cell stage in the haploid-diploid life cycle of the colony-forming haptophyte Phaeocystis antarctica and its ecological implications. *Journal of Phycology*, 46(5), pp.1006–1016.

Gaebler-Schwarz, S., 2009. *Estimation of genetic diversity in the colony forming polar prymnesiophyte species Phaeocystis antarctica*. Universitaet Bremen.

Gage, D.A. et al., 1997. A new route for synthesis of dimethylsulphoniopropionate in marine algae. , 387(6636), pp.891–894.

Gall, M.. et al., 2001. Phytoplankton processes. Part 1: Community structure during the Southern Ocean Iron RElease Experiment (SOIREE). *Deep Sea Research Part II: Topical Studies in Oceanography*, 48(11-12), pp.2551–2570.

Geider, R.J. & La Roche, J., 1994. The role of iron in phytoplankton photosynthesis, and the potential for iron-limitation of primary productivity in the sea. *Photosynthesis research*, 39(3), pp.275–301.

Grabherr, M.G. et al., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology*, 29(7), pp.644–52.

Guillard, R.R. & Ryther, J.H., 1962. Studies of marine planktonic diatoms. I. Cyclotella nana Hustedt, and Detonula confervacea (cleve) Gran. *Canadian journal of microbiology*, 8, pp.229–39.

Haas, B.J. et al., 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature protocols*, 8(8), pp.1494–512.

Haas, B.J., Trinity RNA-Seq Assembly / Mailing Lists. Available at: http://sourceforge.net/p/trinityrnaseq/mailman/message/31622125/ [Accessed May 12, 2014].

Hackett, J.D. et al., 2007. Phylogenomic analysis supports the monophyly of cryptophytes and haptophytes and the association of rhizaria with chromalveolates. *Molecular biology and evolution*, 24(8), pp.1702–13.

Hall, J.A. & Safi, K., 2001. The impact of in situ Fe fertilisation on the microbial food web in the Southern Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography*, 48(11-12), pp.2591–2613.

Hamana, K. & Matsuzaki, S., 1982. Widespread occurrence of norspermidine and norspermine in eukaryotic algae. *Journal of biochemistry*, 91(4), pp.1321–8.

Hartnett, A. et al., 2012. Iron transport and storage in the coccolithophore: Emiliania huxleyi. *Metallomics : integrated biometal science*, 4(11), pp.1160–6.

Heinonen, J.K., 2001. *Biological Role of Inorganic Pyrophosphate*, Springer.

Henderiks, J., 2008. Coccolithophore size rules — Reconstructing ancient cell geometry and cellular calcite quota from fossil coccoliths. *Marine Micropaleontology*, 67(1-2), pp.143–154.

Hickel, W., 1984. Seston retention by Whatman GF/C glass-fiber filters. *Marine Ecology Progress Series*, 16(1-2), pp.185–191.

Van Hilst, C.M. & Smith, W.O., 2002. Photosynthesis/irradiance relationships in the Ross Sea, Antarctica, and their control by phytoplankton assemblage composition and environmental factors. *Marine Ecology Progress Series*, 226, pp.1–12.

Hockin, N.L. et al., 2012. The response of diatom central carbon metabolism to nitrogen starvation is different from that of green algae and higher plants. *Plant physiology*, 158(1), pp.299–312.

Hoffmann, L., Peeken, I. & Lochte, K., 2007. Effects of iron on the elemental stoichiometry during EIFEX and in the diatoms Fragilariopsis kerguelensis and Chaetoceros dichaeta. *Biogeosciences*.

Hoffmann, L.J. et al., 2006. Different reactions of Southern Ocean phytoplankton size classes to iron fertilization. *Limnology and oceanography*, 51(3), pp.1217–1229.

Hoppe, C.J.M. et al., 2013. Iron limitation modulates ocean acidification effects on southern ocean phytoplankton communities. E. V. Thuesen, ed. *PloS one*, 8(11), p.e79890.

Horton, P. et al., 2006. Protein Subcellular Localization Prediction with WoLF PSORT. In *Proceedings of the 4th Asia-Pacific Bioinformatics Conference*. Series on Advances in Bioinformatics and Computational Biology. Imperial College Press, pp. 39–48.

Hutchins, D.A. & Bruland, K.W., 1998. Iron-limited diatom growth and Si:N uptake ratios in a coastal upwelling regime. *Nature*, 393(6685), pp.561–564.

