Thermal adaptation of *Thalassiosira pseudonana* using experimental evolution approaches

A thesis submitted to the School of School of Environmental Sciences of the University of East Anglia in partial fulfilment of the requirements for the degree of Doctor of Philosophy

By Katrin Schmidt

June 2016

Quote

The creatures which can stand the 'storm of stress' of [...] changes which occur in the environment, by undergoing modification of [...] the structures which they get congenitally - these creatures will live; while those which cannot, will not.

Baldwin (1896), first seen in Chevin (2013)

Für meine Familie

Abstract

Diatoms contribute about 50% of global primary production and are on of the most diverse phytoplankton groups. Additionally, they form the basis of most marine food webs and play an important role in elemental cycles such as carbon and silica. Global warming impacts the diversity and productivity of marine ecosystems as temperature is considered a strong selecting agent underpinning global diversity patterns of marine phytoplankton. In order to gain insights into diatom distribution and diversity in the Atlantic Ocean, we analysed 18S rDNA ribotypes over a broad spatial scale from the Fram Strait to the South Atlantic. Diversity patterns were related to environmental metadata in order to identify main drivers. Our results indicate that salinity had a negative effect on diatom diversity in the Fram Strait transect with stations showing low diversity at high salinities. In contrast, diatom diversity in the Atlantic Ocean was negatively correlated to temperature with high temperature showing low diatom diversity. The order of Coscinodiscales showed a, formerly unknown, cosmopolitan distribution and was the overall most abundant species. With this study we provided an updated estimate of diatom distribution and diversity in the Atlantic Ocean.

Phytoplankton physiology is highly temperature dependent and despite the importance of temperature as a major driver of marine phytoplankton evolution, the molecular mechanisms of adaptive evolution under temperature selection are largely unknown but instrumental for predicting how marine phytoplankton will respond to a changing ocean. Here we provide evidence, based on experimental evolution experiments with the marine model diatom *Thalassisoria pseudonana* that thermal tolerance can rapidly evolve within 300 generations. Our results indicate that upper and lower temperature limits were fixed, however temperature optima

for growth shifted towards the selection temperature. Furthermore, temperature had a significant impact on average cell diameter, bSi content and cellular stoichiometry (C:N:P).

Physiological adaptation to high temperature was underpinned by differential expression of genes related to protein metabolism (protein binding and folding), and down-regulation of mismatch repair mechanisms potentially causing a high number of SNPs in the genome. Furthermore, several transposable elements showed strong, temperature specific up-regulation suggesting epigenetic enabled genome plasticity. Our results highlight the relation of adaptive pheno- and genotypes driven by temperature selection. This knowledge is key to our understanding of how the environment shapes the evolution of microbes and the biogeochemical processes they drive.

Acknowledgment

Four years of my journey in getting my PhD have come to an end and I look back with gratitude and happiness. I really enjoyed my time at UEA and the School of Environmental Sciences and owe much of the success of this thesis to my primary supervisor Thomas Mock. He pushed hard and showed me how to reach my potential. For his constant support and provided opportunities I am very grateful. Thank you for showing me the way to become a good scientist. I'd also like to thank my secondary supervisor Vincent Moulton for his bioinformatical support and feedback. Thank you School of Environmental Sciences for your generous financial support allowing me to dessiminate my research at several international conferences. Thank you Klaus Valentin for making it possible to join two research cruises on board Polarstern. Thank you ARK 27-1 for the amazing time in the Arctic Ocean. Michael Ginzburg, you were the best cruise buddy ever! ANT 29-1 was blast because of you.

The last four years would have been only half as much fun if it wasn't for all the great people I've met along the way. Thank you former (Jan, Amy, Andrew, Ana, Martin, Martin, Annemarie) and current members (Andrew, Krizstina, Amanda, Rob, Irina, Nigel) of the Mock lab. I greatly enjoyed working along side all of you. A special thanks to Rob Utting! Without your technical and practical support, most of the experiments would have not been possible. Thank you thesis-feeback-committee (Jan Strauss, Amy Kirkham, Andrew Toseland and Cock van Oosterhout) for your helpful comments on several drafts of this document. Cock van Oosterhout, thank you very much for your fruitful discussions of my experimental evolution results and your uplifting words. I enjoyed our discussions very much!

You guys in the office (Moritz, Alison, Emma, Raf, Cansu, Imke, Penelope, Anna, Nat, Ollie L., Ollie, Max, Barney, Krizstina, Amanda, Clare), it's been an a great 4 years because of you. Thank

you for the moral and cookie support. Special thanks to Ollie L, Ollie and Max for the table-tennis and squash matches. Thank you Alba, Laia, Fanchon and Elena for the distractions from writing and for keeping my spirits up.

Last but not least I'd like to thank my family who had to put up with the roller coaster called PhD studentship. Thank you for having my back the last years. I would have not been able to reach my goal without your great moral and financial support. This thesis is also in loving memory of Marita, Oma Hannelore and Erich who unfortunately passed away during my PhD.

Auf deutsch:

Zu guter Letzt, moechte ich mich bei meiner Familie bedanken die mich waehrend meines turbulent PhDs unterstuetzt haben. Vielen Dank fuer eure Rueckendenkung in den letzten Jahren. Ich haette es ohne eure moralische und finanzielle Unterstuetzung nicht geschafft mein Ziel zu erreichen. Diese Doktorarbeit ist auch in gedenken an Marita, Oma Hannelore und Erich die leider waehrend meiner Doktorarbeit verstorben sind.

Contents

Abstract **II**

Acknowledgement **IX**

Contents **XVII**

List of tables **XIX**

List of figures **1**

1 General Introduction **3**

1.1 The changing ocean environment 3

1.1.1 Effects of ocean warming on marine phytoplankton 4

1.1.1.1 Biological pump 4

1.1.1.2 Stratification 6

1.1.1.3 Distribution shifts 8

1.1.1.4 Changes in communities and food webs 9

1.1.1.5 Photosynthesis 11

1.1.1.6 Stoichiometry and size 13

1.2 The role of phenotypic plasticity and epigenetics in experimental evolution with phytoplankton 14

1.2.1 Epigenetic control of transposable elements 17

1.3 Diatoms 20

1.3.1 Endosymbiont theory 23

1.3.2 Thalassiosira - *Thalassiosira pseudonana* (Hustedt) Hasle et Heimdal . . 24

1.4 Thesis outline and aims 26

2 Sea of change - Diatom ecology in the Atlantic Ocean 29

2.1 Introduction 29

2.2 Material and Methods 31

2.2.1 Field campaign 31

2.2.1.1 Sampling 34

2.2.2 DNA extractions 34

2.2.3 DNA sequencing 35

2.2.4 18S sequence analysis 36

2.2.4.1 Building of reference tree 36

2.2.4.2 18S copy number normalization of reads 36

2.2.5 Statistical analysis 37

2.2.5.1 Co-correlation of environmental factors 38

2.2.5.2 Diatom abundance in relation to environmental factors 39

2.2.5.3 Richness in relation to environmental factors 40

2.2.5.4 Bray-Curtis dissimilarity in relation to environmental factors and spatial distance 40

2.3 Results . . . 41
2.3.1 Polar transect . . . 43
2.3.1.1 Metadata . . . 43
2.3.1.2 18S analysis . . . 47
2.3.2 Atlantic Section . . . 55
2.3.2.1 Metadata . . . 55
2.3.2.2 18S analysis . . . 60
2.4 Discussion . . . 66
2.5 Summary and conclusion . . . 72

3 Phenotypic outcomes of adaptive evolution to temperature in *Thalassiosira pseudonana* 73

3.1 Introduction . . . 73
3.2 Material and Methods . . . 76
3.2.1 General experimental setup . . . 76
3.2.2 Cell cultivation and transfers . . . 77
3.2.3 Sampling and extractions . . . 78
3.2.3.1 Particulate organic carbon and nitrogen (POC/N) . . . 78
3.2.3.2 Particulate organic phosphate (POP) . . . 79
3.2.3.3 Biogenic silica (bSi) . . . 79
3.2.3.4 Chlorophyll a . . . 80
3.2.3.5 Statistical analysis of stoichiometric results . . . 80
3.2.3.6 Light microscopy and scanning electron microscopy . . . 81
3.2.4 Reciprocal transplant assays . . . 82

3.2.4.1 Statistical analysis of transplant assay 82
3.2.5 Temperature response curves 83
3.2.5.1 Temperature response curves - Modelling 83
3.2.5.2 Statistical analysis of temperature response curves 84
3.3 Results 85
3.3.1 Physiological responses to temperature changes 85
3.3.1.1 Growth and photosynthesis 85
3.3.1.2 Temperature response curves (TRC) 90
3.3.2 Cellular composition of *Thalasssiosira pseudonana* under high and low temperature selection 92
3.4 Discussion 99
3.4.1 Physiological responses to high and low temperature selection 99
3.4.2 Elemental and pigment changes under temperature selection 102
3.5 Summary and conclusion 105

4 Molecular underpinnings of temperature adaptation in *Thalassiosira pseudonana* 107

4.1 Introduction 107
4.2 Material and Methods 112
4.2.1 Transcriptome analysis 112
4.2.2 RNA sampling and extraction 112
4.2.2.1 Library preparation and sequencing 112
4.2.2.2 Sequence analysis 112
4.2.2.3 Differential gene expression analysis 114
4.2.2.4 Multidimensional scaling of transcriptomes 114

4.2.2.5 Gene ontology (GO) analysis of differentially expressed genes 115
4.2.2.6 Pathway analysis 115
4.2.3 Genome resequencing of high temperature replicate 116
4.2.3.1 DNA sampling and extraction 116
4.2.3.2 Library preparation and sequencing 116
4.2.3.3 Single nucleotide polymorphism (SNP) analysis of high temperature genotype 117
4.2.3.4 GO analysis of genes with SNPs 117
4.3 Results 118
4.3.1 Transcriptome sequence analysis 118
4.3.2 Differential gene expression of low and high temperature cultures 118
4.3.2.1 GO analysis 124
4.3.2.2 InterPro Domain analysis - Overrepresented domains 131
4.3.2.3 Pathway analysis 134
4.3.2.4 Transposable elements (TEs) 140
4.3.2.5 Genome resequencing - SNP analysis of high temperature replicate 142
4.4 Discussion 146
4.5 Summary and conclusion 151

5 General discussion **153**

5.1 Summary of main results 153
5.2 Discussion 154
5.2.1 Broader impact of results 156

5.3 Future work 158

5.3.1 From genomes to biomes 159

6 Appendix **161**

6.1 Supplementary information for chapter 2 161

6.2 Supplementary information for chapter 3 163

6.2.1 SeaChem Marine SaltTM 163

6.2.2 Reciprocal transplant assay growth curves 164

6.2.3 Chlorophyll a calibration curve 165

6.2.4 Temperature response curve of the 32°C selection lines at two different light intensities 166

6.3 Supplementary information for chapter 4 166

6.4 Publication output 166

6.4.1 A novel cost effective and high-throughput isolation and identification method for marine microalgae 166

6.4.2 The role of phenotypic plasticity and epigenetics in experimental evolution with phytoplankton 182

6.4.3 A cell wall protein underlies cell size in centric diatoms 184

7 References **187**

List of Tables

2.2.1 Generalized Variance Inflation Factors for polar transect environmental covariates. 38

2.2.2 Generalized Variance Inflation Factors for the Atlantic transect environmental covariates. 39

2.3.1 Bacillariphyta order summary 42

2.3.2 Metadata for Polar transect 45

2.3.3 Metadata for Atlantic transect 57

3.3.1 Two-way ANOVA output of reciprocal transplant assay analysis 89

3.3.2 AIC values of model fits to the temperature response curve of control lines 91

4.3.2 Transcriptome sequencing read output throughout the analysis pipeline. 119

4.3.1 Number of differentially expressed genes in low and high temperature cultures. . 122

4.3.3 Overrepresented InterPro domains of down-regulated genes at T1 9°C (50 generations) 131

4.3.4 Overrepresented InterPro domains of up-regulated genes after 50 generations at 9°C 132

4.3.5 Overrepresented InterPro domains of down-regulated genes at TE 9°C 133

4.3.6 Overrepresented InterPro domains of up-regulated genes at TE 9°C 134

4.3.7 Differential expression of TEs for each temperature treatment 141

4.3.8 Differential expression of DNA-specific methyltransferases for all three selection temperatures 142

4.3.9 Overview of variants generated by SNPs per chromosome. 143

4.3.10 Overview of missense, nonsense and silent SNPs 144

6.2.1 SeaChem Marine SaltTM composition 164

6.4.1 Overrepresented InterPro domains of down-regulated genes at T1 9°C (50 generations) 168

6.4.2 Overrepresented InterPro domains of up-regulated genes after 50 generations at 9°C 170

6.4.3 Overrepresented InterPro domains of down-regulated genes at TE 9°C 172

6.4.4 Overrepresented InterPro domains of up-regulated genes at TE 9°C 174

6.4.5 Overrepresented InterPro domains of down-regulated genes at TE 32°C 177

6.4.6 Overrepresented InterPro domains of up-regulated genes at TE 32°C 179

List of Figures

1.1.1 Biological pump. 5

1.1.2 Expected changes to water column stratification under global warming for high and low latitudes. 8

1.2.1 Species and genotype sorting in mixed phytoplankton communities. 15

1.3.1 Schematic representation of the valve of a centric diatom 21

1.3.2 Haeckel drawings of diatom species, indicating their beautiful silicified cell wall. 22

1.3.3 Schematic representation of the endosymbiont theory 24

1.3.4 Scanning electron micrograph of *Thalassiosira pseudonana* 25

2.1.1 Overview of currently known distribution and abundance of Bacillariophyta in the Atlantic 30

2.2.1 Map of sampled stations during three field campaigns 33

2.2.2 18S copy number normalization 37

2.2.3 Stylisted RDA plot. 40

2.3.1 Temperature and Salinity profiles of Fram Strait transect 44

2.3.2 Diatom order richness in correlation with temperature. 47

2.3.3 Diatom order richness in correlation to silicate concentrations. 48

2.3.4 Relative diatom order abundance by station/longitude. 49

2.3.5 Beta diversity dendrogram of polar transect. 50

2.3.6 Correlation of environmental data 51

2.3.7 CAP of diatom orders and environmental variables for polar section 52

2.3.8 Correlation of diatom orders with temperature. 53

2.3.9 Correlation of diatom beta-diversity with distance. 54

2.3.10 Temperature and Salinity profiles of the Atlantic transect 55

2.3.11 Alpha diversity in correlation with environmental data 60

2.3.12 Relative abundance of diatom orders across the Atlantic stations ordered by station/latitude. 61

2.3.13 Beta diversity dendrogram for the Atlantic transect. 62

2.3.14 Co-correlation of environmental data for the Atlantic transect. 63

2.3.15 RDA of Atlantic section 64

2.3.16 Correlation of diatom orders with temperature. 65

2.3.17 The bray-curtis index as a function of distance. 66

3.2.1 Experimental setup of reciprocal transplant assays. 9°C selection cultures are assayed at 9°C, 22°C and 32°C. The same is done for 22°C and 32°C selection cultures. 82

3.3.1 Boxplot of growth rates at the beginning and end of the experiment for control cultures at selection temperatures 85

3.3.2 Overview of generations produced between each transfer for all three experimental setups 86

3.3.3 Fv/Fm of selection cultures over the course of the experiment 87

3.3.4 Average diameter trends over the course of the experiment for all selection lines . 88

3.3.5 Boxplot of growth rates (μ/μ_{max}) of each temperature selection line (9°C (blue), 22°C (green), 32°C (red)), after 300 generations, at assay temperatures. A two-way ANOVA showed that there was a statistically significant interaction between the effect of selection and assay temperature on growth rate, (F(4,40)=470.8 (p-value $<$0.001)) (Table 3.3.1). A post-hoc Tukey HSD test was used to test for significant differences (p-value $<$0.05) between growth rates of each selection line at the assay temperatures. Small letters indicate significant differences between boxes. Error bars indicate one standard deviation. 89

3.3.6 Tested temperature response models fitted to temperature response curve measurments of control lines 90

3.3.7 Correlation between temperature optimum and maximum estimates and selection temperature. 91

3.3.8 TRC models for low temperature selection lines 1-4 (larger phenotype) and line 5 (smaller phenotype) after 300 generations of temperature selection. 92

3.3.9 Cellular carbon, nitrogen, biogenic silica and chlorophyll a content per cell for all three selection lines. Control cultures in green, high temperature and low temperature treatments in red and blue respectively. Thick black lines in the box plot represent mean while thin black lines respond to the median. Small letters next to the box indicate significant differences to other boxes. Boxes sharing the same letter(s) are not significantly different, while different letters indicate significant differences. Values were significantly different at a p-value cut-off of 0.05. 94

3.3.10 Cellular carbon, nitrogen, biogenic silica and chlorophyll a content per cellvolume for all three selection lines. Control cultures in green, high temperature and low temperature treatments in red and blue respectively. Thick black lines in the box plot represent mean while thin black lines respond to the median. Small letters next to the box indicate significant differences to other boxes. Boxes sharing the same letter(s) are not significantly different, while different letters indicate significant differences. Values were significantly different at a p-value cut-off of 0.05. 96

3.3.11 Carbon to nitrogen, carbon to biogenic silica, carbon to chlorophyll a and nitrogen to phosphate ratios for all three temperature treatments after 300 generations of temperature selection. Control cultures in green, high temperature and low temperature treatments in red and blue, respectively. Thick black lines in the box plot represent median. Small letters next to the box indicate significant differences to other boxes. Boxes sharing the same letter(s) are not significantly different, while different letters indicate significant differences. Values were significantly different at a p-value cut-off of 0.05. 98

4.2.1 Work flow of bioinformatic analysis of transcriptome profiles. 113

4.3.1 Multidimensional scaling of low, high and control temperature transcriptomes . . 122

4.3.2 Transcriptome heatmap for low temperature selection lines and two sampling time points 123

4.3.3 Transcriptome heatmap for high temperature selection lines and sampling time points 124

4.3.4 ReVIGO GO-term scatterplots for low temperature cultures after 50 generations. 126

4.3.5 ReVIGO GO-term scatterplots for low temperature cultures after 300 generations. 128

4.3.6 ReVIGO GO-term scatterplots for high temperature cultures after 300 generations. 130

4.3.7 Heatmap of expressed genes related to biosynthesis of fatty acids scaled to their RPKM values for all three temperature treatments 135

4.3.8 Heatmap of expressed genes involved in desaturation fatty acids scaled to RPKM values for all three temperature treatments 136

4.3.9 Heatmap of expressed genes part of carotenoid synthesis scaled to RPKM values for all three temperature treatments 137

4.3.10 Heatmap of expressed genes part of porphyrin and chlorophyll metabolism scaled to RPKM values for all three temperature treatments 138

4.3.11 Heatmap of expressed ribosomal genes scaled to RPKM values for all three temperature treatments 139

4.3.12 Heatmap of expressed genes part of the mismatch repair pathway scaled to RPKM values for all three temperature treatments 140

4.3.13 Summary of SNPs per genomic region 144

4.3.14 Treemap of overrepresented GO-terms (BH adjusted p-value<0.05) for all genes with SNPs 145

4.3.15 Treemap of overrepresented GO-terms (BH adjusted p-value<0.05) for chromosome 14 genes with SNPs 145

4.3.16 Overview of amino acid changes caused by SNPs 146

6.1.1 Correlation between richness and salinity for Fram Strait transect. 162

6.1.2 RDA for the Fram Strait transect including salinity instead of temperature 162

6.2.1 Growth curves of selection cultures during reciprocal transplant assay 165

6.2.2 Calibration curve for chlorophyll a concentration calculations. 165

6.2.3 Temperature response curve of the high temperature replicate at two different light intensities. 166

Chapter 1

General Introduction

1.1 The changing ocean environment

Changing and fluctuating (biotic and abiotic) environments have been the driving force for adaptive evolution over the past billion years, leading from the first simple single cell to complex multicellular organisms. With the ongoing climate change, conditions are changing at an unprecedented rate (Doney *et al.*, 2012) and it is unclear how the ocean biota will respond and what the future ocean will look like. Since the industrialisation, carbon dioxide (CO_2) and other green house gas concentrations have increased. This increase had a severe impact on air and sea surface temperatures as water has a greater heat inertia than air, acting as a heat sink. Hence, sea surface temperatures have increased in recent years by on average 0.61°C (IPCC, 2014), and by almost twice as much in the Arctic (Graversen *et al.*, 2008) - known as "Arctic amplification".

In addition to increases in sea surface temperature, oceans are getting more acidic (ocean acidification) due to the stabilisation of the CO_2 equilibrium with the atmosphere. The acidification of the oceans will decrease the carbonate saturation state and cause enhanced calcium carbonate dissolution with severe consequences for calcifying organisms (Doney *et al.*, 2009) such as coccolitophores or pteropods.

In the following sections we will focus mainly on the indirect and direct effects of ocean warming on marine phytoplankton and more specifically on diatoms.

1.1.1 Effects of ocean warming on marine phytoplankton

1.1.1.1 Biological pump

The biological and solubility pump draw carbon from the atmosphere driven by biochemical and physical processes which results in a net flux of carbon to the deep oceans (Figure 1.1.1, biological pump left, solubility pump right). The solubility pump creates an equilibrium between the ocean and atmosphere, so when atmospheric carbon concentrations increase, so do the oceans, leading to ocean acidification. Additionally, the solubility of carbon in the ocean depends on sea surface temperature with cooler waters having greater solubility than warmer waters. With the formation of cooler deep waters in the higher latitudes (Arctic Ocean), denser and more acidic waters enter the thermohaline circulation eventually up-welling close to the equator and carbon dioxide outgasing due to the higher sea surface temperatures (Raven and Falkowski, 1999).