IPCC Working Group II, 2001. *IPCC Third Assessment Report: Climate Change 2001 (TAR) Working Group II: Impacts, Adaptation and Vulnerability.*,

Itoh, N. & Shinya, M., 1994. Seasonal evolution of bromomethanes from coralline algae (Corallinaceae) and its effect on atmospheric ozone. *Marine Chemistry*, 45(1-2), pp.95–103.

Jacques, J.-F. et al., 2006. RyhB small RNA modulates the free intracellular iron pool and is essential for normal growth during iron limitation in Escherichia coli. *Molecular microbiology*, 62(4), pp.1181–90.

Jeffery, B. et al., 2004. Amnesic shellfish poison. *Food and chemical toxicology : an international journal published for the British Industrial Biological Research Association*, 42(4), pp.545–57.

John, U. et al., 2010. Genomic characterisation of the ichthyotoxic prymnesiophyte Chrysochromulina polylepis, and the expression of polyketide synthase genes in synchronized cultures. *European Journal of Phycology*, 45(3), pp.215–229.

Kennett, J.P., 1977. Cenozoic evolution of Antarctic glaciation, the circum-Antarctic Ocean, and their impact on global paleoceanography. *Journal of Geophysical Research*, 82(27), pp.3843–3860.

Kettle, A.J. et al., 1999. A global database of sea surface dimethylsulfide (DMS) measurements and a procedure to predict sea surface DMS as a function of latitude, longitude, and month. *Global Biogeochemical Cycles*, 13(2), pp.399–444.

Kettle, A.J. & Andreae, M.O., 2000. Flux of dimethylsulfide from the oceans: A comparison of updated data sets and flux models. *Journal of Geophysical Research*, 105(D22), p.26793.

Krogh, A. et al., 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of molecular biology*, 305(3), pp.567–80.

Lagesen, K. et al., 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic acids research*, 35(9), pp.3100–8.

Lalli, C.M. & Parsons, T.R., 2004. *Biological Oceanography an Introduction*, Elsevier B.V.

Landry, M., Ondrusek, M., et al., 2000. Biological response to iron fertilization in the eastern equatorial Pacific (IronEx II). I. Microplankton community abundances and biomass. *Marine Ecology Progress Series*.

Landry, M., Constantinou, J., et al., 2000. Biological response to iron fertilization in the eastern equatorial Pacific (IronEx II). III. Dynamics of phytoplankton growth and microzooplankton grazing. *Marine Ecology Progress Series*.

Law, C. et al., 1998. Sulphur hexafluoride as a tracer of biogeochemical and physical processes in an open-ocean iron fertilisation experiment. *Deep-Sea Research Part II*.

Van Leeuwe, M. & Stefels, J., 1998. Effects of iron and light stress on the biochemical composition of Antarctic Phaeocystis sp.(Prymnesiophyceae). II. Pigment composition. *Journal of Phycology*, 34(3), pp.496–503.

Levasseur, M. et al., 2006. DMSP and DMS dynamics during a mesoscale iron fertilization experiment in the Northeast Pacific—Part I: Temporal and vertical distributions. *Deep Sea Research Part II: Topical Studies in Oceanography*, 53(20), pp.2353–2369.

Li, B. & Dewey, C.N., 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC bioinformatics*, 12(1), p.323.

Li, F., Orban, R. & Baker, B., 2012. SoMART: a web server for plant miRNA, tasiRNA and target gene analysis. *The Plant journal : for cell and molecular biology*, 70(5), pp.891–901.

Li, S. et al., 2006. Phylogenomic analysis identifies red algal genes of endosymbiotic origin in the chromalveolates. *Molecular biology and evolution*, 23(3), pp.663–74.

Liss, P.S. et al., 1994. Dimethyl sulphide and Phaeocystis: A review. *Journal of Marine Systems*, 5(1), pp.41–53.

Lommer, M. et al., 2012. Genome and low-iron response of an oceanic diatom adapted to chronic iron limitation. *Genome biology*, 13(7), p.R66.

Lorenzen, C., 1967. Determination of chlorophyll and pheopigments: spectrophotometric equations. *Limnology and Oceanography*, 12(2), pp.343–346.

Lorenzen, C. & Jeffrey, S., 1980. Determination of chlorophyll in seawater. *Unesco tech. pap. mar. sci*.

Maldonado, M. et al., 2001. Iron uptake and physiological response of phytoplankton during a mesoscale Southern Ocean Iron enrichment. *Limnology and …*, 46(7), pp.1802–1808.

Maldonado, M. et al., 2002. The effect of Fe and Cu on growth and domoic acid production by Pseudo-nitzschia multiseries and Pseudo-nitzschia australis. *Limnology and Oceanography*.

Maldonado, M.T. et al., 2005. Acquisition of iron bound to strong organic complexes, with different Fe binding groups and photochemical reactivities, by plankton communities in Fe-limited subantarctic waters. *Global Biogeochemical Cycles*, 19(4), p.n/a–n/a.

Marchetti, A. et al., 2012. Comparative metatranscriptomics identifies molecular bases for the physiological responses of phytoplankton to varying iron availability. *Proceedings of the National Academy of Sciences of the United States of America*, 109(6), pp.E317–25.

Marchetti, A. et al., 2009. Ferritin is used for iron storage in bloom-forming marine pennate diatoms. *Nature*, 457(7228), pp.467–70.

Marchetti, A. et al., 2008. Identification and assessment of domoic acid production in oceanic Pseudo-nitzschia (Bacillariophyceae) from iron-limited waters in the northeast subarctic Pacific. *Journal of Phycology*, 44(3), pp.650–661.

Marchetti, A. et al., 2006. Phytoplankton processes during a mesoscale iron enrichment in the NE subarctic Pacific: Part I—Biomass and assemblage. *Deep Sea Research Part II: Topical Studies in Oceanography*, 53(20), pp.2095–2113.

Marchetti, A. & Cassar, N., 2009. Diatom elemental and morphological changes in response to iron limitation: a brief review with potential paleoceanographic applications. *Geobiology*, 7(4), pp.419–31.

Martin, J., 1990. Glacial-interglacial CO2 change: The iron hypothesis. *Paleoceanography*.

Martin, J.H. et al., 1994. Testing the iron hypothesis in ecosystems of the equatorial Pacific Ocean. *Nature*, 371(6493), pp.123–129.

Martin, J.H. & Fitzwater, S.E., 1988. Iron deficiency limits phytoplankton growth in the north-east Pacific subarctic. *Nature*, 331(6154), pp.341–343.

Martin, J.H., Fitzwater, S.E. & Gordon, R.M., 1990. Iron deficiency limits phytoplankton growth in Antarctic waters. *Global Biogeochemical Cycles*, 4(1), pp.5–12.

Martin, P. et al., 2013. Iron fertilization enhanced net community production but not downward particle flux during the Southern Ocean iron fertilization experiment LOHAFEX. *Global Biogeochemical Cycles*.

Martinez-Garcia, A. et al., 2014. Iron Fertilization of the Subantarctic Ocean During the Last Ice Age. *Science*, 343(6177), pp.1347–1350.

Martínez-Garcia, A. et al., 2011. Southern Ocean dust-climate coupling over the past four million years. *Nature*, 476(7360), pp.312–5.

Martins, T.G., Computing and visualizing PCA in R | Thiago G. Martins on WordPress.com. Available at: http://tgmstat.wordpress.com/2013/11/28/computing-and-visualizing-pca-in-r/ [Accessed May 10, 2014].

Massé, E. & Gottesman, S., 2002. A small RNA regulates the expression of genes involved in iron metabolism in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, 99(7), pp.4620–5.

Maxwell, K. & Johnson, G.N., 2000. Chlorophyll fluorescence--a practical guide. *Journal of Experimental Botany*, 51(345), pp.659–668.

Mazzocchi, M.G. et al., 2009. A non-diatom plankton bloom controlled by copepod grazing and amphipod predation: Preliminary results from the LOHAFEX iron-fertilisation experiment. In *GLOBEC International NewsletterOctober*.

Medlin, L., 2009. Haptophyte algae (Haptophyta). In *The Timetree of Life*.

Medlin, L. & Zingone, A., 2007. A taxonomic review of the genus Phaeocystis. *Biogeochemistry*, 83(1-3), pp.3–18.

Merzouk, A. et al., 2006. DMSP and DMS dynamics during a mesoscale iron fertilization experiment in the Northeast Pacific–Part II: Biological cycling. *Deep Sea Research Part II: Topical Studies in Oceanography*, 53(20-22), pp.2370–2383.