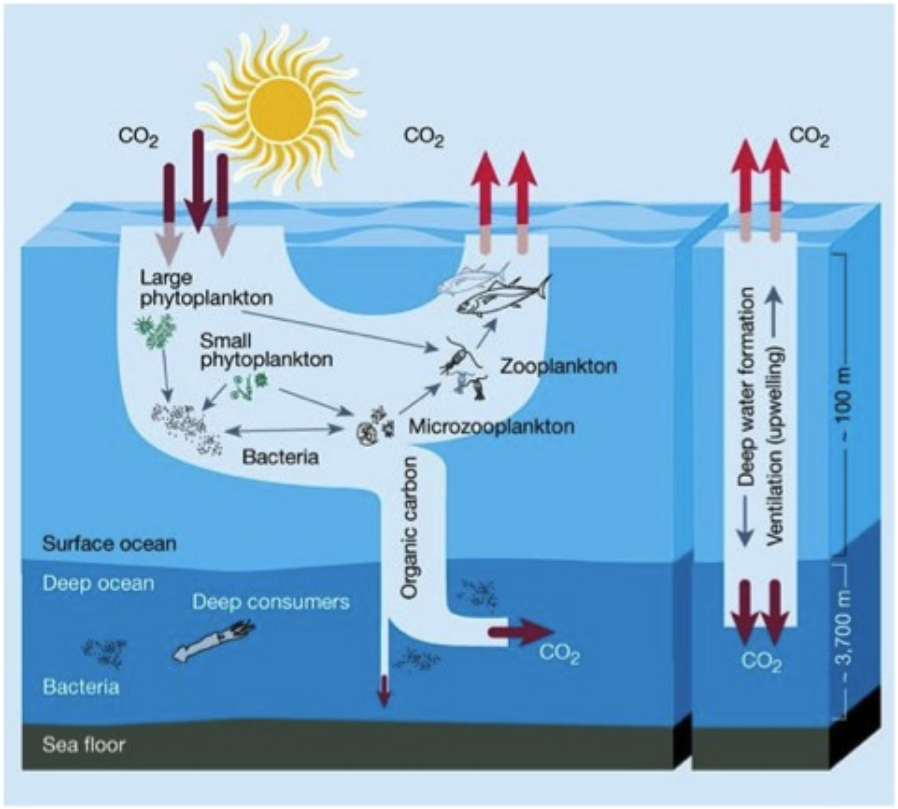

Figure 1.1.1: The biological (left) and solubility pump (right) maintain CO_2 gradients between the atmosphere and ocean. Phytoplankton convert CO_2 to organic matter forming the basis of the marine food web. This organic carbon is mostly respired supporting new primary production. However, a fraction of the organic carbon is incorporated into aggregates or fecal pellets. These particles sink through the water column were they are either remineralized by bacteria or reach the deep sea where the carbon is stored over long timescales. The solubility pump is driven by concentration differences between the atmosphere and the ocean and depends on temperature. The polar oceans uptake more carbon from the atmosphere than the tropical oceans as colder waters have a higher carbon solubility. As carbon dioxide concentrations in the atmosphere increase, the ocean becomes more acidic due to the solubility pump. Global warming has reduced the efficiency of the solubility pump to due warmer sea surface temperatures (Graphic from Chisholm (2000).

The biological pump is the biologically driven flux of carbon, through either primary production (organic carbon pump) or incorporation in shells as calcium carbonate (calcium carbonate pump) (Volk and Hoffert, 1985; Passow and Carlson, 2012), with the organic carbon pump having a magnitude higher efficiency in exporting carbon than the calcium carbonate pump (Sarmiento *et al.*, 2002). Phytoplankton primary production fixes CO_2 from the atmosphere through photosynthesis to produce their organic and inorganic carbon compounds (e.g. lipids, sugars, carbohydrates). Most of the primary production in the sunlit surface is respired and the nutrients recycled into new primary production (Raven and Falkowski, 1999). The formation of sinking particles through aggregation or in form of fecal pellets, is the main driver of particulate organic

carbon (POC) export to the deep ocean and sediments (Armstrong *et al.*, 2002). While these particles travel through the water column, they are remineralized by bacteria, providing the nutrients for new production (Passow, 2002). Only a small fraction reaches the sea floor and is stored over long timescales. The export efficiency of POC can be enhanced through the ballasting effect of silicate or calcium carbonate. Calcium carbonate has been shown to have a higher efficiency than opal as ballast (Klaas and Archer, 2002), however, calcium carbonate is more susceptible to dissolution by ocean acidification (Riebesell *et al.*, 2000). In contrast, diatoms form larger aggregates of higher density, and quicker than coccolithophores (Iversen *et al.*, 2010). Ocean warming might enhance the ballasting effect of diatom frustules as some diatoms grown under higher temperatures were more silicified (Paasche, 1980). A similar effect has been observed as a response to predation through copepods (Ragueneau *et al.*, 2006).

Particles sinking through the water column are subjected to respiration by bacteria. Through respiration inorganic carbon and nutrients are released and mixed into the water column. These nutrients then get recycled and are available for new primary production. Respiration is temperature sensitive as it depends on enzymes released by the bacteria colonizing the particle surface. Thus, with ocean warming, respiration rates are expected to increase, enhancing the recycling of organic matter in the upper ocean and preventing it from being exported to the deep sea. It is currently unclear how climate change might effect the efficiency of the biological pump and if and how the net export of carbon is affected. However, as primary production is a vital contributor to the biological pump, phytoplankton response to ocean warming can provide potential insights into biogeochemical cycles of the future ocean.

1.1.1.2 Stratification

Stratification in the ocean is driven by temperature and density changes in the water column with dense, cold waters at the bottom and warmer, lighter water masses above, forming a barrier between the water masses. This reduces turbulence in the water column, decreasing mixing depth which is vital for nutrient fluxes from nutrient rich deep waters to the sunlit surface waters. However, in temperate and high latitude regions, such as the North Atlantic, the annual onset of stratification by ocean surface warming and shoaling of the mixing depth, increases light availability and induces phytoplankton spring blooms (Brody and Lozier, 2015; Mahadevan

et al., 2012; Siegel *et al.*, 2002). Continued, strong stratification, however, can prevent nutrient fluxes to the productive, nutrient depleting, surface waters, ultimately reducing phytoplankton productivity and biomass.

Increases in sea surface temperatures, due to global warming, will affect lower and higher latitudes differently (Figure 1.1.2). Phytoplankton in the tropical Atlantic Ocean (lower latitudes) are typically nutrient limited as year round thermal stratification limits vertical mixing. Primary production will become more limited with ocean warming as vertical mixing and, thus nutrient supply, is further reduced (Figure 1.1.2, Upper) (Riebesell *et al.*, 2009). These nutrient-poor oligotrophic areas in the ocean have increased by 15% between 1998 and 2006 as a consequence of ocean warming and resulting enhanced stratification (Polovina *et al.*, 2008).

Phytoplankton production at higher latitudes is usually light limited due to a deep mixing depth, reducing the average light intensity. Increasing sea surface temperature and melt water input from sea ice could increase net primary production as the vertical mixing depth is reduced and thus the average light intensity increases (Figure 1.1.2, Lower) (Riebesell *et al.*, 2009).

Ocean warming has been suggested to result in a 10°C temperature increase within several decades for the Arctic Ocean, making polar phytoplankton communities more vulnerable than their tropical counterparts (Lovejoy *et al.*, 2006). Indeed, changes in water column stratification have been identified as drivers of changes in phytoplankton abundance, spatial distribution, size and community structure (Behrenfeld *et al.*, 2006; Finkel *et al.*, 2010; Doney *et al.*, 2012), potentially impacting marine food webs as almost all life in the ocean depends on phytoplankton productivity (Behrenfeld *et al.*, 2001).

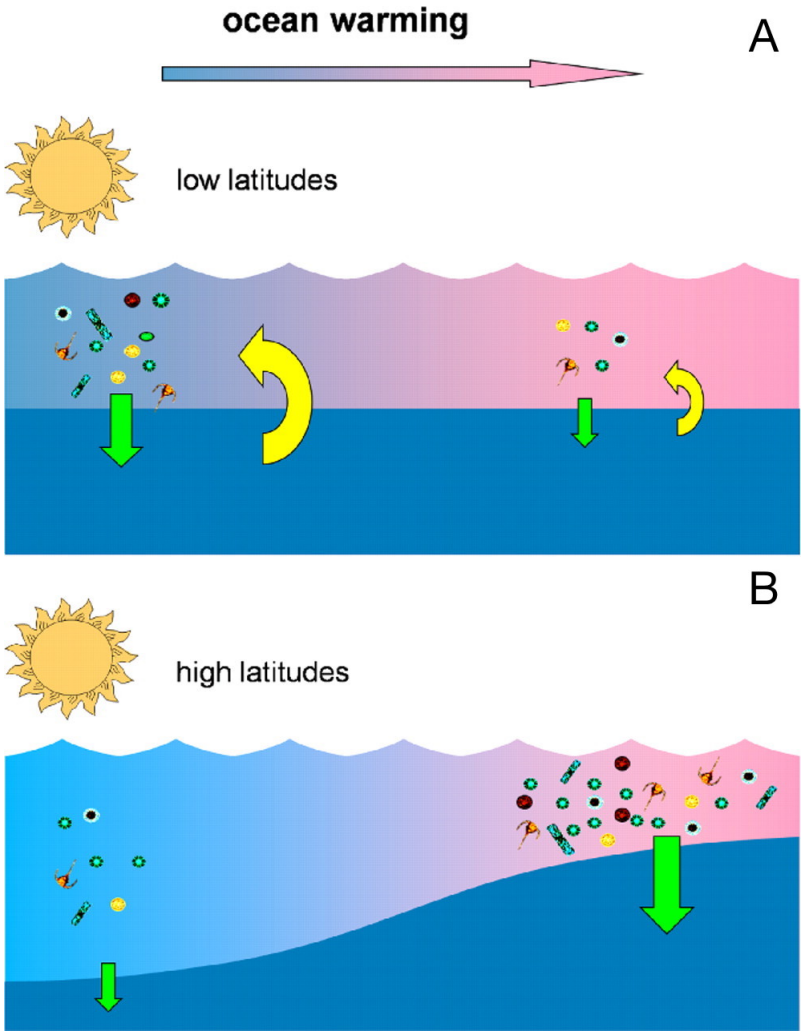

Figure 1.1.2: Summary of expected changes to water column stratification for low (A) and high (B) latitudes (from Riebesell et al. (2009).

1.1.1.3 Distribution shifts

Marine organisms, experiencing continuous change in their habitat can respond by acclimating their physiology within their current limits, or by tracking their niche in space (distribution shifts) and time (phenological shifts). Several studies have provided evidence that distribution shifts are already occurring in many terrestrial organisms but also in the marine environment. These distribution shifts can change local communities by outcompeting local species which might have negative effects on local food webs. Furthermore, a phenological shift might result in a decoupling of trophic levels, e.g. phytoplankton and zooplankton bloom (De Senerpont Domis *et al.*, 2007). However, not only habitat tracking by communities or species can cause community changes. The northern polar oceans are especially vulnerable to invasion of new species. With

the reducing ice coverage in the Arctic, the north-west passage allows water exchange between the North Pacific and North Atlantic (Reid *et al.*, 2007).

In 1999 a diatom species (Neodenticula seminae) endemic to North Pacific waters was reported in the North Atlantic after a 800,000 year absence (Reid *et al.*, 2007). The diatom now seems to be well established in the North Atlantic dominating phytoplankton blooms (Reid *et al.*, 2007). Its invasion might have been possible due to periods of an ice-free Arctic and other species can be expected to follow. However, this is one extreme example of distribution shifts. A more obvious scenario, for the Arctic Ocean, is the northern shift of temperate species. Ocean warming might result in temperature increases in the polar ocean allowing temperate species to colonize this former secluded environment. Both, warming and species invasion, is a great danger for the well adapted polar species. Polar species inhabit a small thermal regime and sea surface temperature (SST) increases might fall outside of their thermal range. The question therefore arises, how these changes will affect the polar ocean in terms of it food web and biogeochemical cycling, but also, if polar species can adapt to the rising temperature.

Tropical species are faced with similar situation. The high SSTs in the equatorial Atlantic are already at the thermal maximum for some species. Further temperature increases might be outside of their normal range forcing species to either track changes or adapt locally. Temperature tracking has been observed for many terrestrial species but also for some marine species like fish, zooplankton and phytoplankton. On average, marine species have travelled 54.6 $\pm$11.7 km decade^{-1} northwards from the equator, a magnitude higher than reported for terrestrial organisms (Poloczanska *et al.*, 2013). Additionally, highly mobile phytoplankton species have been found to expand their range by 469.9 $\pm$115.3 km decade^{-1} (Poloczanska *et al.*, 2013). These distribution shifts are likely to cause changes to phytoplanklton communities and consequently, to marine food webs.

1.1.1.4 Changes in communities and food webs

Ocean warming is affecting primary production rates and producer distribution in the world oceans. These changes to phytoplankton communities might have a great impact on marine food webs and ecosystem functioning. A long-term study by Li 2002 has shown that smaller phytoplankton species, picoplankton, increase under ocean warming in nutrient poor lower

latitudes. Further evidence was provided by Trainor et al. 1993 were Scenedesmus showed smaller phenotypes under higher temperatures. Increased stratification due to ocean warming caused declining zooplankton populations of the coast of California as the limited nutrient input supported less primary production. This has a trickle down effect as zooplankton is a major part of the food web and important to many species of fish and birds (Roemmich and McGowan, 1995). In contrast, phytoplankton abundance in the higher latitudes has been suggested to increase due to increasing temperatures and stratification (Richardson and Schoeman, 2004; O'Connor *et al.*, 2009). However, there might be a potential mis-match begin phytoplankton and zooplankton abundance as phytoplankton spring blooms occur earlier in the year.

Zooplankton abundance and food web stability is also influenced by the nutritional value of primary producers. Cyanobacteria are generally considered less quality food due to their low phosphate content while diatoms and dinoflagellates are of higher quality. Proteins, lipids, fatty acids and other cellular components contribute to the carbon to nutrient ratio of phytoplankton and determine the nutritional value as zooplankton food (Sterner and Elser, 2002). As the nutritional value of phytoplankton changes temperature-driven towards high carbon to nutrient content (De Senerpont Domis *et al.*, 2014), zooplankton can become nutrient limited. This can be counteracted by some zooplankton species through compensatory feeding. However, this potentially increases their carbon intake in relation to nutrients and zooplankton species with a higher nutrient demand (e.g. Daphnia, Brachionus) have been shown to respond by decreasing their grazing rates (Sterner and Elser, 2002; Urabe *et al.*, 2003; Jensen, 2004). This may affect herbivore communities as zooplankton species with a lower nutrient demand (e.g. copepods) might be able to sustain their growth on high carbon to nutrient ratio phytoplankton (Waal *et al.*, 2010). Indeed, Atkinson et al. 2004 have reported an increase of salp species (low nutrient demand) in the Southern Ocean, replacing krill (high nutrient demand).

Even though climate change driven changes to phytoplankton and zooplankton communities could potentially decouple trophic relationships (Winder and Schindler, 2004) and affect food-web functioning (De Senerpont Domis *et al.*, 2007) plankton adaptation might mitigate such effects (Van Doorslaer *et al.*, 2007). Inevitably, higher tropic levels such as fish and seabirds will be affected by changes in primary and secondary producers (Richardson and Schoeman, 2004).

1.1.1.5 Photosynthesis

Temperature modulates diatom physiology through its effect on enzyme kinetics (enzymes and ribozymes) and viscosity of membranes. The temperature dependence of processes, catalysed or uncatalysed, can be expressed as the temperature coefficient Q_{10} where a higher (>1) value is indicative of thermal dependence, a value smaller than 1 indicates a negative temperature dependence (Eppley, 1972). Photochemical processes, for example, have a Q_{10} of 1 and thus are temperature independent. In contrast, Rubisco carboxylase activity in *Phaeodactylum tricornutum* has a Q_{10} of 2.66, meaning that with increasing temperature, Rubisco activity increases (Raven and Geider, 1988). The combination of temperature sensitive and insensitive processes in photosynthesis can result in an imbalance between energy absorption and metabolic sink (e.g. Calvin cycle) when temperature conditions change (Raven and Geider, 1988; Ensminger *et al.*, 2006). Additionally, physiochemical properties of membranes in response to temperature changes can affect membrane-bound processes. High temperature photoinhibition has long been attributed to reduced electron transport due to changes in membrane fluidity and disassociation of light harvesting complexes from the thylakoid (Horvath *et al.*, 1987; Raison *et al.*, 1982; Havaux, 1993). More recent work suggests that the reactivation of Rubisco is the limiting step as, under increased temperature, inactivation rates are higher than Rubisco activation (Feller *et al.*, 1998; Salvucci and Crafts-Brandner, 2004). This is due to the low temperature optimum of Rubisco activase and thus not being able to compensate for the faster rates of Rubisco inactivation (Salvucci and Crafts-Brandner, 2004). In red algae, high temperature photoinhibition has been linked to diminished electron transport between phycobilisomes and PSII (Kuebler *et al.*, 1991), while thermal instability of thylakoid membranes has been linked to photoinhibition in plants (Berry and Raison, 1981).

In contrast, low temperature acclimation mimics responses to high irradiance (Anning *et al.*, 2001; Maxwell *et al.*, 1994; Krol *et al.*, 1997). Low temperature conditions decrease enzyme kinetics and decrease membrane fluidity. This can lead to surplus of energy absorption over photosynthesis causing chlorosis (Geider *et al.*, 1998). If the absorbed light energy exceeds the cells capacity to distribute it towards photosynthesis or to dissipate it (non-photochemical quenching NPQ), photodamage occurs as the rate of D1 protein destruction at PSII is higher than the repair (Melis, 1999). Reduced chlorophyll a (chl a) concentrations can counteract low

temperature photoinhibition, in addition to the production of photoprotective pigments such as carotenoids (Davison, 1991; Geider, 1987; Geider *et al.*, 1998). Indeed, chl a concentration have been shown to correlate inversely with growth temperature for green algae and diatoms (Anning *et al.*, 2001; Thompson *et al.*, 1992*a*; Krol *et al.*, 1997). The xanthophyll cycle plays a crucial role in dissipation of excess energy (Horton *et al.*, 1999). Conversion of light harvesting pigments to quenching pigments results in an energy sink, as they are unable to pass their excitation energy to chl a (Horton *et al.*, 1999; Niyogi *et al.*, 2005). Thus, increases in carotenoids and other xanthophyll could be expected under low temperature conditions. This photoprotective response has been found in cyanobacteria, diatoms and green algae (Tang and Vincent, 1999; Várkonyi *et al.*, 2002; Anning *et al.*, 2001; Krol *et al.*, 1997). In *Cylindrospermopsis raciborskii* carotenoids have been suggested to increase photoprotection due to restricted movement of quenchers under low temperature (Várkonyi *et al.*, 2002).

Membranes consist of lipids built from fatty acids. Saturation states of fatty acids and thus membrane fluidity can be adjusted by desaturases. Desaturases introduce double bonds into the chains of fatty acids. This can be done to already established membranes, however, increasing saturation can only be accomplished by novel synthesis of saturated fatty acids in the chloroplast (Ohlrogge and Browse, 1995). Membrane fluidity is decreased under low temperature conditions potentially reducing electron transport through the thylakoid membrane (Morgan-Kiss *et al.*, 2002) facilitated by Q_A^- and the plastoquinone pool (PQ) (Horvath *et al.*, 1987). This can be counteracted by the desaturation of fatty acids of membranes (White and Somero, 1982; Thompson *et al.*, 1992*b*). Additionally membrane fluidity plays a crucial role in stabilizing the state of LHCII (Dubacq and Tremolieres, 1983).

The fluidity of membranes has been suggested as a mechanism with which cyanobacteria sense temperature (Vigh *et al.*, 1993). Expression of desaturases and heat shock proteins has been suggested to be induced by the physical state of thylakoid membranes (Vigh *et al.*, 1993; Los *et al.*, 1993; Horváth *et al.*, 1998). Furthermore, the PQ can translate energy signals into biochemical signals regulating expression of genes supporting low temperature acclimation of photosynthesis and cellular metabolism (Ensminger *et al.*, 2006).

1.1.1.6 Stoichiometry and size

The elemental composition of diatom cells is modulated by growth rate as well as environmental conditions, such as temperature. Changes in macromolecule abundance and composition impact the elemental composition as lipids (mostly C), proteins (mostly N) and ribosomes (mostly P) contribute to different element pools (Geider and La Roche, 2002). Redfield (Redfield, 1958) proposed a general elemental ratio of 106:16:1 (C:N:P) for marine phytoplankton which was based on element concentrations below the thermocline. However, Geider and La Roche 2002 compiled data from a range of marine algae and cyanobacteria, showing that the C:N:P ratio is rather plastic and can range from 3-13 (C:N) and 7-19 (N:P). Changes to the C:N:P ratio have been shown to be linked to growth rate (Allen and Gillooly, 2009), but also to temperature induced changes of resource allocation (Gillooly *et al.*, 2005; Elser *et al.*, 2000). Higher growth rates have been suggested to require higher concentrations of ribosomes to support the increased demand of proteins (Elser *et al.*, 1996; Klausmeier *et al.*, 2004). This has consequences for the cellular phosphate (P) content as ribsomes consist of P-rich rRNA and thus lead to low N:P ratios (Geider and La Roche, 2002). Therefore, changes in the N:P ratio highlight the balance between protein and rRNA biogenises (Loladze and Elser, 2011; Klausmeier *et al.*, 2004). Additionally, low temperature conditions have been suggested to result in increased cellular P quotas as lowered reaction speeds of protein synthesis can be off-set by more ribosomes (Toseland *et al.*, 2013).

The observed relation between temperature and organisms size, the ”temperature-size rule” (TSR), predicts a 2.5% increase in size for every 1°C decrease (Atkinson, 1994). Most diatoms seem to follow this TSR (Adams *et al.*, 2013), however, sublethal high temperatures can cause size increases (Atkinson *et al.*, 2003). Additionally, temperature modulates metabolic rates due to the temperature dependence of enzymes (Gillooly *et al.*, 2001; Peterson *et al.*, 2007). A reduction in membrane viscosity can have a negative impact on nutrient acquisition as the uptake through the membrane is slowed. Polar diatoms have a higher acquisition rate for silicic acid compared to temperate species, but their utilization for growth is less (Stapleford and Smith, 1996). Nitrate acquisition has been suggested to be inversely related to temperature in marine algae and bacteria independent of their thermal niche (Reay *et al.*, 1999). In addition, phosphate utilization in marine phytoplankton appears to be temperature independent (Reay *et al.*, 2001). Reduced

nutrient acquisition can be compensated by means of decreasing cell size, increasing the surface area for nutrient uptake in relation to cell volume (Hein *et al.*, 1995).