Miller, E.P. et al., 2014. Surface-bound iron: a metal ion buffer in the marine brown alga Ectocarpus siliculosus? *Journal of experimental botany*, 65(2), pp.585–94.

Moisan, T.A. & Mitchell, B.G., 1999. Photophysiological acclimation of Phaeocystis antarctica Karsten under light limitation. *Limnology and oceanography*, 44(2), pp.247–258.

Moore, J.K. et al., 2000. The southern ocean at the Last Glacial Maximum: A strong sink for atmospheric carbon dioxide. *Global Biogeochemical Cycles*, 14(1), pp.455–475.

Moorthie, S., Mattocks, C.J. & Wright, C.F., 2011. Review of massively parallel DNA sequencing technologies. *The HUGO journal*, 5(1-4), pp.1–12.

Morey, J.S. et al., 2011. Transcriptomic response of the red tide dinoflagellate, Karenia brevis, to nitrogen and phosphorus depletion and addition. *BMC genomics*, 12(1), p.346.

Morrissey, J. & Bowler, C., 2012. Iron utilization in marine cyanobacteria and eukaryotic algae. *Frontiers in microbiology*, 3, p.43.

Mortlock, R.A. et al., 1991. Evidence for lower productivity in the Antarctic Ocean during the last glaciation. *Nature*, 351(6323), pp.220–223.

Mos, L., 2001. Domoic acid: a fascinating marine toxin. *Environmental Toxicology and Pharmacology*, 9(3), pp.79–85.

Moustafa, A., 2009. *Evolutionary and functional genomics of photosynthetic eukaryotes*. University of Iowa.

Moustafa, A. et al., 2009. Genomic footprints of a cryptic plastid endosymbiosis in diatoms. *Science (New York, N.Y.)*, 324(5935), pp.1724–6.

Moustafa, A. et al., 2010. Transcriptome profiling of a toxic dinoflagellate reveals a gene-rich protist and a potential impact on gene expression due to bacterial presence. *PloS one*, 5(3), p.e9688.

Moustafa, A., Reyes-Prieto, A. & Bhattacharya, D., 2008. Chlamydiae has contributed at least 55 genes to Plantae with predominantly plastid functions. R. DeSalle, ed. *PloS one*, 3(5), p.e2205.

Moxon, S. et al., 2008. A toolkit for analysing large-scale plant small RNA datasets. *Bioinformatics (Oxford, England)*, 24(19), pp.2252–3.

Murphy, E.R. & Payne, S.M., 2007. RyhB, an iron-responsive small RNA molecule, regulates Shigella dysenteriae virulence. *Infection and immunity*, 75(7), pp.3470–7.

Naito, K., Imai, I. & Nakahara, H., 2008. Complexation of iron by microbial siderophores and effects of iron chelates on the growth of marine microalgae causing red tides. *Phycological Research*, 56(1), pp.58–67.

National Center for Biotechnology Information, NCBI Taxonomy browser (Haptophyceae). Available at: http://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi [Accessed April 30, 2014].

Newton, J., 2002. Analysis of Phytoplankton Pigments: Determination of Chlorophyll a. Available at: http://courses.washington.edu/ocean220/misc/Chla_method_newton.pdf [Accessed December 10, 2013].

Nodop, A. et al., 2008. Transcript profiling reveals new insights into the acclimation of the mesophilic fresh-water cyanobacterium Synechococcus elongatus PCC 7942 to iron starvation. *Plant physiology*, 147(2), pp.747–63.

Norden-Krichmar, T.M. et al., 2011. Characterization of the small RNA transcriptome of the diatom, Thalassiosira pseudonana. I. Friedberg, ed. *PloS one*, 6(8), p.e22870.

Ohsawa, N. et al., 2001. Physiological function of bromoperoxidase in the red marine alga, Corallina pilulifera: production of bromoform as an allelochemical and the simultaneous elimination of hydrogen peroxide. *Phytochemistry*, 58(5), pp.683–692.

Olson, R.. et al., 2000. Effects of iron enrichment on phytoplankton in the Southern Ocean during late summer: active fluorescence and flow cytometric analyses. *Deep Sea Research Part II: Topical Studies in Oceanography*, 47(15-16), pp.3181–3200.