1.2 The role of phenotypic plasticity and epigenetics in experimental evolution with phytoplankton[1]

Experimental evolution (EE) approaches are increasingly used to investigate evolutionary processes of organisms to controlled experimental conditions and can be applied to study several facets of adaptation and evolution (e.g. estimate mutation rates or evaluate evolutionary theories and concepts) (Kawecki *et al.*, 2012). Examples of EE experiments include studies on organisms ranging all three domains of life: Bacteria, Archaea and Eukaryota (e.g. *Chlamydomonas reinhardtii* (e.g. Bell 1997; Ratcliff *et al.* 2013; Collins and De Meaux; Perrineau *et al.* 2013), *Escherichia coli* (e.g. Plucain *et al.* 2014; Lenski *et al.* 1991; Blount *et al.* 2012), *Saccharomyces cerevisiae* (e.g. Lewis *et al.* 1995; Dhar *et al.* 2011; Ratcliff *et al.* 2012), *Arabidopsis thaliana* (e.g. Bond and Baulcombe 2015; Rahavi 2011; Blödner *et al.* 2007) and *Drosophila melanogaster* (e.g. Azevedo *et al.* 2015; Burke *et al.* 2010)) shedding light on the general processes but also specific outcomes of evolution. More recently, researchers have started to use EE approaches to provide insight into phytoplankton responses to climate change. The ideal organisms for studying evolution, on a reasonable time scale, have a high turnover rate, are easy to cultivate, store, manipulate, revive and have different available clones (Collins, 2013). Organisms investigated in evolution experiments so far included a variety of species from prokaryotes (e.g. *Escherichia coli*, *Pseudomonas sp.*) to eukaryotes (e.g. *Chlamydomonas*, *Saccharomyces*, *Drosophila*, *Emiliania huxleyi*, *Ostreococcus tauri*, guppies (*Poecilia reticulata*), and Diatoms) (Collins 2013; van Wijk *et al.* 2013) experiencing varying selection pressures (e.g. carbon source, predation pressure, temperature, pH, salinity). Marine phytoplankton possess the above mentioned qualities, making them good candidates for experimental evolution studies investigating the effects of global ocean change on their physiology, dispersal and survival.

Marine phytoplankton contribute about 50% of annual global carbon fixation and form the basis of most marine food webs (Field, 1998; Falkowski *et al.*, 2008). Furthermore, major functional

[1] **This section has been published by Schmidt et al. in "The role of phenotypic plasticity and epigenetics in experimental evolution with phytoplankton", Perspectives in Phycology, 2016.**

groups of phytoplankton play important roles in elemental cycles as carbon fixers, nitrogen fixers, calcifiers and silicifiers (Tréguer *et al.*, 1995). Under current climate change, phytoplankton experience multiple, simultaneous and interacting changes (Boyd *et al.*, 2015*a*) in their environment, which affect primary production, community composition and biogeochemical cycles. How phytoplankton respond to climate change depends on their phenotypic plasticity (the ability of a single genotype to produce multiple phenotypes) and adaptation (via novel genetic changes) of organisms within populations which ultimately results in species sorting that affects the composition of the phytoplankton community (Figure 1.2.1). Quantifying the role each response plays in phytoplankton adaptation to climate change is an increasingly important field of research in experimental evolution.

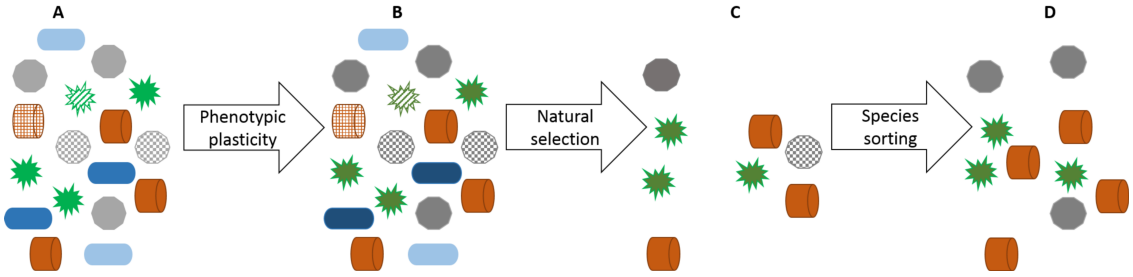

Figure 1.2.1: Species and genotype sorting in mixed phytoplankton communities.

The ability of phytoplankton to respond to climate change on the level of an individual cell can be largely considered by taking into account their genetic repertoire that underpins the initial response to a change in environmental conditions and therefore the degree of phenotypic plasticity. Larger genomes contain a higher diversity and redundancy of genes and therefore enable the expression of multiple phenotypes from a single genotype (high phenotypic diversity) and therefore usually generate a higher phenotypic plasticity than for organisms with a smaller genome (e.g. cyanobacteria) (Koonin and Wolf, 2008; Bentkowski *et al.*, 2015). On the level of populations, effective population size (Ne) determines how a population is able to respond to environmental change by genotype sorting without contributions from *de novo* mutation (Via and Lande, 1985). Genotype sorting, in asexual organisms, has been shown to lead to the dominance of one genotype in the population, such that the changes in genotype frequencies then impacts the average population response (e.g. (Tatters *et al.*, 2013*a*,*b*; Collins, 2011; Lohbeck *et al.*, 2013)). Over selection time scales, mutational supply (mutation rate x effective population size) is also

important in generating novel heritable variation in fitness on which natural selection can act. Phytoplankton with a high effective population size are thus likely to have (as standing variation) or produce (through mutation) standing genetic variation that is available to natural selection. Furthermore, in large populations changes in allele frequencies are governed more by selection than by random genetic drift, which means that genetic variants that are nearly-neutral can be expected to contribute to an adaptive evolutionary response under environmental change (Fisher, 1930; Hare *et al.*, 2011).

Although large microbial populations can have or produce large amounts of genetic variation, adaptation may be slowed, by the absence of sexual reproduction (Kaltz and Bell, 2002) if selection happens mainly during asexual growth. This is because sex facilitates recombination and generates novel genotypes, and counteracts clonal interference, and so the rate of sexual reproduction will affect the speed of adaptation of microbial populations. Given the difficulty in identifying sexual reproduction in some phytoplankton species (Chepurnov *et al.*, 2004), an important future research avenue is to examine the evidence of recombination in phytoplankton populations using population genetic tools. In addition to adaptive evolution, microbes can also respond through phenotypic plasticity. It is well known that many phytoplankton species are very responsive to changing growth conditions due to quickly adjusting their transcriptomes, proteomes and metabolomes (Armbrust, 2009). However, what orchestrates the regulation of gene expression, the activity of proteins and metabolites is still largely unknown but key for our understanding how phenotypic plasticity contributes to adaptation. From models and studies examining gross phenotypic changes such as growth or photosynthesis rates, there is some evidence that populations from highly heterogeneous environments, such as coastal seas at higher latitudes, have larger plastic responses than populations from less variable environments (Schaum *et al.*, 2013), which further increases their chances of evolving (Draghi and Whitlock, 2012; Schaum and Collins, 2014). Thus, phenotypic plasticity may provide a buffer against environmental fluctuations for many phytoplankton that could allow large populations to persist (Charmantier *et al.*, 2008) and adapt.

What is the mechanistic basis of the physiological differences of highly plastic and less plastic organisms? Because phenotypic plasticity requires responding to environmental cues, variations

in sensing the environment and regulation of pathways might explain the observed differences of plasticity. However, the molecular basis of environmental sensing and regulation of plasticity in phytoplankton is currently unclear. Physiological responses imposed by a rapidly changing environment are mainly the consequence of signalling cascades that sense a change in the environment and response regulators translate signals into an appropriate physiological response. Several of these sensors and regulators have been identified in plants (mainly *Arabidopsis*) and a few in marine phytoplankton (see below).

In addition to changes in the sensory apparatus available to different genotypes or species, epigenetic marks, like DNA methylation, RNAi and histone modifications dynamically regulate gene expression that can alter phenotypes (Duncan *et al.*, 2014). Epigenetics (non-genetic regulation of physiological responses) is known to play an important role in regulating gene and protein activity in many organisms including phytoplankton in response to environmental changes which could also underlie differences in plastic responses between genotypes or species.

1.2.1 Epigenetic control of transposable elements

Epigenetics is the study of phenotypic changes that are not caused by changes of the genotype. Single-celled phytoplankton are typically studied during asexual reproduction, even though many also reproduce sexually, non-lethal epigenetic marks can be transferred through both reproductive modes (Maumus, 2009). The major mechanisms involve DNA methylation, histone modification and RNAi (Bossdorf *et al.*, 2008; Bird, 2007; Fablet and Vieira, 2011) activating, reducing or disabling gene expression, protein function or histone activity. Epigenetic regulation has been studied intensively in animals, plants and fungi (Maumus *et al.*, 2011), whereas studies on phytoplankton are still in their infancy. The most prominent example in plants is the temperature-mediated epigenetic regulation of a locus responsible for flowering of Arabidopsis in spring. During colder months, the repressor FLOWERING LOCUS C (FLC) is active, preventing production of flowers. However, prolonged cold progressively reduces expression of FLC, which is mediated through methylation marks. Once silenced and favourable temperature and light conditions are sensed by the plant, flowering is activated (Li *et al.*, 2014). Thermosensory in Arabidopsis has been shown to be facilitated by histone modification (Kumar and Wigge, 2010). Under cold conditions the histone H2A.Z replaces H2A wrapping DNA more tightly. This limits

the translational accessibility of the wrapped region inhibiting gene expression. Increasing temperatures overcome the tight wrapping giving evidence to a temperature dependent regulation. More importantly, H2A.Z activity has been shown to be independent of transcription suggesting direct mechanisms by which temperature epigenetically regulates gene expression.

Genome methylation has been the focus of many epigenetic studies including studies on phytoplankton and it has been shown that the degree and patterns of DNA methylation seem to vary greatly across species, ranging from 18% in dinoflagellates (Rae, 1976) and about 6% in the diatom *Pheodactylum tricornutum* (Veluchamy *et al.*, 2013). The process of DNA methylation is less accurate than DNA synthesis (1 error for every 10^4 CG pair (Schmitz *et al.*, 2011) compared to 1 error for every 10^6-10^9 bases copied (Chevin *et al.*, 2010; Klironomos *et al.*, 2013) therefore, it can more rapidly lead to a diversification of methylation patterns (Bird, 2007), compared to genetic variation, on which selection can act (Klironomos *et al.*, 2013; Flores *et al.*, 2013). Thus, epigenetic variation can enhance phenotypic variation at much shorter timescales as the genetic mutation frequency is much lower (Boyko and Kovalchuk, 2011). Furthermore, epigenetic variation can exceed genetic based estimates of variation and environmentally induced adaptive changes in the epigenome can be transferred to future generations (Richards, 2006). For example, ecotypes of Arabidopsis and single genotypes of yeast showed that natural variation of epigenetic marks in the absence of genetic variation lead to diverse and adaptive phenotypes (Bossdorf *et al.*, 2010; Riddle and Richards, 2002; Herrera *et al.*, 2012). Environmentally induced changes in the methylation pattern can therefore, have similar phenotypic effects as DNA mutations (Biémont and Vieira, 2006). Epigenetic mechanisms in marine phytoplankton are currently poorly understood. Recent research indicates that DNA methylation, RNAi and histone modification might be mechanisms used by phytoplankton to control transposable elements (TEs) (Veluchamy *et al.*, 2013) and foreign DNA (De Riso *et al.*, 2009). A study on DNA methylation in the diatom *Pheodactylum tricornutum* revealed that most methylated genes encode important metabolic pathways and are differentially expressed under different environmental conditions (Veluchamy *et al.*, 2013). Another study on a marine cyanobacterium (*Trichodesmium*) suggested that high nitrogen fixation rates evolved under high CO_2 conditions are related to DNA-methylation activity, however the ultimate mechanism remains unknown (Hutchins *et al.*, 2015). Other findings suggest that different phytoplankton groups, depending on their evolutionary history, show functional divergence of epigenetic mechanisms compared to other

eukaryotes (De Riso *et al.*, 2009). For example, despite the lack of detecting histone proteins in the dinoflagellate *Lingulodinium*, expression of histone genes has been reported (Roy and Morse, 2012). Furthermore, only distantly related genes for RNAi core components have been found in some phytoplankton groups and their absence in others indicates an evolutionary loss of machinery (De Riso *et al.*, 2009; Maumus *et al.*, 2011).

Transposable elements (TEs) are DNA sequences that can copy or cut a duplicate of themselves at different places in the genome disrupting gene function, altering levels of gene expression or triggering chromosomal rearrangements (Maumus *et al.*, 2009). Activation of TEs has been shown to be related to stressful abiotic and biotic conditions like temperature changes, UV, nutrient availability and can also be controlled by their methylation status (Hermann *et al.*, 2014). TEs can cause a great deal of *de novo* intraspecific genetic variability that is likely to have phenotypic effects, and can thus produce heritable variation that natural selection can act on (Casacuberta and González, 2013; Fablet and Vieira, 2011). Retrotransposons are the most abundant class of transposable elements found in diatoms (Maumus *et al.*, 2011), which can be induced by environmental stress, and could be a useful tool for marine diatoms to fast-track adaptation (Mirouze and Paszkowski, 2011). *Pheodactylum tricornutum* showed, for example, a highly enhanced activity of a retrotransposon to nitrate starvation, potentially enhancing its adaptation capability (Maumus *et al.*, 2009). Another class of transposons has recently been shown to be responsible for the evolution of a irreversible B12-dependent clone in the usually B12-independent green algae *Chlamydomonas reinhardtii*. Gene disruption was reversible for some evolved clones, yet in one clone reversion left a transposon footprint which caused genetically fixed B12-dependency (Helliwell *et al.*, 2015). However, in both, Pt and *Chlamydomonas*, the transposon regulation remains unclear. The epigenome might poses an additional system on which selection can act on since some epigenetic modifications are independent of the genetic code (Bossdorf *et al.*, 2008). However, while the lack of empirical evidence of these processes in marine phytoplankton limits our understanding, it opens up a whole new field in phycology. An important challenge for todays phycologists will be to address epigenetic questions in the context of climate change. For example, experimental evolution studies investigating phytoplankton plasticity to climate change potentially hint to the role of epigentic processes as some studies (Schaum and Collins, 2014) revealed that phenotypic plasticity might be attributed to epigenetic modifications. Sequencing the epigenomes of key

species under changing environmental conditions can therefore provide valuable insight into the epigenetic toolbox of phytoplankton plasticity.

1.3 Diatoms

Diatoms (Bacillariophyceae) are small (2µm to 2 mm), unicellular, photosynthetic, eukaryotic algae surrounded by a symmetrical ornately structured silica cell wall (frustule). This frustule is their name giving feature, derived from the greek ”diatom”: dia ”through” and temein ”to cut in half”. Bacillariophyceae emerged from a secondary endosymbiosis between a photosynthetic eukaryote, most likely red algal-like, and a heterotrophic eukaryote (Falkowski, 2004), providing them a distinct genetic make up. Diatom total diversity is estimated to be about 100,000 species, across 250 genera (Round *et al.*, 1990), however only 2,000 distinct species have been described genetically (based on rDNA sequences) (Vargas *et al.*, 2015), and are one of the most diverse eukaryotic lineages. Diatoms can be found in almost every aquatic environment and show a global distribution. They are key players in biogeochemical cycles of silicate (Smetacek, 1999) and carbon (Tréguer *et al.*, 1995) and are the basis of most marine food webs. Furthermore, they are the most successful taxa in cold sea ice ecosystems (Wilhelm *et al.*, 2006; Arrigo *et al.*, 2010). In our current marine ecosystems, diatoms constitute one of the most successful groups (about 1400 to 1800 species) of phytoplankton and produce about 50% of the annual marine primary production (Nelson *et al.*, 1995; Bowler *et al.*, 2010) which is just as much as the production of all tropical rain forests.

Traditionally, diatoms are divided into two orders, the radial symmetric *Bidulphiales* (centrales), thought to have arisen around 180 Million years ago (Sims *et al.*, 2006), and the bilateral symmetric *Bacillariales* (pennales), arisen approximately 90 Million years later (Kooistra and Medlin, 1996). The pennales are furthermore divided into raphids and araphids, depending on the presence or absence of a longitudinal groove in the valve (raphe). Raphid diatoms are capable of ”sliding” along surfaces via secretion of polysaccharides through their raphe. This allows pennate diatoms to colonise surfaces, forming biofilms, while centric species are more abundant in the open ocean where they can form long chains. Both orders share a common shape consisting of two halves, a epi- (upper part) and a hypovalve (lower part), which are linked by

girdle bands (cingulum) (Figure 1.3.1). These parts are held together by a surrounding organic layer which can obscure frustule observations in the scanning electron microscope.

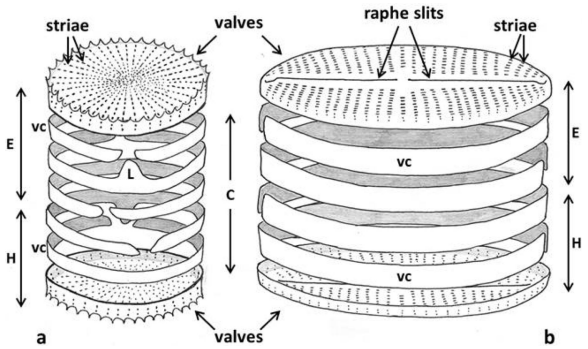

Figure 1.3.1: The diatom is composed of an epitheca and hypotheca that fit together like a petri dish. Each theca is composed of the valve face (epivalve or hypovalve) and the valve mantle. The overlap region of the valves is surrounded by structures referred to as girdle bands and this region is referred to as the cingulum (From Cox 2014).

The ornamented valves of diatoms are their most prominent and identifying feature (Figure 1.3.2). Classification of diatoms is based on the extent and fineness of pores (areolae), and the arrangement of other defining structures such as raphes, spines or portulae. Each valve feature has an important function for the diatom cell, for example, areolae allow gas exchange as well as nutrient uptake, spines support vertical movement of the cell.

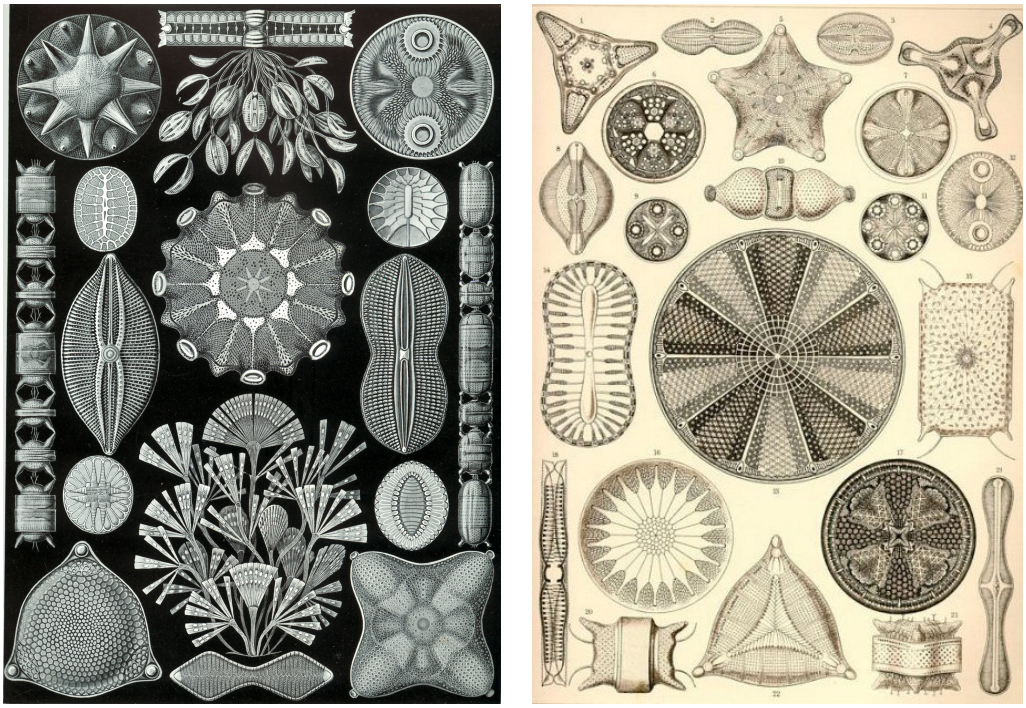

Figure 1.3.2: Haeckel drawings of diatom species, indicating their beautiful silicified cell wall.

Diatoms usually reproduce asexually by mitosis, however sexual reproduction has also been observed. During asexual reproduction, each daughter cell takes a parent valve as its new epitheca. Within the limits of the parent cell a new hypotheca is produced. This process is thought to cause a cell size reduction with every division. Once a certain size threshold is reached, cell size is restored by (sexual) auxospore production. However, some taxa restore size without auxospore production or show no decrease in size.

Observation of sexual reproduction in diatoms is inherently difficult and has only been reported for a small fraction of the described species (Round *et al.*, 1990; Armbrust and Galindo, 2001). Sexual reproduction differs between centric and pennate species. Gametogenesis occurs at the end of meiosis and zygotes develop into auxospores. Within the boundaries of the auxospores, a new, bigger cell is produced.

Another stage in the diatom life cycle, can be the formation of resting spores. Their formation can be induced by environmental stress like temperature changes or nutrient depletion. The ability of forming resting spores differs between species and individual cells. For example, only cells of a certain size are able to produce resting spores while for some species every cell has the potential to form one (French and Hargraves, 1985). Resting spores can be much denser than vegetative cells and can therefore sink to the sea floor until they are revived by favourable

conditions. This is thought to be one mechanism of spring bloom seeding in the polar ocean.

1.3.1 Endosymbiont theory

Diatoms have a distinct genetic make up that is heavily influenced by their evolutionary history. While glaucophytes, green and red algae plastids are based on a primary endosymbiosis event, diatoms and other members of the chromalveolate supergroup originated from a (single) secondary endosymbiosis event (Figure 1.3.3) (for a most up to date review see Hopes and Mock 2015).

Primary endosymbiosis involved a non-photosynthetic eukaryote engulfing a cyanobacterium resulting in a two membrane photosynthetic apparatus (Falkowski, 2004). This explains the monophyletic origins of all eukaryotic plastids which has occurred around 1.8 billion years ago (Gould *et al.*, 2008; Keeling, 2010). Later evolution of multicellularity in the green algal lineages, gave rise to terrestrial plants (Wodniok *et al.*, 2011). As part of the establishment of the cyanobacterium in the host, a thousand cyanobacterial genes were transferred to the host genome, while the chloroplast genome was reduced to a few hundred genes (Prihoda *et al.*, 2012).