Osterman, A. & Overbeek, R., 2003. Missing genes in metabolic pathways: a comparative genomics approach. *Current Opinion in Chemical Biology*, 7(2), pp.238–251.

Oudot-Le Secq, M.-P. et al., 2007. Chloroplast genomes of the diatoms Phaeodactylum tricornutum and Thalassiosira pseudonana: comparison with other plastid genomes of the red lineage. *Molecular genetics and genomics : MGG*, 277(4), pp.427–39.

Parker, P., Displaying K-means Results. | Adventures in R on WordPress.com. Available at: http://adventuresinr.wordpress.com/2010/10/22/displaying-k-means-results/ [Accessed May 24, 2014].

Parra, G., Blanco, E. & Guigó, R., 2000. GeneID in Drosophila. *Genome research*, 10(4), pp.511–5.

Peperzak, L. et al., 2000. Growth and mortality of flagellates and non-flagellate cells of Phaeocystis globosa (Prymnesiophyceae). *Journal of Plankton Research*, 22(1), pp.107–120.

Peperzak, L. & Gäbler-Schwarz, S., 2012. Current knowledge of the life cycles of Phaeocystis globosa and Phaeocystis antarctica (Prymnesiophyceae). *Journal of Phycology*, 48(3), pp.514–517.

Petersen, T.N. et al., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature methods*, 8(10), pp.785–6.

Phaeocystis.org, Phaeocystis.org: Phaeocystis Research. Available at: http://www.phaeocystis.org/about.html [Accessed May 5, 2014].

Pollack, S., Vanderhoff, G. & Lasky, F., 1977. Iron removal from transferrin An experimental study. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 497(2), pp.481–487.

Porter, K.G. & Feig, Y.S., 1980. The use of DAPI for identifying and counting aquatic microflora. *Limnology and Oceanography*, 25(5), pp.943–948.

Powell, S. et al., 2012. eggNOG v3.0: orthologous groups covering 1133 organisms at 41 different taxonomic ranges. *Nucleic acids research*, 40(Database issue), pp.D284–9.

Prince, E.K., Irmer, F. & Pohnert, G., 2013. Domoic acid improves the competitive ability of Pseudo-nitzschia delicatissima against the diatom Skeletonema marinoi. *Marine drugs*, 11(7), pp.2398–412.

Punta, M. et al., 2012. The Pfam protein families database. *Nucleic acids research*, 40(Database issue), pp.D290–301.

R Core Team, 2013. R: A Language and Environment for Statistical Computing.

Raven, J.A., 2013. Iron acquisition and allocation in stramenopile algae. *Journal of experimental botany*, 64(8), pp.2119–27.

Read, B.A. et al., 2013. Pan genome of the phytoplankton Emiliania underpins its global distribution. *Nature*, 499(7457), pp.209–13.

Redfield, A., 1934. On the proportions of organic derivations in sea water and their relation to the composition of plankton. , pp.177 – 192.

Reyes-Prieto, A., Moustafa, A. & Bhattacharya, D., 2008. Multiple genes of apparent algal origin suggest ciliates may once have been photosynthetic. *Current biology : CB*, 18(13), pp.956–62.

Reynolds, C.S., Oliver, R.L. & Walsby, A.E., 1987. Cyanobacterial dominance: The role of buoyancy regulation in dynamic lake environments. *New Zealand Journal of Marine and Freshwater Research*, 21(3), pp.379–390.

Riegger, L. & Robinson, D., 1998. Photoinduction of UV-absorbing compounds in Antarctic diatoms and Phaeocystis antarctica. *Oceanographic Literature Review*, 45(8).

Riffenburgh, B., 2007. *Encyclopedia of the Antarctic, Volume 1* B. Riffenburgh, ed., Taylor & Francis.

Rousseau, V. et al., 2007. The life cycle of Phaeocystis: state of knowledge and presumptive role in ecology. *Biogeochemistry*, 83(1-3), pp.29–47.

Rubinelli, P. et al., 2002. Cadmium- and iron-stress-inducible gene expression in the green alga Chlamydomonas reinhardtii: evidence for H43 protein function in iron assimilation. *Planta*, 215(1), pp.1–13.

Rue, E. & Bruland, K., 2001. Domoic acid binds iron and copper: a possible role for the toxin produced by the marine diatom Pseudo-nitzschia. *Marine Chemistry*, 76(1-2), pp.127–134.