In contrast, secondary endosymbiosis, giving rise to diatoms (red algal and green endosymbiont) and other chromalveolates (green, and/or red endosymbionts) (Moustafa *et al.*, 2009; Deschamps and Moreira, 2012), resulted in plastids with three or four membranes (Chaal and Green, 2005). From the outside to the inside, these membranes originate from the exosymbiont endomembrane, the plasma membrane of the second endosymbiont and the two membranes from the primary plastid (Chaal and Green, 2005). In addition to the green and red algal endosymbiont, evidence of a chlamydial primary endosymbiont has been found, providing novel functional genes (Allen *et al.*, 2006). The incorporation of endosymbiont genes into the host genome is thought to underpin the chimeric diatom genomes (Armbrust *et al.*, 2004; Falkowski, 2004; Bowler *et al.*, 2008) and probably contributed to the high diversification of Bacillariophyta (Prihoda *et al.*, 2012). This allowed them to adapt to a variety of habitats and cope with a wide range of environmental conditions.

The mix and match nature of diatom genomes becomes even more evident by the following examples. The urea cycle allows metabolising complex organic nitrogen compounds and the excretion of nitrogenous waste products. This cycle was thought to be only present in metazoans, however, it is has been found in several diatoms (e.g. *Thalassiosira pseudonana*) and allows rapid recovery after prolonged nitrogen limitation (Allen *et al.*, 2011). Additionally, diatoms have the ability to breakdown fatty acids (fat) and using the generated energy to produce metabolic intermediates (Chan *et al.*, 2013). This is especially advantageous for survival in polar oceans with long periods of darkness (Strauss, 2012). These and other examples underpin the characterization of diatoms as neither plants nor animals (Armbrust, 2009).

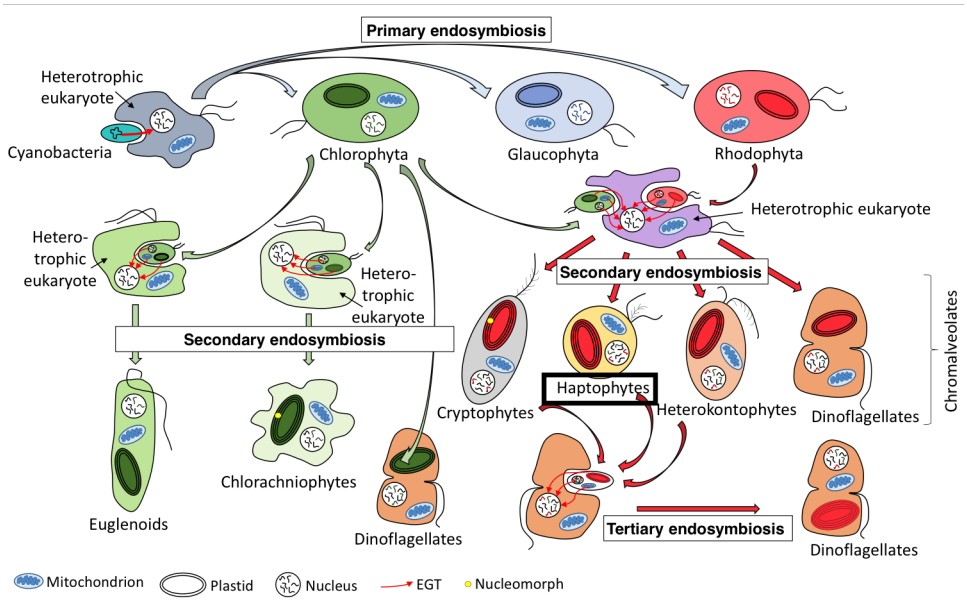

Figure 1.3.3: Schematic representation of the endosymbiont theory with highlighted (black box) heterokonts to which diatoms belong. From Hopes and Mock 2015, courtesy of Amanda Hopes.

1.3.2 Thalassiosira - *Thalassiosira pseudonana* (Hustedt) Hasle et Heimdal

The diatom under investigation in chapter three and four is the species *Thalassiosira pseudonana* (Hustedt) Hasle et Heimdal (Tp) which belongs to the cosmopolitan genus *Thalassiosira*. Tp is a small (less than 10µm) cylindrical diatom with circular valves (Figure 1.3.4). It has a ring of fultoportulae around the valve mantel, which allow excretion of chitin fibres to increase buoyancy (Morin *et al.*, 1986). Tp was first isolated in 1958 from a coastal location at Moriches Bay, New York, US and has since been maintained in culture collections. It has been shown that Tp

maintains high growth rates over a wide range of salinities (0.5-37‰) (Guillard and Ryther, 1962). Tp has been originally classified as *Cylothella nana*, a fresh water species, and it has been shown phylogenetically that Tp is conspecific to the Cylothella genus (Alverson *et al.*, 2011). Alverson et al. (2011) furthermore suggest that based on the morphological and phylogentic analysis, Tp falls within the Cylothella genus and conclude to resurrect Tp's original, *Cylothella nana*.

Tp's genome (34Mb, 24 chromosomes, 11,242 protein-coding genes) has been sequenced in 2004, revealing unanticipated metabolic pathways (e.g. Urea cycle) and providing insights into the mosaic nature of diatom genomes. Additionally, most transcription factors in the Tp genome belong to the heat shock families, giving temperature a important role in regulating gene expression. Other mechanisms to likely impact gene expression are chromatin binding and condensation related or through regulation of RNA processing. These mechanisms act on transcription and translation, however, several transposons have been found in Tp's genome (Armbrust *et al.*, 2004), which can impact protein function, by introducing single nucleotide polymorphisms or insertions. In 2006 Poulsen et al. (2006) reported the first genetic transformation system for Tp. Together with the development of expression vectors, this has enabled diatom genomics and established Tp as a model organism.

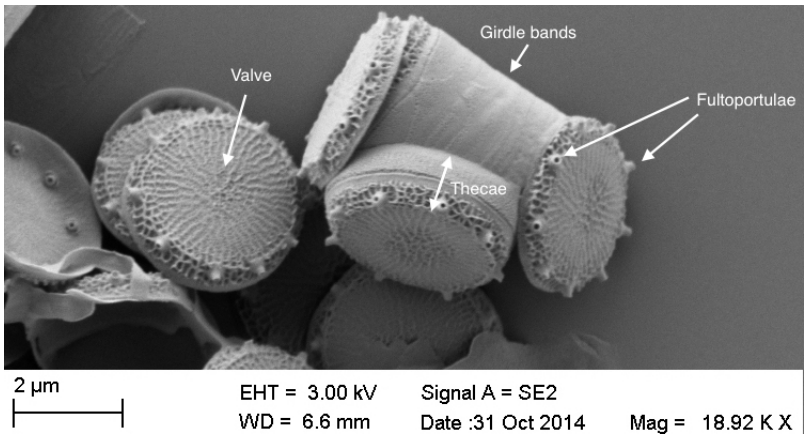

Figure 1.3.4: Scanning electron micrograph (SEM) of cleaned frustules of *Thalassiosira pseudonana* showing thecae, valves and girdle bands.

1.4 Thesis outline and aims

With the ongoing climate change, ocean environments are changing at an unprecedented rate. One main environmental factor changing is sea surface temperature, which has been shown to be the most influential on phytoplankton communities in the ocean. Recent research efforts are trying to identify the impact of ocean warming on natural phytoplankton communities and species adaptation. This thesis contributes to this effort by investigating the following overarching goal:

How does temperature drive diatom communities in the Atlantic Ocean and their pheno- and genotypic basis of thermal adaptation in the laboratory?

To answer this research question we embarked on three research cruises to survey diatom diversity and distribution in the Atlantic Ocean from the polar to the tropical Atlantic (chapter 2). The phenotypic basis of thermal adaptation is investigated in chapter 3 by using experimental evolution approaches to evolve thermo- and coldtolerance in the model diatom *Thalassiosira pseudonana*. Additionally, the genetic basis of thermal evolution was investigated through whole-genome transcriptomics and genome resequencing (chapter 4).

Chapter 2 provides insights, from three field campaigns, into diatom communities in the Atlantic Ocean. The aim of this chapter was to investigate diatom diversity and abundance over broad spatial scales in the Atlantic Ocean and to identify underlying environmental drivers. For this, diatom 18S rDNA ribotypes were analysed at the order level and related to environmental metadata such as temperature, salinity and nutrient concentrations. Hypothesis to be tested were the following:

- Temperature is a major environmental driver of diatom community and distribution.
- Diatom diversity is negatively correlated to temperature.
- Diatom community differences increase with increase distance.
- Abundance and distribution of the diatom order Thalassiosirales was driven by temperature.

In **chapter 3**, an experimental evolution approach was utilised with the aim to characterise the physiological and phenotypic responses to long-term (300 generations) temperature selection of

the model diatom *Thalassiosira pseudonana* (Tp). For this purpose, replicate cultures of Tp were semi-continuously cultivated at high and low temperature settings. Physiological responses and changes in Tp's elemental composition were measured after 300 generations. The generated dataset provided insights into temperature adaptation of this diatom. The experimental setup allowed investigating the following hypothesis:

- High and low temperature selection lines adapt genetically to novel temperature conditions.
- Stoichiometric changes underpin thermal adaptation.
- Cell size follows an inverse relationship to temperature with bigger cells under lower temperatures.
- Light-temperature interactions result in changes of cellular chlorophyll concentrations.
- Temperature optima shift towards selection temperatures while the temperature maximum remains stable.

Chapter 4 aimed to elucidate molecular underpinnings of the evolution of thermal tolerance using whole-genome transcriptome sequencing and genome resequencing. Temperature specific gene expression pattern were identified and potential adaptive differential expression of genes part of cellular pathways were highlighted. Single nucleotide polymorphisms were reported for one high temperature replicate and summarized based on gene function. The following hypothesis were tested with the above described approach:

- Phenotypic responses to selection temperatures in chapter 3 were underpinned by transcriptomic responses.
  - Low temperature adaptation is mediated by increased expression of genes related to the translation machinery.
  - Adjustments to membrane fluidity aided low temperature adaptation of the photosystem.
- Short-term (50 generations) and long-term (300 generations) responses of the cellular transcriptome differ.

- Transposon activity aids temperature adaptation.

Finally, **chapter 5** will conclude with a summary of the major findings of this thesis and a general discussion of the main results followed by suggestions for future research.

Chapter 2

Sea of change - Diatom ecology in the Atlantic Ocean

2.1 Introduction

Diatoms are single-cell autotrophs with a silica frustule cell wall. They are of global significance in biogeochemical cycles (Tréguer *et al.*, 1995) and for the functioning of most aquatic food webs. About 40% of the annual primary production can be attributed to diatoms (Nelson *et al.*, 1995) and they are considered the most successful taxa in polar environments (Wilhelm *et al.*, 2006; Arrigo *et al.*, 2010). There are an estimated 100,000 species (Round *et al.*, 1990) however, only 2,000 are genetically described (Vargas *et al.*, 2015). Diatoms can be found in almost all aquatic ecosystems ranging from the polar Oceans to the tropical regions, however, their diversity and distribution in the Atlantic Ocean is not well known as based on data from the Ocean Biogeographic Information System (OBIS) (Figure 2.1.1). The Continuous Plankton Recorder (CPR) has been recording phytoplankton distribution and diversity over several decades but mainly focused on the temperate northern hemisphere (OBIS). Other annual surveys were conducted along the Atlantic Meridional Transect (AMT) providing insight into the importance of picophytoplankton in the tropical Atlantic (Robinson *et al.*, 2006). Most of these surveys were based on microscopy or flowcytometry, however, the use of metagenomics, as in Sorcerer II (Rusch *et al.*, 2007) and Tara Ocean (Vargas *et al.*, 2015), has revolutionized our understanding

of marine microbes in the world oceans (Mock *et al.*, 2015). Mapping biogeography of marine microbes provides understanding of how phytoplankton distribution is shaped by the environment (Follows *et al.*, 2007). For example, diatom community changes in the tropical Atlantic were mainly governed by temperature (Chust *et al.*, 2013), while temperature and salinity were important in the Fram Strait (polar waters) (Hardge *et al.*, 2015). Similar results were found for other phytoplankton groups like cyanobacteria, dinoflagellates and coccolithophores (Johnson *et al.*, 2006; Chust *et al.*, 2013; Carpenter *et al.*, 1999). With the ongoing ocean warming, it becomes increasingly important to investigate and monitor distribution and abundance of main phytoplankton groups in the global ocean. Over the last decade, changes in phytoplankton communities have become more prominent and were related to climate change.

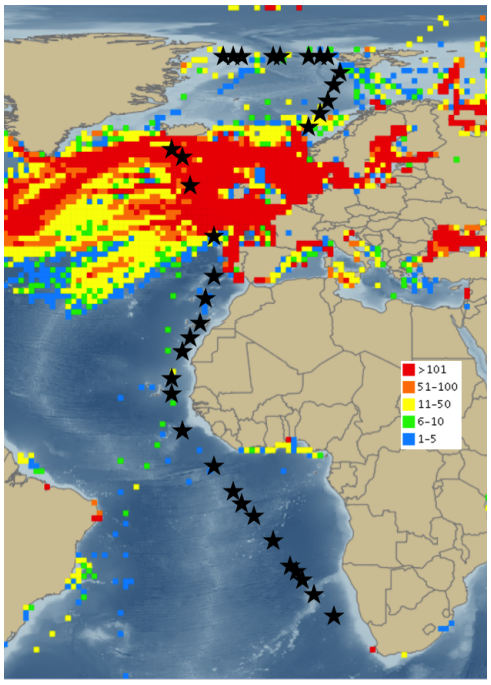

Figure 2.1.1: Abundance of Bacillariophyta in the Atlantic Ocean reaching from the Polar to the Southern Atlantic. Colour coded squares represent abundance of diatom species observed at that location. Black stars show Sea of Change stations.

Furthermore, we were able to add new knowledge about diatom distribution and abundance in the Atlantic Ocean.

Diatom community structure was investigated, using metabarcodes (V4-18S), along a longitudinal polar transect (North Atlantic Ocean) and latitudinal Atlantic transect (Atlantic Ocean) from 76°N

to 30°S. Both transects cover large environmental gradients and biological diversity and were therefore ideal for investigating environmental drivers of phytoplankton diversity and distribution.

Order level diatom metabarcodes were related to environmental metadata such as temperature, salinity and nutrient concentrations.

The Sea of Change project is a collaboration between several research groups through Europe and aims to understand the impact of climate change on phytoplankton diversity and activity in the Atlantic Ocean, as well as the potential of invasion of lower latitude phytoplankton species into the temperate and polar Atlantic Ocean. All collected samples were sequenced using 18S, 16S and ITS metabarcodes as well as metatranscriptomes and -genomes. This huge amount of data was distributed to collaborators and we focused on the 18S dataset. This 18S dataset was split into diatoms and all other autotrophs between myself and Kara Martin (PhD student in the computational department at the University of East Anglia). I chose to focus on diatoms, as they are one of the most important phytoplankton groups, contributing about 50% of the annual primary production in the ocean (Nelson *et al.*, 1995). Additionally, if our first or fourth hypothesis are confirmed, *Thalassiosira pseudonana* would be a good representative for marine diatoms to experimentally investigate their adaptive evolution to temperature changes (see chapters 3 and 4). Hypothesis to be tested were the following:

- Temperature is a major environmental driver of diatom community and distribution.
- Diatom diversity is negatively correlated to temperature.
- Diatom community differences increase with increase distance.
- Abundance and distribution of the diatom order Thalassiosirales was driven by temperature.

2.2 Material and Methods

2.2.1 Field campaign

In order to study species composition and distribution of diatoms in the Atlantic Ocean, samples have been collected during two field campaigns on board the German research vessel Polarstern

(Figure 2.2.1). The first collection was carried out during the ARK27-1 field campaign in the Arctic Ocean from the 14th of June 2012 to the 15th July 2012, starting from Bremerhaven, Germany and sailed to Svalbard and Greenland (black dots). The second set of samples was collected during the ANT29-1 campaign in the Atlantic Ocean from the 27th October to the 27th November 2012 where we sailed from Bremerhaven, Germany to Cape Town, South Africa (yellow dots). Furthermore, we received samples from a field campaign conducted by Klaas Timmermans of the Royal Netherlands Institute for Sea Research (NIOZ, Texel, The Netherlands) (Stratiphyt, red dots), sailing from the Canary Islands to Iceland in April 2011.
Water samples were collected either at the chlorophyll maximum and/or surface, in 12 L Niskin bottles by a Niskin bottle rosette sampler, with an attached Sonde (CTD, conductivity, temperature, depth). As soon as the rosette sampler was back on board, water samples were immediately transferred into plastic containers and transported to the laboratory as quick as possible. All sampled stations are accompanied by metadata (salinity, temperature, collection depth and nutrient concentration).

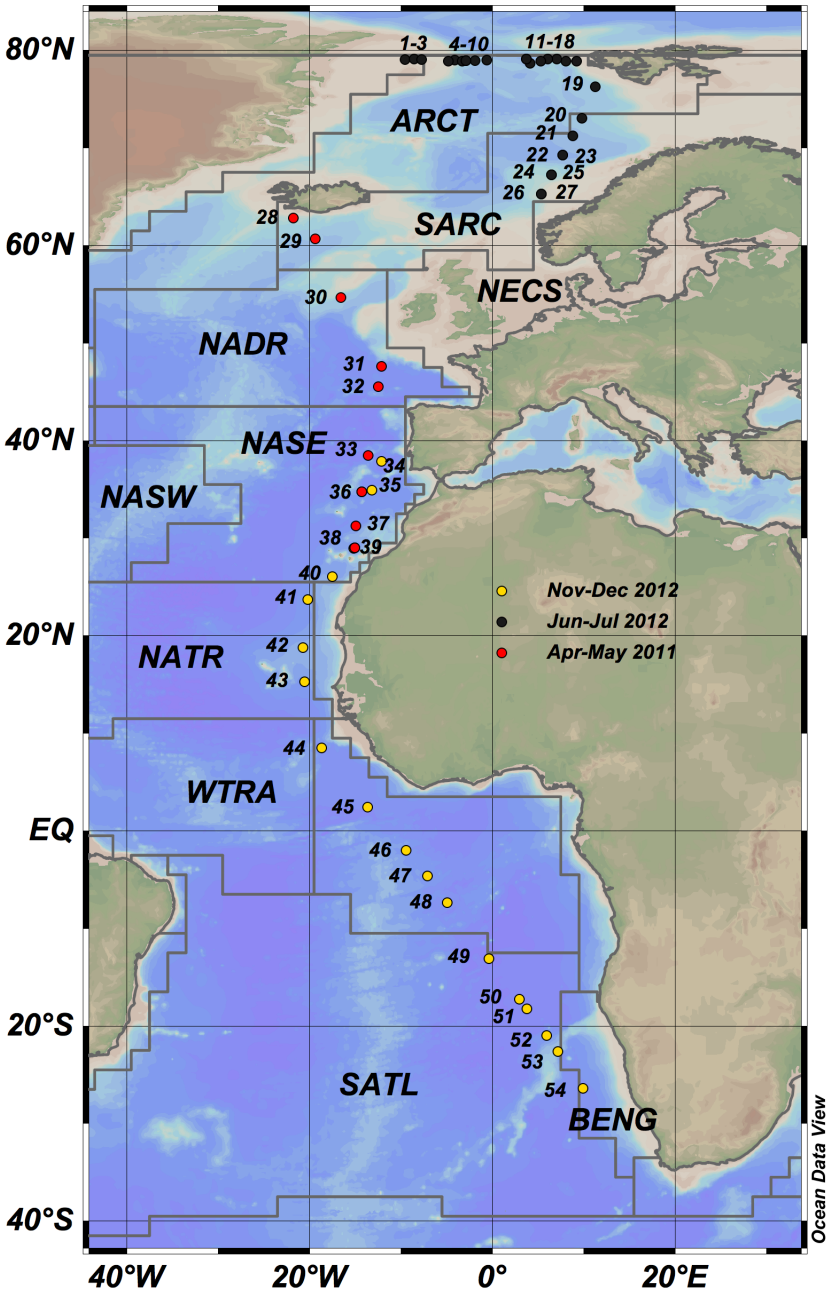

Figure 2.2.1: Map of sampled stations during three field campaigns (ANT (black dots), Stratiphyt (red dots) and ARK (orange dots)). Oceanic provinces defined by Longhurst (2007) are included. SATL=South Atlantic Gyral Province, WTRA=Western Tropical Atlantic Province, NATR=North Atlantic Tropical Gyre Province, ETRA=Eastern Tropical Atlantic Province, NASE=North Atlantic Subtropical Gyral East Province, NASW=North Atlantic Subtropical Gyral West Province, NADR=North Atlantic Drift Province, ARCT=Atlantic Arctic Province

2.2.1.1 Sampling

Water samples were pre-filtered with a 100 µm mesh to remove bigger zooplankton, and distributed into 1.25 L plastic bottles. Bottles were mounted onto prepared filtration racks and all filtrations were conducted under low pressure (<150 mbar Hg) and in under 30 minutes.
Samples for DNA extractions were filtered onto 25 mm 1.2 µm polycarbonate filters (Isopore membrane, Millipore, MA, USA). All filters were snap frozen in liquid nitrogen and stored at -80°C until further analysis.
Duplicate samples for dissolved phosphorus and nitrogen analysis were sterile filtered with 0.2 µm nitrate-cellulose syringe filters into 15 mL tubes and stored at -20°C. Dissolved silicate samples were also filtered with 0.2 µm nitrate-cellulose syringe filters into 15 mL tubes but stored at 4°C to avoid silica precipitation due to freezing.

Stratiphyt cruise samples Klaas Timmermans (The Royal Netherlands Institute for Sea Research, Texel, The Netherlands) and Willem van de Poll (Rijksuniversiteit Groningen, Groningen, The Netherlands) kindly provided one filter for each of the 22 stations sampled between the Canary Islands and Iceland. These samples were taken by a Niskin bottle rosette sampler with attached CTD and filtered onto 47 mm 0.2 µm polycarbonate filters without pre-filtration and snap frozen in liquid nitrogen. Samples were stored at -80°C until extraction. Accompanying environmental data was provided by Willem van de Poll.

2.2.2 DNA extractions

DNA extractions ARK and ANT field campaign DNA samples were extracted with the EasyDNA Kit (Invitrogen, Carlsbad, CA, USA) with some adjustments. Cells were washed off the filter with pre-heated (65°C) Solution A (EasyDNA Kit, Invitrogen, Carlsbad, CA, USA) and the supernatant was transferred into a new tube with one small spoon of glass beads (425-600 µm, acid washed) (Sigma-Aldrich, St. Louis, MO, USA). The samples were then vortexed three times in intervals of 3 seconds to break the cells. RNase A was added to the samples and incubated for 30 min at 65°C. The supernatant was transferred into a new tube and Solution B (EasyDNA Kit, Invitrogen, Carlsbad, CA, USA) was added followed by a chloroform phase

separation and an ethanol precipitation. DNA was pelleted by centrifugation and washed several times with isopropanol, air dried and suspended in 100 µL TE buffer. DNA concentration was measured with a Nanodrop (Thermo Fisher Scientific, Waltham, MA, USA), samples snap frozen in liquid nitrogen and stored at -80°C until sequencing.

DNA extractions Stratiphyt samples As we received only one filter per station from the Stratiphyt field campaign, we had to opt for a different DNA extraction method as RNA was also needed for the Sea of Change project. Thus, Stratiphyt samples were extracted with the ZR-Duet™DNA/RNA MiniPrep kit (Zymo Research, Irvine, USA) allowing simultaneous extraction of DNA and RNA from one sample filter. Cells were washed from the filters with DNA/RNA Lysis Buffer (ZR-Duet™DNA/RNA MiniPrep kit, Zymo Research, Irvine, USA) and one spoon of glass beads (425-600 µm, Sigma-Aldrich, MO, USA) was added. Samples were vortexed quickly and loaded onto Zymno-Spin™IIIC columns. The columns were washed several times with DNA/RNA Wash Buffer (ZR-Duet™DNA/RNA MiniPrep kit, Zymo Research, Irvine, USA) and DNA was eluted in 60 µL DNase/RNase-free water. Purified DNA samples were measured with a Nanodrop spectrophotometer. Samples were snap frozen in liquid nitrogen and stored at -80°C until sequencing.

2.2.3 DNA sequencing

All extracted DNA samples were sequenced by the Joint Genome Institute (JGI) (Department of Energy, Walnut Creek, CA, USA). Amplicon sequencing was performed with V4-18S region specific primers (Zimmermann *et al.*, 2011) on a Illumina MiSeq instrument with a 2x250 bp reads configuration. This region has been shown to be conserved enough for class/species identification. Sequences were quality controlled, consisting of scanning for contaminations (for example, Illumina adapter sequences and primer sequences) and quality trimming of reads. A total of 54 samples passed quality control after sequencing and on average 128,692 reads per sample were provided.

2.2.4 18S sequence analysis

The bioinformatical analysis of 18S rDNA sequences was done by Kara Martin (School of Computational Science, UEA) with feedback from Andrew Toseland, Thomas Mock, Vincent Moulton and myself. The final output (diatom order count matrix) was then used by myself for ecological analysis.

2.2.4.1 Building of reference tree

Taxonomic identification of our 18S rDNA sequences was done by alignment to a reference tree. This reference tree was built by selection of 1663 high quality sequences of algae and outgroup taxa (Opisthokonta, Cryptophyta, Glaucocystophyceae, Rhizaria, Stramenopiles, Haptophyceae, Viridiplantae, Alveolata, Amoebozoa and Rhodophyta) from the SILVA (SSU) and Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP) database. Closely related sequences were removed to reduce computational costs by clustering reference sequences with CD-HIT using a similarity threshold of 97%. Multiple sequence alignments were optimized in ClustalW by iteration (Thompson *et al.*, 1994) and the resulting alignments were cleaned with TrimAL. Our query sequences were placed onto the reference using pplacer (Matsen *et al.*, 2010), in combination with a RAxML statistics output, providing a description of the phylogenetic model and the likelihood of read placement (cutoff ML 0.5). Annotation of sequences at a high taxonomic level is only possible if a reference sequence is available. In our case, we used 161 diatom species sequences, thus, many sequences could only be annotated reliably at a higher taxonomic level.

2.2.4.2 18S copy number normalization of reads

18S rDNA copy numbers vary greatly between eukaryotes and in order to account for this variation, genome sizes of 186 representatives across the eukaryote tree were related to 18S rDNA copy number (Figure 2.2.2). Genome sizes were not available for all our reference species, therefore, genome sizes of closely related species were averaged.

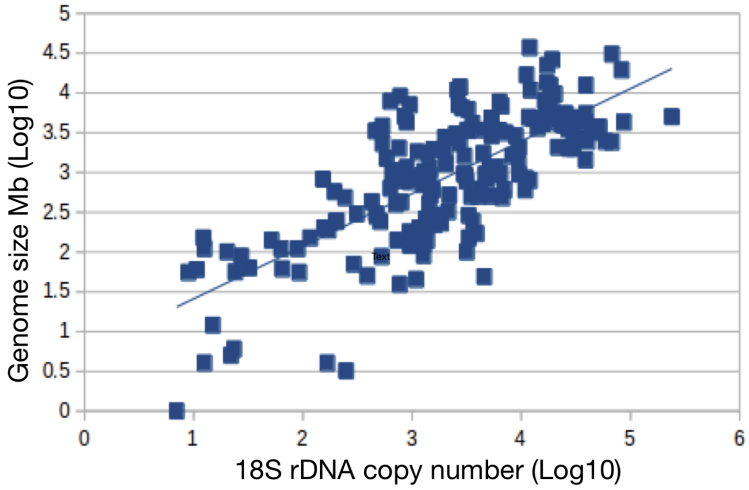

Figure 2.2.2: 18S copy number vs genome size of 186 species showed a significant correlation (pearson coefficient 0.74, p-value <2.2e-16). This resulted in the following equation: 18S copy number = 0.66* genome size + 0.75.

The equation for 18S copy number normalization was generated by correlating (Pearson) genome size and 18S rDNA copy number (Pearson coefficient 0.74). This provided the following equation (R^2 0.74, p-value <2.2e-16):

$$\text{18S copy number} = 0.66 * (\text{genome size}) + 0.75 \quad (1)$$

Normalization to copy number was rounded to full number, and normalized to million reads. Normalization to 18S copy number provided diatom order reads per station being equivalent to diatom order abundance per station. This was then normalized to million reads to account for the different library sizes. The resulting data matrix was used for subsequent community analysis providing relative diatom abundance as diatom order reads per million reads.

2.2.5 Statistical analysis

The aim was to identify environmental drivers of diatom diversity and distribution. Thus, diatom diversity, based on diatom orders, was related to environmental gradients of temperature, salinity, nitrogen, phosphate and silicate across the two transects (Polar and Atlantic) of the Atlantic Ocean.

Out of 54 sampled stations, 53 provided diatom orders. Due to missing metadata, station 13

was deleted from the dataset. All statistical analysis were conducted in the R environment (R Development Core Team 2012).

Diatom order abundance (reads per million reads) was Hellinger transformed to give less weight to highly abundant orders.

2.2.5.1 Co-correlation of environmental factors

Co-correlation of environmental covariates of the polar transect was assessed by Generalized Variance Inflation Factors (GVIF), calculated with the HighstatLibV6 package and the corvif function `http://www.highstat.com/Book2/HighstatLibV6.R`. Sequential dropping of the covariate with the highest GVIF was repeated until all remaining covariates had GVIFs smaller than 3, as suggest by Zuur et al. 2010.

Polar transect A summary of the GVIFs and sequentially dropped covariates of the Polar transect is shown in Table 2.2.1. Covariates measuring sampling depth, temperature, phosphate and silicate were included for further downstream analysis.

Table 2.2.1: Generalized Variance Inflation Factors for polar transect covariates.

Covariate	Full set	Salinity dropped	Nitrate+Nitrite dropped
sampling depth (m)	1.79	1.79	1.37
Temperature (°C)	5.71	1.96	1.36
Salinity (PSU)	6.36	-	-
Nitrate+Nitrite [$\mu molL^1$]	4.17	3.76	-
Phosphate [$\mu molL^1$]	2.33	2.28	1.62
Silicate [$\mu molL^1$]	1.22	1.19	1.18

Atlantic transect A summary of the GVIFs and sequentially dropped covariates of the Atlantic transect is shown in Table 2.2.2. Covariates measuring temperature, salinity, nitrate+nitrite and silicate were included for further downstream analysis.

Table 2.2.2: Generalized Variance Inflation Factors for Atlantic transect covariates.

Covariate	Full set	Phosphate dropped	Sampling depth dropped
sampling depth (m)	5.29	5.18	-
Temperature (°C)	4.25	3.09	2.08
Salinity (PSU)	4.37	3.43	1.9
Nitrate+Nitrite [$\mu molL^1$]	4.41	2.33	2.29
Phosphate [$\mu molL^1$]	7.17	-	-
Silicate [$\mu molL^1$]	3.03	2.92	2.91

2.2.5.2 Diatom abundance in relation to environmental factors

Initial scatter plots of response variables and covariates did not show a strong non-linear pattern and therefore redundancy analysis (RDA) (Legendre and Legendre, 1998) was used to model phytoplankton community composition as a function of environmental variables. A backward selection procedure was used to select explanatory variables that explain significant amount of variation, while non-significant (p-value$>$0.05) terms were removed. Multivariate pseudo-F-value statistic was used for permutation tests of significance with a total of 9999 permutations to estimate p-values (Zuur et al., 2009). Variance explained by the significant terms was calculated by starting with a single term model and sequential adding of model terms and calculating the additional variance explained. This is only accurate when terms are added in the order they appear in the model. Arrows of model terms in a RDA plot can be interpreted by their length and angle to the axis and to each other (Figure 2.2.3). The length of the arrow corresponds to the strength of the underlying gradient (e.g. long arrow strong gradient, short arrow smaller gradient). The angle to the axis shows how well they correlate with the axis and the smaller the angle the higher the correlation. Angles between the arrows points towards correlation, where a small angle represents a high positive correlation, 180°high negative correlation and a 90°angle points towards no correlation.

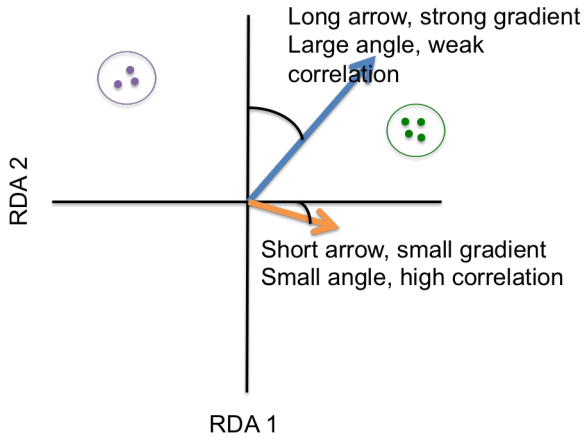

Figure 2.2.3: Stylised RDA plot to explain plot interpretation. The length of the arrow corresponds to the strength of the underlying gradient (e.g. long arrow strong gradient, short arrow smaller gradient). The angle to the axis shows how well they correlate with the axis and the smaller the angle the higher the correlation. Angles between the arrows points towards correlation, where a small angle represents a high positive correlation, 180°high negative correlation and a 90°angle points towards no correlation. Clusters of points that are close together suggest higher similarity to each other than points that are further apart.

2.2.5.3 Richness in relation to environmental factors

Species richness was measured by counting present diatom orders per station. This richness was related to the selected environmental covariates by backwards model building. A general-linearised model (glm) was built with all covariates and investigated for the least significant covariate. This covariate was dropped from the model and the model rerun. This was repeated until all remaining covariates were significant at the 0.05 level.

2.2.5.4 Bray-Curtis dissimilarity in relation to environmental factors and spatial distance

The contribution of environmental factors and geographic distance to diatom diversity was estimated using similarity matrices and Mantel tests (Manly, 1998) across sites of the polar and Atlantic transect. The Mantel test is a non-parametric test to determine how frequently the observed similarity between matrices would arise by chance and calculates a statistic r_m which measures the correlation between the matrices. The dissimilarity between any two diatom communities was estimated based on a Hellinger transformed relative diatom order abundance matrix using the vegdist function of the vegan package (Oksanen *et al.*, 2016). The Bray-Curtis dissimilarity index (B) (Bray and Curtis, 1957) was calculated based on the following equation:

$$B = \frac{\sum |x_{ij} - x_{ik}|}{\sum x_{ij} + x_{ik}} \tag{2}$$

The Bray-Curtis index ranges from 0 to 1 with highly dissimilar communities having an index of 1, and very similar communities an index of 0.

The Bray-Curtis index matrix was related to temperature by correlating the index matrix to the temperature matrix with a Mantel test (Mantel correlation coefficient r_m) (vegan package, (Oksanen *et al.*, 2016)). 999 permutations were calculated to obtain the significance of the Mantel statistic.

Additionally, a distance matrix of sampled stations of both transects was calculated in R with the geosphere package and the distm function. Correlation between the distance matrix and beta diversity was calculated with a mantel statistic (Mantel correlation coefficient r_m), part of the vegan package (Oksanen *et al.*, 2016). Significance of the Mantel statistic was obtained with 999 permutations.

2.3 Results

Sequencing of the V4 18S rDNA region resulted in on average 126,201 reads per sample. After normalisation to 18S rDNA copy number, sequences were analysed to order level resulting in 19 orders (out of 44 known orders), of which Coscinodiscales (56.56%), Leptocylindrales (9.48%) and Thalassiosirales (8.54%) were most abundant (Table 2.3.1).

Table 2.3.1: Identified Bacillariophyta orders and reads per million reads summary across all stations

Order	**reads per million reads**	**% of total**
Coscinodiscales	29.977.601	56.56
Leptocylindrales	5.023.267	9.48
Thalassiosirales	4.526.524	8.54
Corethrales	3.675.214	6.93
Hemiaulales	1.863.028	3.52
Fragilariales	1.848.725	3.49
Chaetocerotales	1.602.128	3.02
0viculales	1.174.194	2.22
Achnanthales	1.070.900	2.02
Rhizosoleniophycidae	836.811	1.58
Aulacoseiraceae	615.283	1.16
Biddulphiales	320.449	0.60
Melosirales	283.614	0.54
Bacillariales	149.190	0.28
Thalassiophysales	16.755	0.03
Surirellales	8.541	0.02
Striatellales	6.104	0.01
Climacospheniales	1.673	0.003

To clarify the influences of variables relating to latitudinal vs longitudinal changes on Bacillariophyta diversity and distribution, the dataset was split into two transects, one from Greenland to Spitsbergen (Longitudinal transect, polar transect) and into a Atlantic transect (Latitudinal transect) from Spitsbergen to South Africa. Metadata such as temperature, salinity and nutrient concentrations (dissolved nitrate, phosphate and silica) were measured along both transects.

2.3.1 Polar transect

2.3.1.1 Metadata

The Fram Strait, between Greenland and Spitsbergen, is a transition zone between the northern North Atlantic and the Arctic Ocean. It represents the only deep connection between the Arctic and Atlantic Ocean (Klenke and Schenke, 2002) and can be separated into three hydrographic regions. The West Spitsbergen Current (WSC) is fed by the northern warm Atlantic Current and flows northward west of Svalbard into the Arctic Ocean. The WSC transports warm (2.7°C to 8°C), salty (about 35 PSU) water into the Arctic Ocean (Figure 2.3.1).

On the west side of the Fram Strait, flowing southward is the East Greenland Current (EGC) characterized by cold (-2°C to 1°C) and fresh (30-33 PSU) water masses coming from the Arctic Ocean. The middle of the Fram Strait is formed by the Returning Atlantic Current (RAC) which is a branch of the WSC. The sampled temperature gradient ranged from 0°C, in the west, to 8°C in the east (Figure 2.3.1,a). The salinity gradient ranged from 31 PSU, in the west, to 35 PSU in the east (Figure 2.3.1,b).

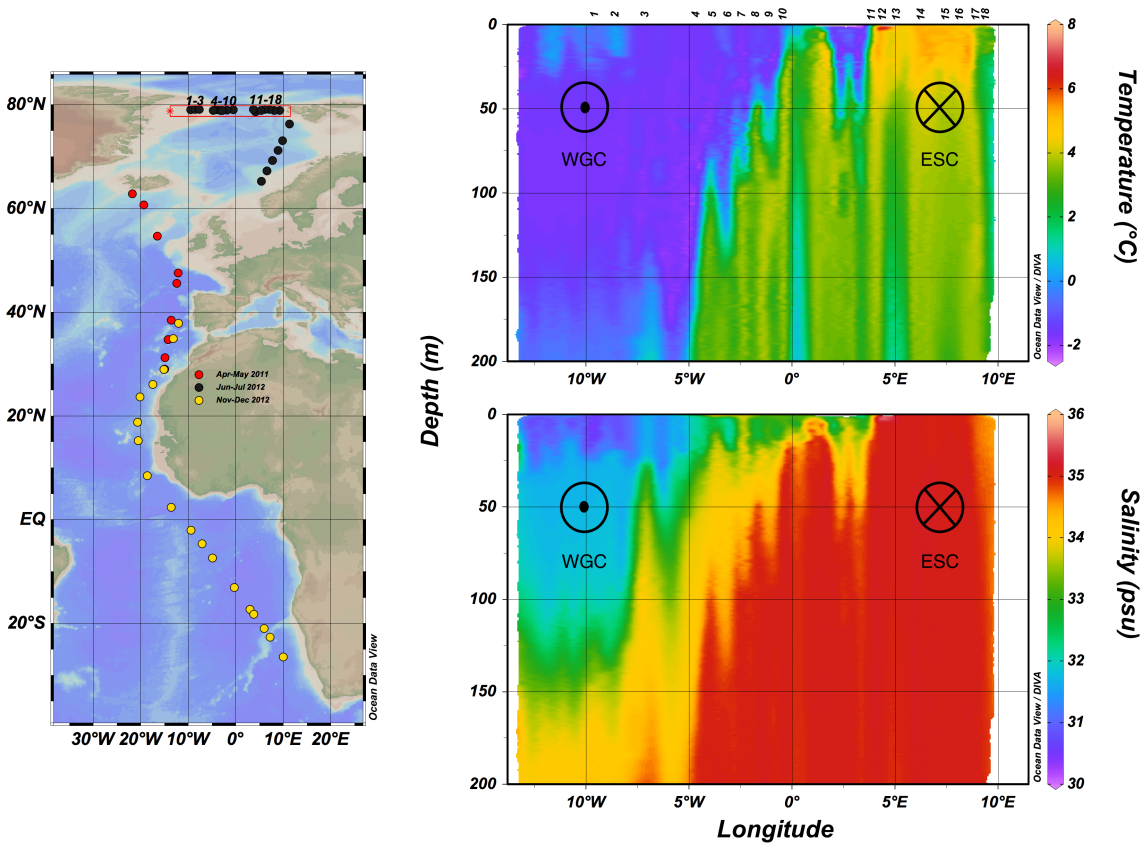

Figure 2.3.1: Temperature (top) and Salinity (bottom) profiles for the Fram Strait in Summer 2012.

Nutrient concentrations of polar transect An overview of measured temperature, salinity and nutrient concentrations (NO_3+NO_2, PO_4 and SiO_3) is given in Table 2.3.2 for the 18 stations sampled in the polar transect. The lowest temperature was measured at station 5 with -1.62°C while the highest temperature occurred at station 17 with 5.27°C. Salinity varied from 31 PSU at the most western stations (1-3) and increased to 35 PSU at the most eastern stations (13-18). Phosphate concentrations remained relatively stable through out the transect, ranging from 0.46 µmol L^{-1} to 1.07 µmol L^{-1}. Peak nitrogen concentrations were measured at station 7 and 16 with about 10 µmol L^{-1} and were lowest at the most western stations, close to Greenland, with close to 0 µmol L^{-1}. Silica concentrations were similar across stations and ranged from 2.16 µmol L^{-1} to 8.97 µmol L^{-1}. Phytoplankton community samples were usually taken at the chlorophyll maximum within the first 25 m of the water column with the exception of station 7 which was taken at 110 m because of a second chlorophyll a maximum.

Table 2.3.2: Metadata summary of all sampled stations in the Polar transect.

Station	Longitude (dec degrees)	Latitude (dec degrees)	sampling Depth (m)	Temperature (°C)	Salinity (PSU)	NO_3+NO_2 (µmol L^{-1})	PO_4 (µmol L^{-1})	SiO_3 (µmol L^{-1})
1	-9.52	79.02	17	-1.03	31.03	0.00	0.47	2.48
2	-8.52	79.08	26	-1.51	31.71	0.14	0.95	2.86
3	-7.67	79.04	20	-1.46	31.33	0.00	0.62	7.17
4	-4.09	79.00	10	-1.51	32.95	0.81	0.58	5.49
5	-4.79	78.86	20	-1.62	32.20	0.60	0.74	8.97
6	-3.23	78.87	15	-0.88	32.35	3.30	0.79	5.47
7	-2.84	78.90	110	3.08	34.99	10.44	1.05	4.46
8	-2.84	78.90	25	-1.37	33.34	0.70	0.46	3.03
9	-1.84	78.94	16	4.40	35.06	4.77	0.84	3.26
10	-0.56	78.98	10	1.37	33.53	0.69	0.50	3.64
11	3.74	79.07	10	0.05	33.59	2.80	0.61	2.49
12	3.74	79.07	5	-0.30	33.15	2.21	0.51	3.12
13	4.15	78.62	15	5.87	35.05	NA	NA	NA
14	5.33	78.84	15	4.49	35.10	3.32	0.73	3.95
15	6.11	79.09	7	4.82	35.11	3.90	0.76	3.88

Continued on next page

Continued from previous page

Station	Longitude (dec degrees)	Latitude (dec degrees)	sampling Depth (m)	Temperature (°C)	Salinity (PSU)	NO_3+NO_2 (µmol L^{-1})	PO_4 (µmol L^{-1})	SiO_3 (µmol L^{-1})
16	7.08	79.07	18	4.56	35.09	9.69	1.07	4.31
17	8.11	78.87	10	5.27	35.07	0.51	0.49	2.75
18	9.23	78.85	25	2.34	34.76	2.93	0.94	2.16

2.3.1.2 18S analysis

Diatom order richness was correlated to environmental covariates with temperature (Figure 2.3.2) and silicate (Figure 2.3.3, together explaining 62.52% of the variation across the Fram Strait. For visualization purposes, temperature and silicate were plotted separately.