Sáez, A.G. et al., 2004. A review of the phylogeny of the Haptophyta. In H. R. Thierstein & J. R. Young, eds. *Coccolithophores: from molecular processes to global impact*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Sánchez Puerta, M.V., Bachvaroff, T.R. & Delwiche, C.F., 2005. The complete plastid genome sequence of the haptophyte Emiliania huxleyi: a comparison to other plastid genomes. *DNA research : an international journal for rapid publication of reports on genes and genomes*, 12(2), pp.151–6.

Santini, S. et al., 2013. Genome of Phaeocystis globosa virus PgV-16T highlights the common ancestry of the largest known DNA viruses infecting eukaryotes. *Proceedings of the National Academy of Sciences of the United States of America*, 110(26), pp.10800–5.

Sarmiento, J.L. et al., 1998. Simulated response of the ocean carbon cycle to anthropogenic climate warming. , 393(6682), pp.245–249.

Schoemann, V. et al., 2005. Phaeocystis blooms in the global ocean and their controlling mechanisms: a review. *Journal of Sea Research*, 53(1), pp.43–66.

Sieburth, J.M., Smetacek, V. & Lenz, J., 1978. Pelagic ecosystem structure: Heterotrophic compartments of the plankton and their relationship to plankton size fractions. *Limnology and Oceanography*, 23(6), pp.1256–1263.

Sigman, D., Hain, M. & Haug, G., 2010. The polar ocean and glacial cycles in atmospheric CO2 concentration. *Nature*.

Sigman, D.M. & Hain, M.P., 2012. The Biological Productivity of the Ocean: Section 2. *Nature Education Knowledge*, 3(10), p.20.

Silver, M.W. et al., 2010. Toxic diatoms and domoic acid in natural and iron enriched waters of the oceanic Pacific. *Proceedings of the National Academy of Sciences of the United States of America*, 107(48), pp.20762–7.

Smetacek, V., 2009. A non-diatom bloom controlled by copepod grazing and amphipod predation: Preliminary results from the LOHAFEX iron fertilization experiment. In *GeoBiosphere Science Centre, Lund University, Sweden*.

Smetacek, V., 2001. A watery arms race. *Nature*, 411(6839), p.745.

Smetacek, V. et al., 2012. Deep carbon export from a Southern Ocean iron-fertilized diatom bloom. *Nature*, 487(7407), pp.313–9.

Smetacek, V. et al., 1997. Ecology and biogeochemistry of the antarctic circumpolar current during austral spring: a summary of southern ocean JGOFS cruise ANT X/6 of R.V. Polarstern. *Deep Sea Research Part II: Topical Studies in Oceanography*, 44(1-2), pp.1–21.

Smetacek, V., Assmy, P. & Henjes, J., 2004. The role of grazing in structuring Southern Ocean pelagic ecosystems and biogeochemical cycles. *Antarctic Science*, 16(4), pp.541–558.

Smetacek, V. & Naqvi, S.W.A., 2008. The next generation of iron fertilization experiments in the Southern Ocean. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, 366(1882), pp.3947–67.

Smith, D.R. et al., 2014. Massive difference in synonymous substitution rates among mitochondrial, plastid, and nuclear genes of Phaeocystis algae. *Molecular phylogenetics and evolution*, 71, pp.36–40.

Smith, D.R. & Lee, R.W., 2014. A Plastid without a Genome: Evidence from the Nonphotosynthetic Green Algal Genus Polytomella. *Plant physiology*, 164(4), pp.1812–9.

Smith, W. et al., 1998. Growth dynamics of Phaeocystis antarctica-dominated plankton assemblages from the Ross Sea. *Marine Ecology Progress Series*, 168, pp.229–244.

Smith, W.O. et al., 2003. The temporal dynamics of the flagellated and colonial stages of Phaeocystis antarctica in the Ross Sea. *Deep Sea Research Part II: Topical Studies in Oceanography*, 50(3), pp.605–617.

Stefels, J. et al., 2007. Environmental constraints on the production and removal of the climatically active gas dimethylsulphide (DMS) and implications for ecosystem modelling. *Biogeochemistry*, 83(1-3), pp.245–275.

Stewart, R.H., 2009. *Introduction to Physical Oceanography*, Orange Grove Texts Plus.