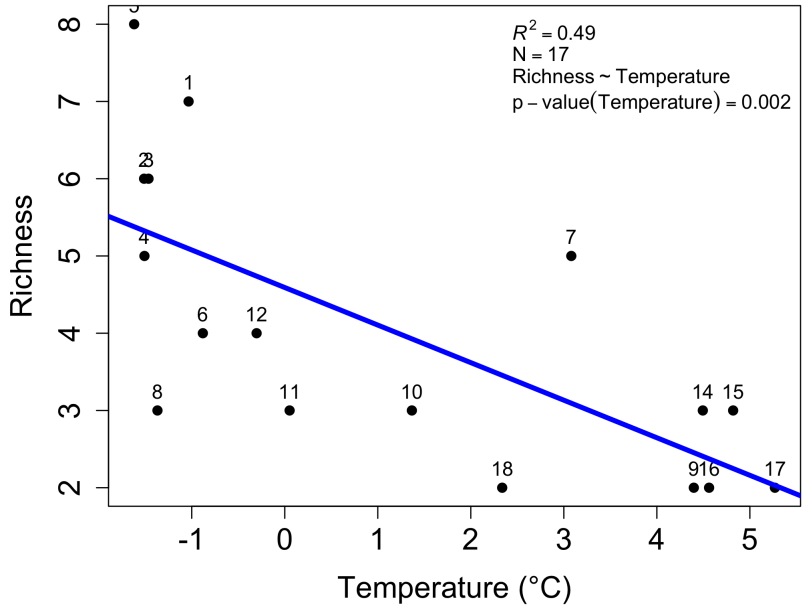

Figure 2.3.2: Richness (y-axis) was negatively correlated (-0.48, R^2=0.49, N=17) to temperature (°C) (x-axis), meaning that stations with higher temperature had a less diverse diatom community.

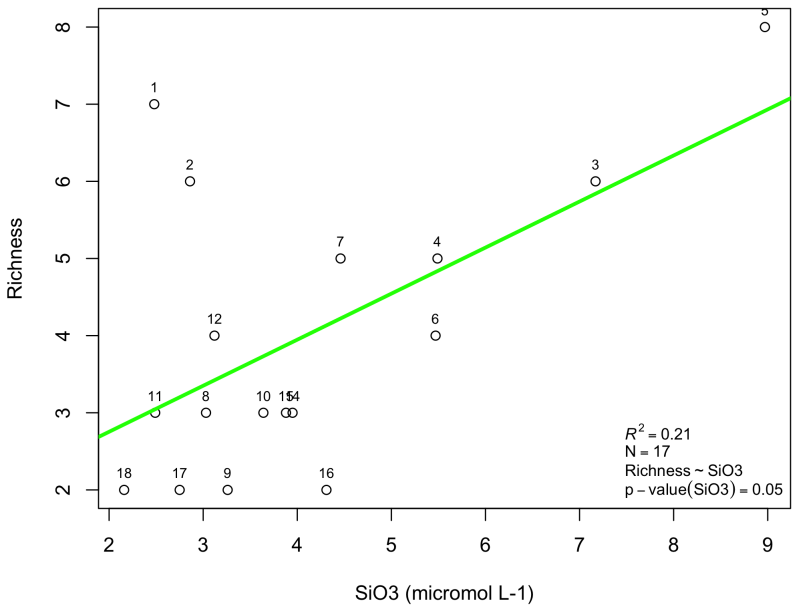

Figure 2.3.3: Richness (y-axis) was positively correlated (0.59, R^2=0.21, N=17) to silicate (°C) (x-axis), meaning that stations with higher silicate concentrations had a more diverse diatom community.

Diatom communities differed between the West and East part of the transect (Figure 2.3.4). Stations 1-7 were dominated by Thalassiosirales, Chaetocerotales and Naviculales. Coscinodiscales dominated the overall transect (65%). Other orders were of minor importance along the transect and showed very low relative abundance levels. Furthermore, the abundance pattern suggest a higher order diversity at the EGC compared to the WSC. Surprisingly, Melosirales and Fragilariales, typical polar species, were absent in this transect.

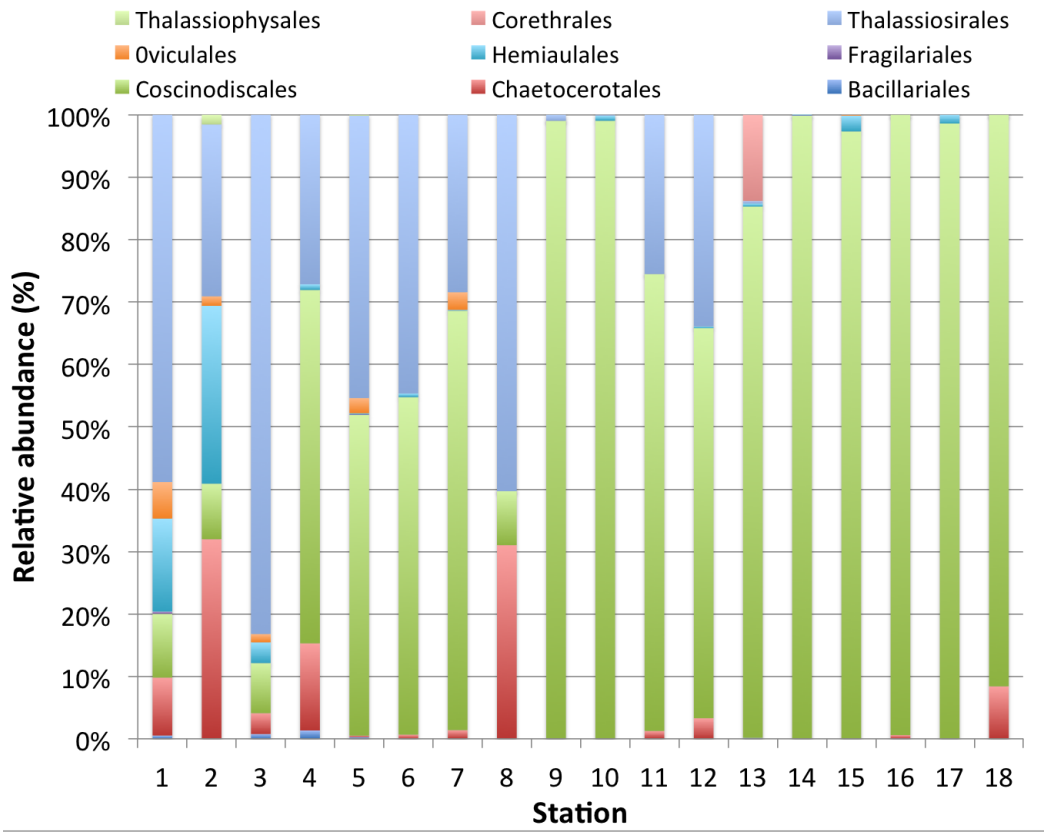

Figure 2.3.4: Relative diatom order abundance of the polar transect ordered by station/longitude.

The abundances of diatom orders over the polar transect was analysed using a distance matrix based on the Bray-Curtis dissimilarity index. This distance matrix was used to visualise clustering between stations. The dendrogram shows three clusters (Figure 2.3.5). Two main clusters were identified, a WSC (red) and a EGC (blue) cluster. Interestingly, a few WSC stations (6,7) cluster with several EGC stations (4,5,11,12).

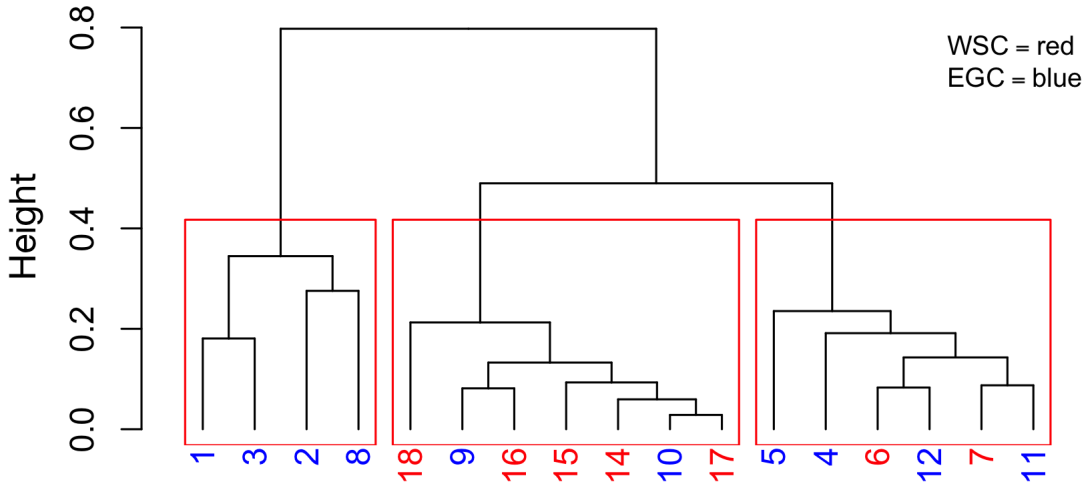

Figure 2.3.5: Beta diversity dendrogram based on bray-curtis dissimilarties and method complete.

Environmental correlation In order to identify the dominant environmental parameters shaping diatom diversity and abundance, diversity and abundance matrices were correlated with environmental metadata. Co-correlation analysis of the measured metadata is shown in Figure 2.3.6 with correlation coefficients in the upper panels. A strong co-correlation was found for temperature with salinity (R^2=0.9) and are thus interchangeable, and for nitrogen with phosphate (R^2=0.73). Based on variance inflation factor (VIF) analysis, salinity and nitrogen were excluded from the metadata.

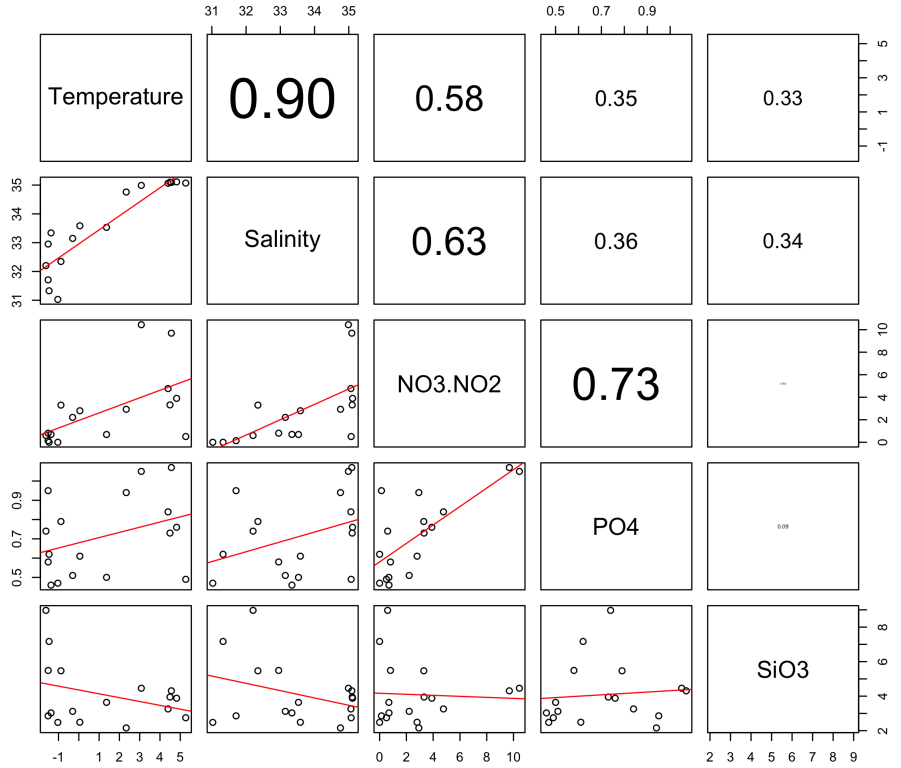

Figure 2.3.6: Test for co-correlation of environmental data.

In contrast, diatom abundance pattern between stations were significantly influenced by temperature and salinity (Figure 2.3.7) with temperature explaining slightly more variation, while nutrients showed no significant effect. Axis 1 (RDA1) explained 35.41% of the variation and was highly correlated with temperature (correlation coefficient 0.99). Note that temperature and salinity showed strong co-correlation and were thus interchangeable, with salinity explaining 35.37% of the variation (Appendix, Figure 6.1.2).

EGC stations 1-3 and 5 were found to cluster at low salinity and temperature while WSC stations 10,14,15 and 17 were associated with higher salinity and temperature.

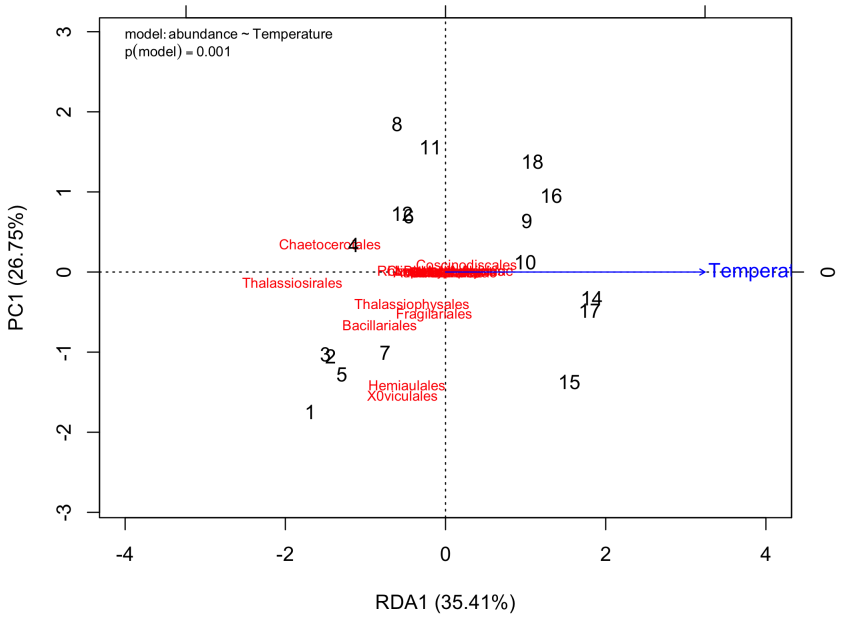

Figure 2.3.7: Correlation of environmental data and diatom diversity.

Figure 2.3.8 provides the correlation (Spearman correlation) of diatom order abundance (Hellinger transformed) with temperature. Coscinodiscales showed strong positive correlation with temperature (correlation coefficient=0.84, p-value <0.001) while Thalassiophysales, Thalassiosirales, Chaetocerotales, Bacillariales showed negative correlation with temperature (-0.56, 0.02; -0.79, <0.001; -0.69, 0.002; -0.61, 0.009, respectively).

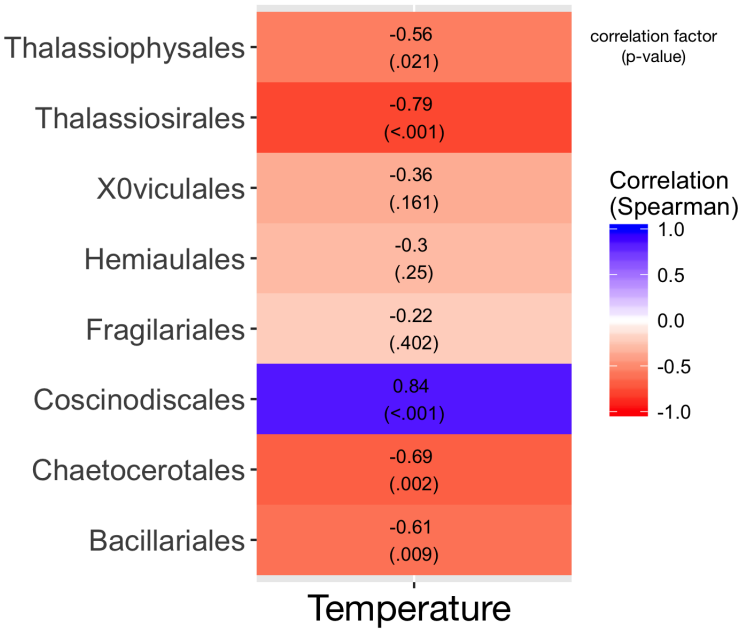

Figure 2.3.8: Correlation (Spearman correlation) of diatom order abundance (Hellinger transformed) with temperature. Spearman correlation factor is color coded with blue showing strong positive correlation and red strong negative correlation. Values in the boxes provide the correlation coefficient and the p-value in brackets underneath.

The bray-curtis index across the polar transect was significantly related to temperature (mantel test, r_m=0.59, p-value=0.001, 999 permutations). This correlation was still significant after correcting for the auto-correlation of temperature and distance (partial mantel test, r_m=0.47, p-value=0.002, 999 permutations). In addition, distance had a significant effect on beta-diversity (mantel test, r_m=0.53, p-value=0.001, 999 permutations, which increased with increasing distance (Figure 2.3.9).

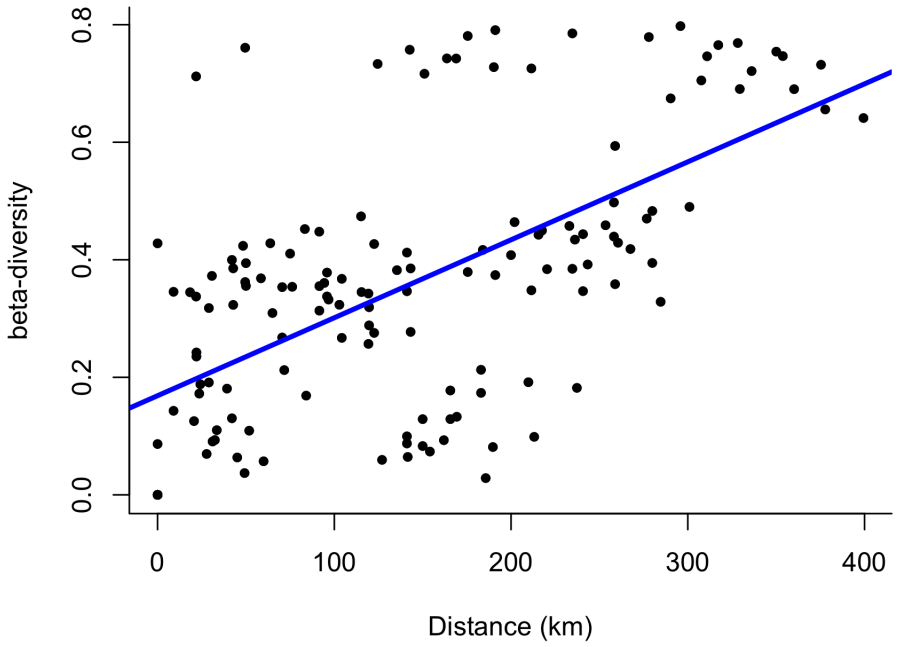

Figure 2.3.9: Correlation of diatom bray-curtis index with distance (km). The bray-curtis index increased with increasing distance (r_m=0.53, p-value=0.001).

2.3.2 Atlantic Section

2.3.2.1 Metadata

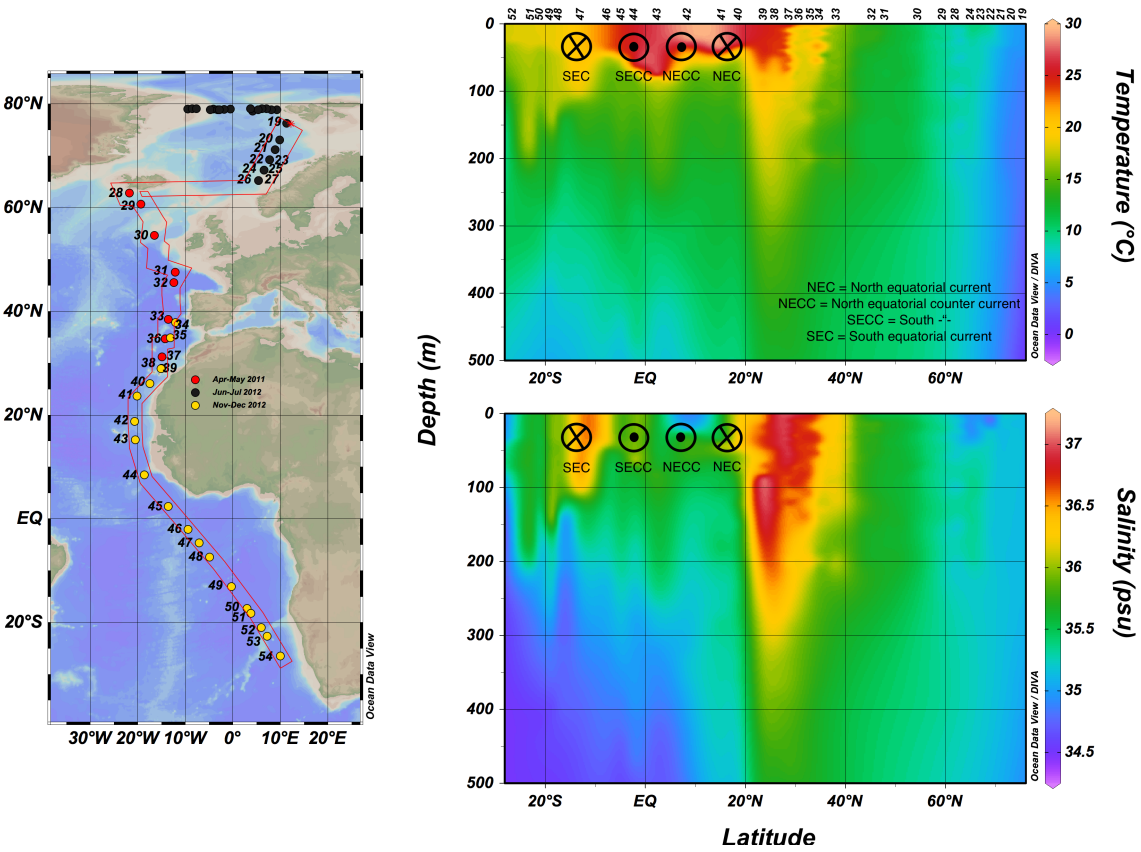

Figure 2.3.10: Temperature and Salinity profiles for the Atlantic transect from Svalbard, Norway to Cape Town, South Africa.

The Atlantic transect was sampled in the Norwegian Sea as well as the Atlantic from Iceland to Cape Town, South Africa. The temperature at the sampled stations ranged from 5°C in the Norwegian Sea to 30°C at the Equator and 20°C close to the South African coast (Figure 2.3.10, a). The deepening of the 18°C isotherm corresponds to the region where the transect crosses the North Atlantic gyre (yellow deep patch between 30 and 20°N) and relates to the up-welling of the NW Africa. Salinity ranged from 35 PSU in the North Atlantic to 37 PSU at the Equator and 35 PSU close to the South African coast (Figure 2.3.10, b). The transect passed through several currents at the equator, transporting water either east to west or west to east. Between 40°North and 20°North we passed through a patch of salty, warm water which was followed by the North equatorial current (NEC). This current transports fresh (35 PSU) surface water from the

east to the west of the Atlantic and is mainly driven by wind. Just above the equator is the North equatorial counter current (NECC) which transports surface water from the west to the east and is also wind driven. Below the equator, between 0°and 10°South is the South equatorial counter current (SECC) followed by the further south South equatorial current (SEC).

Nutrient concentrations Atlantic section An overview of measured temperature, salinity and nutrient concentrations (NO_3+NO_2, PO_4 and SiO_3) is given in Table 2.3.3 for the 35 stations sampled in the Atlantic transect. Sampling depth of each station increased from the North Atlantic (about 20 m) to the equator (about 60 m) and decreased from the equator to South Africa (about 30 m). Phosphate concentrations were low through out the Atlantic transect with concentrations below 1 µmol L^{-1}. Silica concentrations were patchy with highest concentrations (1.52 - 6.69 µmol L_{-1}) at stations 19-32 in the North Atlantic and ten-fold lower concentrations at stations 33-41. Another patch of higher silica concentrations was measured 20°North to 7°South. Similar trends were observed for nitrogen concentrations. The Atlantic transect showed a great temperature gradient from the North Atlantic (about 8°C) to the equator (about 28°C) but less variation in salinity (35 PSU to 36 PSU) in comparison to the Arctic transect (see above).

Table 2.3.3: Metadata summary of all sampled stations in the Atlantic transect

Station	Longitude (dec degrees)	Latitude (dec degrees)	sampling Depth (m)	Temperature (°C)	Salinity (PSU)	NO_3+NO_2 (µmol L^{-1})	PO_4 (µmol L^{-1})	SiO_3 (µmol L^{-1})
19	11.31	76.25	15	5.53	35.15	6.26	0.83	4.08
20	9.86	73.02	20	6.02	35.15	6.55	0.88	3.90
21	8.87	71.20	10	7.18	35.13	4.83	0.77	3.62
22	7.73	69.23	10	9.10	34.73	2.46	0.46	1.52
23	7.73	69.23	5	8.75	34.87	0.86	0.49	1.68
24	6.53	67.23	15	8.94	35.09	4.16	0.74	2.48
25	6.53	67.23	20	8.68	35.10	2.38	0.66	2.12
26	5.42	65.25	20	9.24	34.93	2.85	0.60	2.44
27	5.42	65.25	5	9.75	34.94	1.58	0.59	2.35
28	-21.74	62.80	10	7.69	35.20	13.35	0.86	6.69
29	-19.34	60.68	10	8.75	35.25	11.75	0.76	6.57
30	-16.51	54.63	10	10.79	35.45	7.79	0.49	1.59
31	-12.11	47.57	40	12.07	35.64	5.92	0.41	2.54
32	-12.43	45.53	16	13.96	35.81	0.33	0.05	2.18
33	-13.58	38.42	67	15.90	36.26	0.52	0.05	0.29

Continued on next page

Continued from previous page

Station	Longitude (dec degrees)	Latitude (dec degrees)	sampling Depth (m)	Temperature (°C)	Salinity (PSU)	NO_3+NO_2 (µmol L^{-1})	PO_4 (µmol L^{-1})	SiO_3 (µmol L^{-1})
34	-12.09	37.83	80	20.63	36.44	0.43	0.07	0.48
35	-13.14	34.88	80	21.78	36.65	0.27	0.00	0.29
36	-14.26	34.72	48	16.93	36.42	1.02	0.03	0.29
38	-15.00	29.00	81	18.97	36.83	0.02	0.02	0.75
39	-15.15	28.94	90	23.00	36.60	0.54	0.08	0.75
40	-17.46	26.05	80	24.62	36.90	0.52	0.07	0.41
41	-20.18	23.69	60	25.25	36.57	1.95	0.19	0.69
42	-20.70	18.76	45	26.78	36.17	6.13	0.33	1.19
43	-20.52	15.25	55	28.60	35.64	16.42	0.68	3.37
44	-18.62	8.47	46	29.02	34.88	3.11	0.24	2.59
45	-13.60	2.41	80	27.10	35.61	6.46	0.37	2.06
46	-9.42	-2.05	63	25.64	35.94	10.03	0.52	2.62
47	-7.06	-4.67	47	24.71	35.83	11.32	0.56	3.07
48	-4.90	-7.39	45	23.27	36.23	6.48	0.37	2.02
49	-0.32	-13.10	75	20.66	36.50	4.40	0.24	0.83

Continued on next page

Continued from previous page

Station	Longitude (dec degrees)	Latitude (dec degrees)	sampling Depth (m)	Temperature (°C)	Salinity (PSU)	NO_3+NO_2 (µmol L^{-1})	PO_4 (µmol L^{-1})	SiO_3 (µmol L^{-1})
50	2.98	-17.28	30	18.82	35.97	0.00	0.19	2.63
51	3.80	-18.25	43	18.77	35.99	8.01	0.44	0.86
52	6.00	-20.99	30	18.52	35.62	5.95	0.39	0.07
53	7.19	-22.64	40	18.48	35.70	0.50	0.04	0.43
54	9.98	-26.44	30	18.20	35.20	2.28	0.17	0.69

2.3.2.2 18S analysis

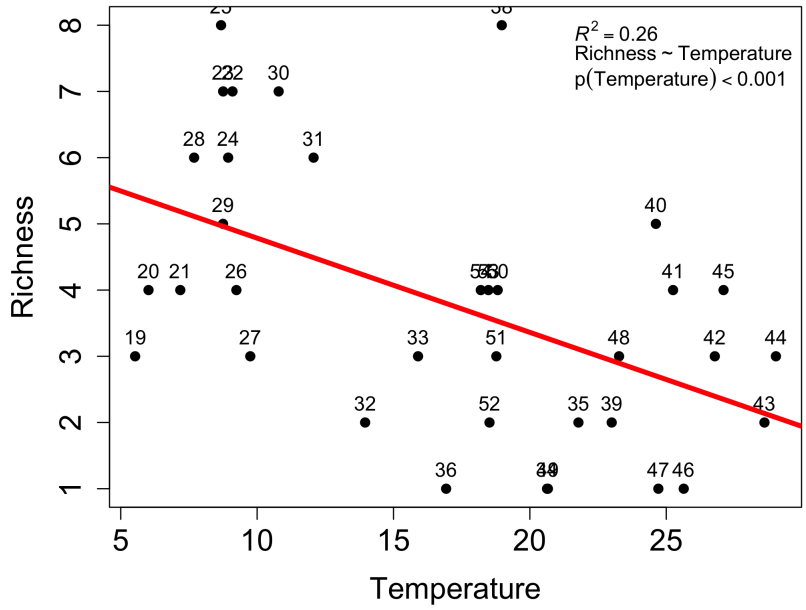

Figure 2.3.11: Alpha diversity was correlated to environmental data. Model selection resulted in temperature showing a negative correlation with alpha diversity (-0.04, R^2=0.26, p-value<0.001), meaning that station with high temperature had a less diverse diatom community.

Alpha diversity was calculated based orders per station and correlated with environmental variables (Figure 2.3.11). Temperature was the most significant variable explaining most of the variation (26%) in alpha diversity along the Atlantic transect. The correlation coefficient between alpha diversity and temperature was -0.04, meaning that stations with high temperature had a less diverse diatom community.

Figure 2.3.12 shows the relative abundance of diatom orders across the Atlantic transect ordered by latitude. Coscinodiscales (52%) were dominating the transect followed by Leptocylindrales (14%) and Corethrales (10%). The latter two orders showed a higher abundance in the temperate part of the section while Coscinodiscales were abundant throughout the whole transect. Diatom diversity in the tropical part of the transect was lower in comparison to the temperate part. The North Atlantic between Iceland and the coast of Spain was dominated by species belonging to Chaetocerotales, Fragilariales, Aulacoseiraceae and Achnanthales and were absent in the

Norwegian Sea and tropical Atlantic.

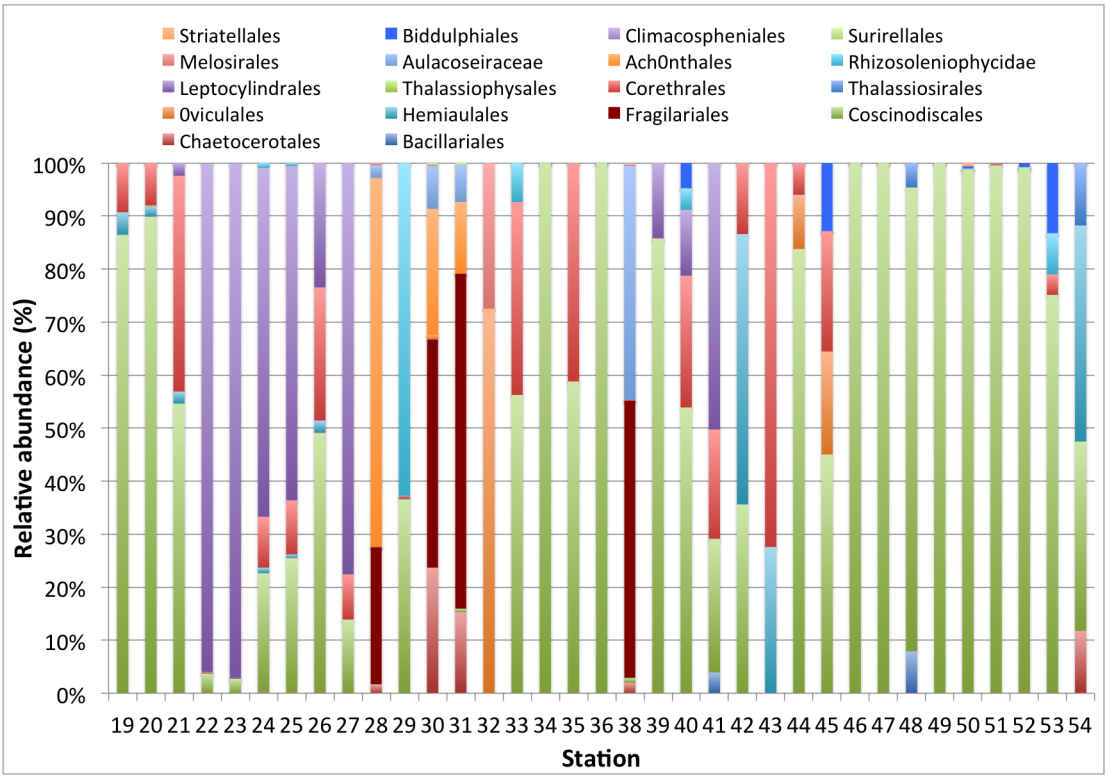

Figure 2.3.12: Relative abundance of diatoms orders across Atlantic stations ordered by latitude.

Cluster analysis showed that the Atlantic transect can be split in four clusters that roughly respond to similar regions. Cluster one groups stations 28,30,31 and 38 which were located in a nutrient rich eddy (Willem van de Poll, personal communication). Cluster two reflects the oligotropic regions of the Atlantic while cluster three grouped stations in nutrient rich temperate regions and up-welling of the coast of Africa.

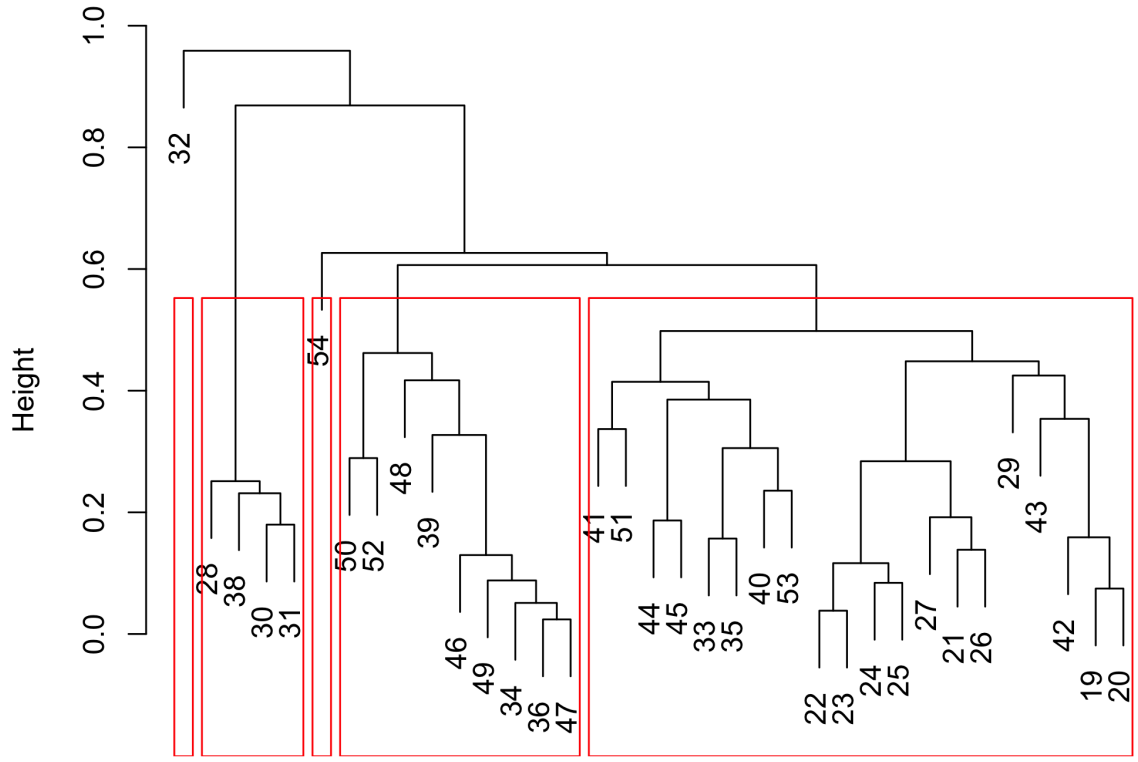

Figure 2.3.13: Beta diversity dendrogram based on bray-curtis dissimilarties and average.

Environmental data for the Atlantic section showed correlation between salinity and phosphate and silica (Figure 2.3.14). Weak correlation was found between temperature and salinity, as well as temperature and nutrients. Based on the VIF analysis phosphate was excluded from the environmental data.

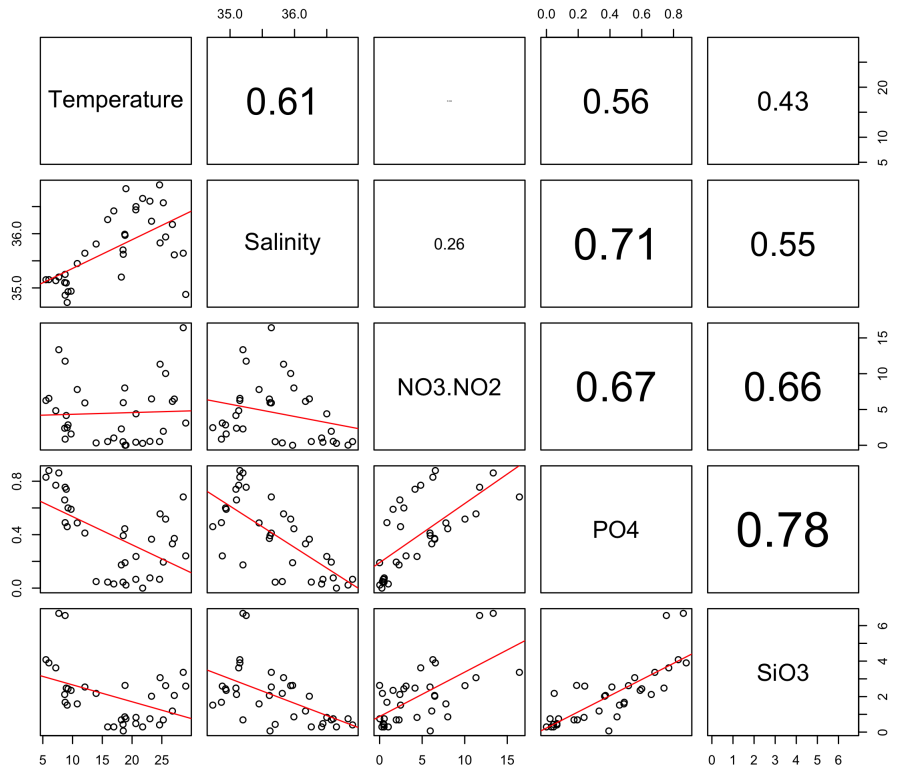

Figure 2.3.14: Test for co-correlation of environmental data for the Atlantic transect.

Diatoms were significantly related to variations in temperature, salinity and nitrogen concentrations (Figure 2.3.15) (p-value<0.001). This model explained a total variation of 15.78% with temperature explaining 5.7%, nitrogen 5.61% and salinity 4.47%.

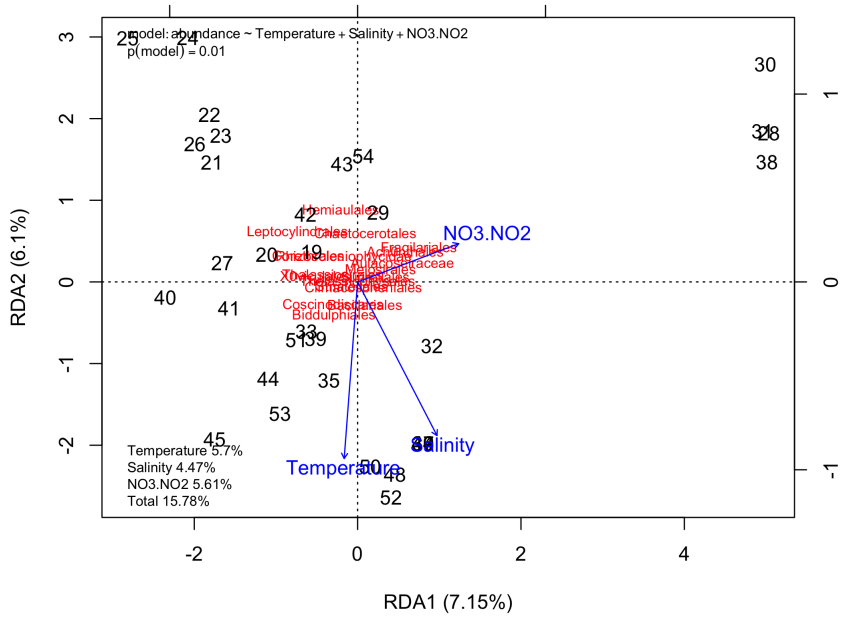

Figure 2.3.15: Correlation of environmental data and diatom order abundance for the latitudinal transect (Atlantic Ocean).

Figure 2.3.16 provides a summary of diatom order abundance (Hellinger transformed) correlating with temperature the correlation. Out of the 17 orders identified in the Atlantic Ocean, only two correlated significantly with temperature. Chaetocerotales and Hemiaulales both correlated negatively with temperature (-0.39, p(0.02) both).

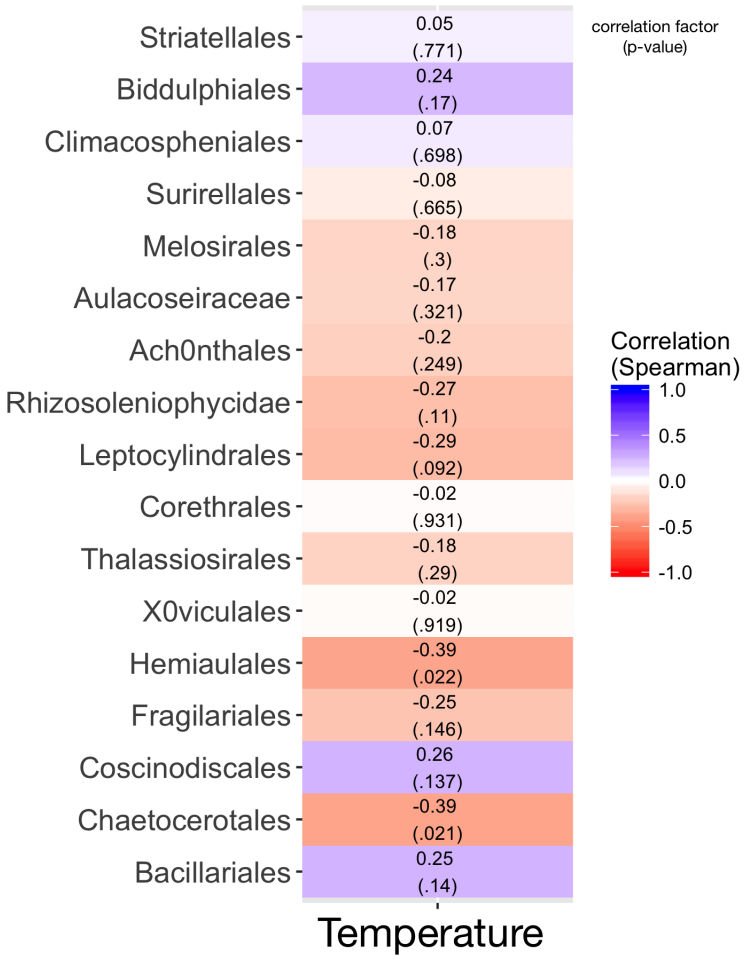

Figure 2.3.16: Correlation (Spearman rho) of diatom order abundance (Hellinger transformed) with temperature. Spearman correlation factor is color coded with blue showing strong positive correlation and red strong negative correlation. Values in the boxes provide the correlation coefficient and the p-value in brackets.

The bray-curtis index across the Atlantic transect was significantly related to temperature (mantel test, r_m=0.11, p-value=0.034, 999 permutations). This correlation was not significant after correcting for the auto-correlation of temperature and distance (partial mantel test, R^2=0.12, p-value=0.109, 999 permutations). In contrast, distance had no significant effect on bray-curtis index (mantel test, r_m=0.08, p-value=0.096, 999 permutations, which increased with increasing distance (Figure 2.3.17).

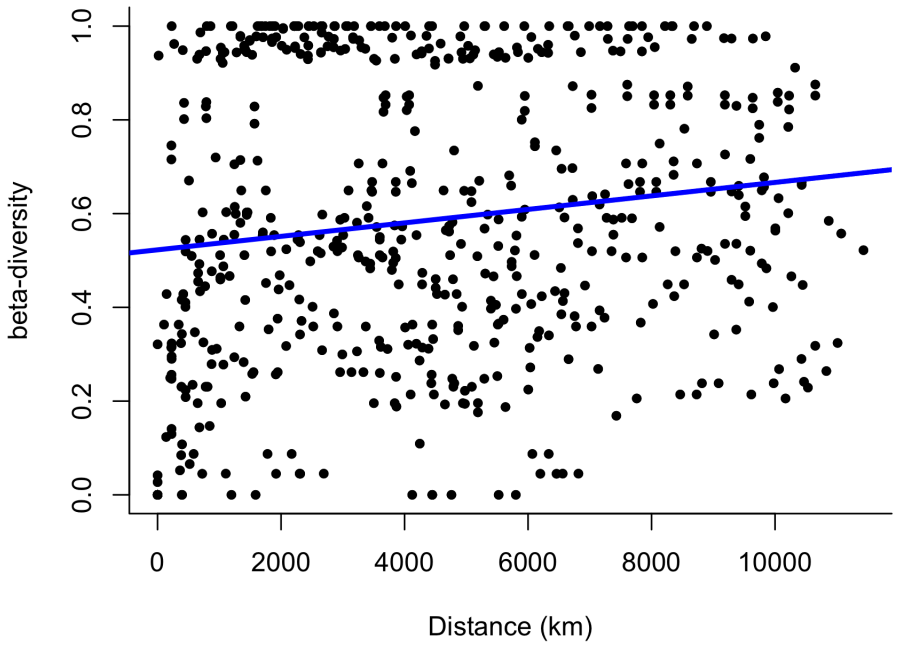

Figure 2.3.17: Correlation (Spearman) of the bray-curtis index with distance (km). Beta-diversity showed no significant correlation with distance (r_m=0.08, p-value=0.096).