Stoll, H., Shimizu, N. & Langer, G., 2009. Coccolith B/Ca ratios using SIMS ion probe analysis. *EGU General Assembly Conference Abstracts*, 11, p.6729.

Strong, A. et al., 2009. Ocean fertilization: time to move on. *Nature*, 461(7262), pp.347–8.

Strong, A., Cullen, J. & Chisholm, S., 2009. Ocean fertilization: Science, policy, and commerce. *Oceanography*, 22(3), pp.236–261.

Strzepek, R.F. et al., 2011. Adaptive strategies by Southern Ocean phytoplankton to lessen iron limitation: Uptake of organically complexed iron and reduced cellular iron requirements. *Limnology and Oceanography*, 56(6), pp.1983–2002.

Summers, P.S. et al., 1998. Identification and Stereospecificity of the First Three Enzymes of 3-Dimethylsulfoniopropionate Biosynthesis in a Chlorophyte Alga. *Plant Physiology*, 116(1), p.369.

Sunda, W.G. & Huntsman, S.A., 1995. Iron uptake and growth limitation in oceanic and coastal phytoplankton. *Marine Chemistry*, 50(1-4), pp.189–206.

Sutak, R. et al., 2012. A comparative study of iron uptake mechanisms in marine microalgae: iron binding at the cell surface is a critical step. *Plant physiology*, 160(4), pp.2271–84.

Suttle, C.A., 2007. Marine viruses--major players in the global ecosystem. *Nature reviews. Microbiology*, 5(10), pp.801–12.

Tagliabue, A. et al., 2014. Surface-water iron supplies in the Southern Ocean sustained by deep winter mixing. *Nature Geoscience*, 7(4), pp.314–320.

Takahashi, T. et al., 2009. Climatological mean and decadal change in surface ocean pCO2, and net sea–air CO2 flux over the global oceans. *Deep Sea Research Part II: Topical Studies in Oceanography*, 56(8-10), pp.554–577.

Talley, L.D. et al., 2011. *Descriptive Physical Oceanography: An Introduction* 6th ed., Academic Press.

Tang, K.W. et al., 2009. Survival and recovery of Phaeocystis antarctica (Prymnesiophyceae) from prolonged darkness and freezing. *Proceedings. Biological sciences / The Royal Society*, 276(1654), pp.81–90.

Ternois, Y. et al., 1997. Evaluation of long-chain alkenones as paleo-temperature indicators in the Mediterranean Sea. *Deep Sea Research Part I: Oceanographic Research Papers*, 44(2), pp.271–286.

Thamatrakoln, K. et al., 2012. Whole-genome expression analysis reveals a role for death-related genes in stress acclimation of the diatom Thalassiosira pseudonana. *Environmental microbiology*, 14(1), pp.67–81.

The UniProt Consortium, 2014. Activities at the Universal Protein Resource (UniProt). *Nucleic acids research*, 42(1), pp.D191–8.

Thiele, S. et al., 2012. Microbial community response during the iron fertilization experiment LOHAFEX. *Applied and environmental microbiology*, 78(24), pp.8803–12.

Tilman, D., 1994. Competition and Biodiversity in Spatially Structured Habitats. *Ecology*, 75(1), p.2.

Tomczak, M. & Godfrey, J.S., 2004. *Regional Oceanography: An Introduction* 2nd ed., Daya Publishing House.

Trick, C.G. et al., 2010. Iron enrichment stimulates toxic diatom production in high-nitrate, low-chlorophyll areas. *Proceedings of the National Academy of Sciences of the United States of America*, 107(13), pp.5887–92.

Tsuda, A. et al., 2003. A mesoscale iron enrichment in the western subarctic Pacific induces a large centric diatom bloom. *Science (New York, N.Y.)*, 300(5621), pp.958–61.

Tsuda, A. et al., 2007. Evidence for the grazing hypothesis: Grazing reduces phytoplankton responses of the HNLC ecosystem to iron enrichment in the western subarctic pacific (SEEDS II). *Journal of Oceanography*, 63(6), pp.983–994.

Turner, S.M. et al., 1996. Increased dimethyl sulphide concentrations in sea water from in situ iron enrichment. *Nature*, 383(6600), pp.513–517.

Turner, S.M. et al., 2004. Iron-induced changes in oceanic sulfur biogeochemistry. *Geophysical Research Letters*, 31(14), p.L14307.

Utermöhl, H., 1958. Zur Vervollkommnung der quantitativen Phytoplankton-Methodik. *Mitt. int. Ver. theor. angew. Limnol.*, 9, pp.1 – 38.

Vaillancourt, R.D. et al., 2003. Primary productivity and in situ quantum yields in the Ross Sea and Pacific Sector of the Antarctic Circumpolar Current. *Deep Sea Research Part II: Topical Studies in Oceanography*, 50(3-4), pp.559–578.

Vaulot, D. et al., 1994. Morphology, ploidy, pigment composition, and genome size of cultured strains of Phaeocystis (Prymnesiophyceae). *Journal of Phycology*, 30(6), pp.1022–1035.

Verity, P.G. et al., 2007. Current understanding of Phaeocystis ecology and biogeochemistry, and perspectives for future research. *Biogeochemistry*, 83(1-3), pp.311–330.

Wang, X.J. et al., 2009. Regulation of phytoplankton carbon to chlorophyll ratio by light, nutrients and temperature in the Equatorial Pacific Ocean: a basin-scale model. *Biogeosciences*, 6(3), pp.391–404.

Wang, Z., Gerstein, M. & Snyder, M., 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews. Genetics*, 10(1), pp.57–63.

Watson, A., Liss, P. & Duce, R., 1991. Design of a small-scale in situ iron fertilization experiment. *Limnology and Oceanography*.

Wells, M.L. et al., 2005. Domoic acid: The synergy of iron, copper, and the toxicity of diatoms. *Limnology and Oceanography*, 50(6), pp.1908–1917.

Wever, R. et al., 1993. Bromoperoxidases: Their Role in the Formation of HOBr and Bromoform by Seaweeds. In R. Oremland, ed. *Biogeochemistry of Global Change SE - 43*. Springer US, pp. 811–824.

Whitney, F.A., Crawford, D.W. & Yoshimura, T., 2005. The uptake and export of silicon and nitrogen in HNLC waters of the NE Pacific Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography*, 52(7-8), pp.1055–1067.

Wolf, C. et al., 2012. Eukaryotic picoplankton composition and succession during the iron fertilization experiment LOHAFEX in the Southern Ocean. In *IPY 2012 Conference, Montréal, Canada*.

Wolf, C. et al., 2013. Regional variability in eukaryotic protist communities in the Amundsen Sea. *Antarctic Science*, 25(06), pp.1–11.

Wolfe, G. V., Steinke, M. & Kirst, G.O., 1997. Grazing-activated chemical defence in a unicellular marine alga. , 387(6636), pp.894–897.

Yamada, T. et al., 2011. iPath2.0: interactive pathway explorer. *Nucleic acids research*, 39(Web Server issue), pp.W412–5.

Yang, I. et al., 2011. Growth- and nutrient-dependent gene expression in the toxigenic marine dinoflagellate Alexandrium minutum. *Harmful Algae*, 12, pp.55–69.

Yoon, H.S. et al., 2004. A molecular timeline for the origin of photosynthetic eukaryotes. *Molecular biology and evolution*, 21(5), pp.809–18.

Yoon, H.S., Hackett, J.D. & Bhattacharya, D., 2002. A single origin of the peridinin- and fucoxanthin-containing plastids in dinoflagellates through tertiary endosymbiosis. *Proceedings of the National Academy of Sciences of the United States of America*, 99(18), pp.11724–9.

Zhi-Liang, H., Bao, J. & Reecy, J., 2008. CateGOrizer: a web-based program to batch analyze gene ontology classification categories. *Online J Bioinformatics*, 9(2), pp.108–112.

Zingone, A. et al., 1999. Morphological and genetic characterization of Phaeocystis cordata and P. Jahnii (Prymnesiophyceae), two new species from the Mediterranean Sea. *Journal of Phycology*, 35(6), pp.1322–1337.

Zingone, A. et al., 2011. Morphological characterization of Phaeocystis antarctica (Prymnesiophyceae). *Phycologia*, 50(6), pp.650–660.

Zingone, A. et al., 2006. Potentially toxic and harmful microalgae from coastal waters of the Campania region (Tyrrhenian Sea, Mediterranean Sea). *Harmful Algae*, 5(3), pp.321–337.