2.4 Discussion

The Sea of Change dataset provided new insight into diatom community structures in the Atlantic Ocean, ranging from the Fram Strait to the South Atlantic. Our analysis is based on the hypervariable region V4 of 18S rRNA locus. Metabarcodes are a useful tool for high throughput environmental sequence surveys allowing species level identification. Indeed, the V4 region has been shown to be able to resolve taxa at the species level (Zimmermann *et al.*, 2011), however with limited resolution of cryptic species (Luddington *et al.*, 2012). Even though, some species are not well resolved by the V4 region, it was nonetheless well suited to explore order-level diatom diversity. The availability of reference sequences is one potential caveat of metabarcoding studies. The SILVA database `www.arb-silva.de` contains 11,993 Bacillariophyta 18S SSU-rRNA gene sequences providing a substantial reference volume. However, due to limitation in computational power, we used 161 unique references sequences covering 31 (out of 44 identified (Round *et al.*, 1990)) diatom orders. The low sequencing depth suggest that we were

most likely able to pick up the most abundant species and potentially missed rarer ones.

Another issue of metabarcoding project is the presence of multiple copies of 18S genes, especially in dinoflagellates (Godhe *et al.*, 2008). To overcome this issue, we related genome size to 18S copy number, which has been shown previously to correlate in eukaryotes (Prokopowich *et al.*, 2003), and use this relationship to normalize our diatom dataset. The generated equation resulted in underestimation (10-fold) of 18S copy numbers, however, previously proposed normalization to cell size (Godhe *et al.*, 2008; Vargas *et al.*, 2015), overestimated some species by over 100-fold (Kara Martin, personal communication). Therefore, our abundance estimate of identified diatom orders are conservative and need to be taken with caution. Furthermore, copy number variation is lower in diatoms (10s-100s) compared to dinoflagellates (200s- $<$10000s) (Zhu *et al.*, 2005; Godhe *et al.*, 2008) and resolving orders rather than species provides more confidence in our approach. Since cell size is related to 18S copy number (Zhu *et al.*, 2005), this is therefore less likely to effect our community analysis as we sampled size classes between 1.2 µm and 100 µm due to prefiltration over a 100 µm mesh and retaining cells on a 1.2 µm filter. However, it is not possible to completely exclude organisms that are smaller or larger than our desired size classes due to broken cells passing through the prefilter, or clogging of the filters retaining smaller sized cells. The identification of larger zooplankton and dinoflagellate species indicates a certain degree of ”contamination” of our samples (Kara Martin, personal communication).

Taking above mentioned pitfalls into account, we were able to show how diatom diversity and abundances varies over a polar and Atlantic Ocean transect. With this dataset we were able to add important information on diatom distribution in the Atlantic Ocean, especially tropical and Southern Atlantic. A search on the OBIS database (`http://iobis.org/mapper/?taxon_id=739452`) showed that only little data on Bacillariophyta occurrence and abundance is available for the tropical South Atlantic Ocean as well as the polar Atlantic (Figure 2.1.1). Thus, we are closing a gap in the existing datasets and provide a clearer picture on distribution of diatom orders in the polar and tropical Atlantic Ocean. Three of our most southern stations (47,49 and 54) overlapped roughly with locations of the TARA Ocean dataset (Malviya *et al.*, 2015) but provided similar results in terms of dominant

diatom orders (Coscinodiscales) even though they were sampled in different years.
Surprisingly, Coscinodiscales were the most widespread and abundant diatom order, and were unknown in the polar Atlantic and tropical South Atlantic, however this lack of records could be related to missing sampling efforts. This order comprises of 1,288 species, with some having a wide thermal niche (0°C to >20°C), salinity range (24-35 PSU) and nutrient range (Round *et al.*, 1990; Durselen and Rick, 1999) potentially allowing them a widespread distribution in the Atlantic Ocean. Furthermore, local adaptation of species within the order of Cosciniodiscales could also explain their cosmopolitan distribution (Malviya *et al.*, 2015). Vanormelingen (2008) suggested that on higher taxonomic levels (genera and upwards) diatoms are cosmopolitan, but community diversity can vary greatly across locations (Malviya *et al.*, 2015).

Based on RDA, temperature (35.4%) explained most of the variation the polar section, while nutrients (nitrogen/phosphate), and temperature seemed to have played an important role in the Atlantic section. However, only 15.78% of the variation could be explained in the Atlantic section. This might be due to patches of high nutrient concentrations that occurred over the transect (2.3.3), for example, due to upwelling at the Cape Verde islands and Benguela upwelling zone (Longhurst, 1998). Furthermore, low sequencing depth and an incomplete reference database could have resulted in lack of orders. Notably, typical polar diatom orders known for their high abundance (Melosirales (*Melosira arctica*), Fragilariales (*Fragilariopsis*), Bacillariales (*Nitzschia*) (Boetius *et al.*, 2013; Hegseth and Sundfjord, 2008; Poulin *et al.*, 2011, 2014; Booth and Horner, 1997) were underrepresented in our dataset. The lack of *Melosira* related sequences was especially surprising due to their high abundance in the sea ice associated water column (Booth and Horner, 1997) and observation by on-board microscopy (personal observation). Potential primer biases could have contributed to the underrepresentation of some orders, however, the used 18S-V4 primers have been shown to be universal and allowing sufficient discrimination of diatom orders (Zimmermann *et al.*, 2011; Malviya *et al.*, 2015). Thus it might be more likely that the 100 µm prefiltration excluded *Melosira* as they form long and sturdy chains (Poulin *et al.*, 2014).

Several diatom orders showed distinct geographical preferences. Chaetocerotales and

Thalassiosirales are well known abundant species in the polar waters of the North Atlantic (Booth and Horner, 1997; Poulin *et al.*, 2011; Hegseth and Sundfjord, 2008; Poulin *et al.*, 2014; Gosselin *et al.*, 1997) and were mostly found in cold, low saline waters that are characteristic of the EGC suggesting they might have originated from the high Arctic. In contrast, Corethrales were mainly found in temperate and coastal areas like the canary islands, in agreement with previous reports (Malviya *et al.*, 2015). Leptocylindrales showed patches of high abundance in the Norwegian Sea and the upwelling region off NW Africa which made it the overall most abundant order. These parts are nutrient rich (Longhurst, 1998) and support blooms of larger diatoms like Corethales and Rhizosoleniphycidae which co-occurred with Leptocylindrales (Yallop, 2001). These findings are in agreement with measurements of the coast of Iceland in June 2001 also showing high abundances of Leptocylindrus species which might suggest that this order comprises an important component of the phytoplankton community in the North East Atlantic (Yallop, 2001). Diatom diversity showed a negative relationship with increasing temperature for the Atlantic transect, meaning that the temperate parts had a more diverse diatom community than the tropical parts. This is in accordance with other reports where diatom diversity is reduced in the tropical, oligotrophic parts of the Atlantic Ocean and are outcompeted by picocyanobacteria such as Proclorococcus and Synechococcus (Tarran *et al.*, 2006; Carpenter *et al.*, 1999; Marañón *et al.*, 2000; Zubkov *et al.*, 1998; Heywood *et al.*, 2006; Mojica *et al.*, 2015). The smaller size of picocyanobacteria presents a competitive advantage over bigger groups, e.g. diatoms, due to lowered nutrient requirements, large surface area to volume ratio and small diffusion boundary layers (Chisholm, 1992; Finkel *et al.*, 2010).

Phytoplankton community in the Fram Strait varied according to hydrographic characteristics (temperature and salinity) of the distinct water masses from the Arctic (East Greenland current (EGC), low salinity, low temperature) and Atlantic (West Spitsbergen current (WSC), high salinity, higher temperature) (Figure 2.3.7. Overall, temperature explained about 35% of the variation in diatom abundance across the polar section, in agreement with other studies conducted in the Fram Strait (Kilias *et al.*, 2013; Hardge *et al.*, 2015; Fragoso *et al.*, 2016).
The amount of variation explained by the RDA models made us confident that we captured major trends within our data, however, the remaining variation could be an indication of additional

factors structuring phytoplankton communities in the Atlantic Ocean. For example, a study by Mojica et al. (2015) found that stratification was a important factor in driving phytoplankton communities in the North East Atlantic. However, stratification is influenced by temperature and salinity and also impacts nutrient availability and is difficult to disentangle.

Since we only measured at one location at one point in time, we have no information about temporal changes in phytoplankton diversity and distribution. Studies from the Pacific and North Atlantic have shown, however, that phytoplankton communities show seasonal changes but have a similar composition each year (Tarran *et al.*, 2006; Kilias *et al.*, 2013; Hardge *et al.*, 2015; Fragoso *et al.*, 2016). Nevertheless, climate change driven community changes have already occurred in the North Atlantic (Bluhm *et al.*, 2011). For example, *Neodenticula seminae*, a endemic diatom to the Pacific, has been reported throughout several years (first detection in 1999) in the North Atlantic and might have invaded through the Canadian Arctic Archipelago and/or Fram Strait which is thought to have been possible due to extensive low ice cover North of Canada (Reid *et al.*, 2007). Furthermore, the coccolithophore *Emiliania huxleyi* has been shown to expand further north in the Atlantic Ocean into areas formerly occupied by diatoms assumed to be possible due to changes in nutrient regimes (Rousseaux and Gregg, 2015) and positive temperature anomalies (Hegseth and Sundfjord, 2008; Stockwell *et al.*, 2001). Thus, competition between endemic Arctic species and temperate Atlantic species might increase with ongoing ocean warming with currently unknown consequences for the polar ecosystem. Therefore, an assessment of diatom diversity and distribution in general and of vulnerable ocean parts like the Arctic is needed to form a baseline to which changes can be compared.

Diatom community richness was negatively correlated to temperature on the polar transect, with a richer diatom community close to Greenland. Additionally, temperature seems to play a more important role in driving community differences in the polar ocean than distance. This supports our hypothesis that temperature is a important driver of diatom diversity and distribution. Physical barriers, such as water mass fronts, could have caused the found biogeographic distribution of diatom communities in the Fram Strait. The West Spitsbergen Current (WSC) transports warmer, denser, water masses from the North Atlantic into the Arctic ocean, thus

communities in this region could be classified as temperate communities. On the other side of the Fram Strait, the East Greenland Current (EGC) transports cold, fresher, Arctic Waters into the North West Atlantic. Thus communities in the western Fram Strait could be classified as polar communities (Gradinger and Baumann, 1991). Indeed, the geographic disconnection of plankton communities by temperature and salinity gradients has been suggested to underpin community speciation (Kilias *et al.*, 2014; Hardge *et al.*, 2015; Wei and Kennett, 1988; Mousing *et al.*, 2016; Malviya *et al.*, 2015). This disconnection of the diatom communities might be problematic for climate change resilience as high connectivity increases genetic diversity, potentially maximising adaptive potential (Momigliano *et al.*, 2015). Nevertheless, plankton communities have been shown to track changes in environmental conditions through their dispersal abilities (Cermeño *et al.*, 2010; Barton *et al.*, 2016) and biodiversity recovery on large geological timescales (Cermeño *et al.*, 2010). It remains unclear however, how contemporary communities will respond to the current fast-paced climate change and how their response might affect biogeochemical cycles and food webs in the ocean.

Spatial processes seemed to have played no role in driving diatom communities in the Atlantic Ocean transect as there was no significant correlation between diatom beta-diversity and distance. Thus, diatom communities in the Atlantic Ocean do not seem to be dispersal limited. This stands in contrast to a study by Chust et al. (2013) where distance played significant roles in shaping diatom communities in the western part of the Atlantic Ocean.

The lack of correlation of beta-diversity with distance in the Atlantic Ocean might be related to the patchy nature of diatom communities and environmental conditions. The tropical Atlantic is generally nutrient poor supporting other phytoplankton groups than diatoms. However, several up-welling regions provide enough nutrients to support diatom blooms (Malviya *et al.*, 2015). As we sampled some of these up-welling regions on our transect there might be shorter spatial gradients within our dataset. Nevertheless, temperature correlated significantly with beta-diversity pattern (r_m=0.08, p-value=0.03) as well as richness (R^2=0.26, p-value$<$0.001) suggesting that temperature shapes diatom communities in the Atlantic Ocean. However, temperature explained only 8% of community differences indicating that other factors or combinations of factors, such as grazing, mixed layer depth or competition, might play significant role in shaping community

composition.

The study of 18S ribotypes provides a much clearer picture of phytoplankton distribution and diversity than the utilization of flow cytometry, microscopy or probe based approaches. Reduction in sequencing cost and increases in computational power will make studies like this more feasible. However, 18S studies have already allowed unveiling of unknown species and a more detailed view of diatom distribution in the global Ocean (Vargas *et al.*, 2015; Malviya *et al.*, 2015; Stecher *et al.*, 2015).

2.5 Summary and conclusion

Diatom diversity (based on 18S ribotypes) and distribution was investigated over a longitudinal transect in the Fram Strait and a latitudinal transect in the Atlantic Ocean. Diatom diversity and biogeography in the Fram Strait was negatively correlated with salinity/temperature, and with temperature in the Atlantic Ocean. Coscinodiscales was the most abundant diatom order and dominated both transects, however, their wide distribution in the Atlantic Ocean was formerly unknown. With this metabarcoding study we added additional information about diatom distribution and diversity in the contemporary Atlantic Ocean. As, in this study, temperature has been suggested to drive diatom diversity, global warming might pose challenges to diatom communities, with potential negative effects on marine food webs and biogeochemical cycles.

Chapter 3

Phenotypic outcomes of adaptive evolution to temperature in *Thalassiosira pseudonana*

3.1 Introduction

Global change has been shown to have a strong impact on the ocean environment. CO_2 in the atmosphere is absorbed by the ocean (Caldeira and Wickett, 2003), which causes pH values of the surface ocean waters to decrease with negative effects on, mainly, calcifying organisms such as corals, molluscs and coccolithophores (Feely *et al.*, 2004). Increasing water temperatures result in increasing stratification, limiting the nutrient flux to the surface ocean (Cermeno *et al.*, 2008; Hofmann and Schellnhuber, 2009; Boyd, 2011). Furthermore, temperature has a direct impact on organism physiology and ecology. With increasing stratification, phytoplankton species are likely to experience higher light intensities by being trapped in surface layers of the ocean. Out of these environmental factors, temperature has been shown to have the strongest effect on marine autotrophs (Boyd *et al.*, 2013; Tatters *et al.*, 2013*a*; Schlüter *et al.*, 2014). And, as shown in chapter 2, temperature drives diversity and distribution of diatoms in the ocean. Furthermore, temperature interacts with other environmental parameters (light, carbon dioxide concentration) on physiological process (growth rate, metabolic rate, photo-physiology or

calcification) (Xu *et al.*, 2014; Feng *et al.*, 2009).

The temperature size rules suggests a negative correlation between cell volume and temperature with 2.5% volume increase for every °C decrease in temperature (Atkinson, 1994). This has been shown to hold true for most diatoms (Adams *et al.*, 2013), however, diverging responses to sublethal high temperatures have been reported (Atkinson *et al.*, 2003). Metabolic rate increases with increasing temperatures due to the temperature dependence of enzymes (Gillooly *et al.*, 2001; Peterson *et al.*, 2007). Lowered temperature affects metabolic rates by reducing enzyme reaction speeds and membrane viscosity. This can cause an imbalance between energy absorption and metabolic sink in autotrophs (Ensminger *et al.*, 2006). Reduced membrane viscosity can impact nutrient acquisition as nutrient transport through the membrane is slowed. Nutrient acquisition in relation to temperature has been shown to vary for temperate and polar microorganisms. Polar diatoms have been shown to have a potentially higher acquisition rate of silicic acid than temperate species, however, they are less efficient in utilizing it for growth (Stapleford and Smith, 1996). Reay et al. 1999 showed that nitrate acquisition is inversely related to temperature for marine bacteria and algae independent of their thermal niche. No clear temperature dependence of phosphate utilization was measurable in natural phytoplankton assemblages (Reay *et al.*, 2001). Reduced nutrient acquisition can be compensated by means of decreasing cell size, increasing the surface area for nutrient uptake in relation to cell volume (Hein *et al.*, 1995).

Long-term low temperature adaptation include modifications of membrane viscosity by alteration of unsaturated fatty acid composition (Hazel, 1995), adjustments in pigment composition and concentration to adjust energy absorption (Geider, 1987; Ensminger *et al.*, 2006) and increased expression of enzymes and ribosomes to compensate for lowered reaction speeds (Jørgensen, 1968) under lower temperatures. These cellular adjustments impact the elemental composition as lipids (mostly C), proteins (mostly N) and ribosomes (mostly P) contribute to different element pools (Geider and La Roche, 2002).

Multiple short-term experiments have investigated the immediate (incubation of several days) response of natural and artificial diatom assemblages, and single species to changes in temperature, while only few studies tested for longer-term (weeks to months) responses. Furthermore, most long-term studies focused on growth responses, excluding measurements of

changes in cell quotas of elements like nitrogen, phosphate or silicon. However, these measurements are important for estimating the effects of ocean warming on biogeochemical cycles and food quality (Rossoll *et al.*, 2012). Short-term studies indicate that temperature causes significant changes in the elemental composition of phytoplankton cells. For example, Thompson (1999) showed for Tp, that a decrease in temperature results in increased cellular nitrogen and carbon quotas. However, no changes in the carbon to nitrogen ratio were observed. In contrast, cellular carbon quotas decreased for *Chaetoceros calcitrans* when temperature decreased (Anning *et al.*, 2001). Furthermore, carbon quotas increased with increasing temperature (Anning *et al.*, 2001).

Bacillariophyta, which *Thalassiosira pseudonana* (Tp) belongs to, are key players in the contemporary ocean in terms of both abundance and ecological functionality. They produce approximately 50% of marine primary production, playing an integral role in the global carbon cycle as well as the cycling of other elements such as, silicon, nitrogen and phosphorus (Field, 1998; Smetacek, 1999, 2000; Ragueneau *et al.*, 2006). Throughout their evolutionary history of approximately 180 million years, diatoms have successfully adapted to changes in their marine environment (Sims *et al.*, 2006; **?**). However, the ocean environment is changing at an unprecedented rate with temperature having the strongest impact on diatom diversity and distribution (see Chapter 2). Little is currently known how phytoplankton, and diatoms, will respond evolutionary to ocean warming and how their responses might affect food webs and biogeochemical cycles.

In order to investigate the long-term response to temperature changes of the model diatom *Thalassiosira pseudonana*, an experimental evolution approach was implemented, selecting Tp, for approximately 300 generations, at its upper (32°C) and lower (9°C) temperature limit under nutrient replete conditions. Changes to Tp's cellular composition (carbon, nitrogen, phosphate, silicon and chlorophyll a) were measured at the end of the experiment. Furthermore, physiological parameters, such as photosynthetic efficiency and size, were monitored over the course of the experiment. All measured parameters were related to selection temperature providing insight into the phenotypic outcomes of the evolution of thermal tolerance.

The experimental setup allowed investigating the following hypothesis:

- High and low temperature selection lines adapt to novel thermal conditions.

- Stoichiometric changes underpin thermal adaptation.
- Cell size follows an inverse relationship to temperature with bigger cells under lower temperatures.
- Light-temperature interactions result in changes of cellular chlorophyll concentrations.
- Temperature optima shift towards selection temperatures while the temperature maximum will remain stable.

3.2 Material and Methods

In order to investigate the mechanisms of thermal adaptation in diatoms, we conducted a long-term temperature study exposing *Thalassiosira pseudonana* (Tp) (CCMP1335) to high and low temperatures.

3.2.1 General experimental setup

Populations *Thalassiosira pseudonana* (CCMP1335) (formerly know as *Cyclotella nana*) were cultivated at optimal (22°C), high (32°C) and low (9°C) temperatures. Populations with potential standing genetic variation were used, instead of a single clone to simulate natural populations and maximise the evolutionary response to temperature selection.
Cultures growing at optimal temperature were kept in a climatically controlled room with light from above with an intensity of 75 $\pm$ 5 µmol photons·m^{-2}·sec^{-1}. Light for all three experimental setups was provided by Sylvania Grolux fluorescent light tubes (120 cm, 36 W, Grolux T8/Minilynx, Havells Sylvania Europe Ltd, London, UK) with a higher output in the red and blue wavelengths. The high and low temperature cultures were kept in plant growth chambers (Panasonic Biomedical Sales Europe BV, Etten Leur, The Netherlands) with lights from three sides with an intensity of 75 $\pm$ 5 µmol photons·m^{-2}·sec^{-1}. Light intensity and spectra were measured with a spectroradiometer (Macam Photometrics LTD., Livingston, Scotland).

All cultures were cultivated in artificial seawater made up with Seachem Marine Salt™, F

strength macronutrients, F/2 strength micronutrients (Guillard 1975) and sterile filtered with Millipore Polycap HD 0.65/0.45 µm and Sterivex 0.2 µm filters. The pH of the media was adjusted with HCl to 7.9, salinity was adjusted to 33 PSU by adding more Seachem Marine Salt™or MilliQ water. Media was stored in polycarbonate 2.5 L bottles at respective temperatures in black plastic boxes, excluding 32°C . Media permanently stored in the dark at 32°C showed signs of bacterial growth. Therefore, media for the 32°C cultures was stored at 22°C and acclimated overnight to 32°C before the transfers.

3.2.2 Cell cultivation and transfers

All experimental cultures were cultivated semi-continuously in 25 cm^2 polystyrene flasks with vented caps (VWR, Lutterworth, UK) and transfers were conducted every third day until each experimental setup reached 300 generations. Development of the cultures was monitored at every transfer by measuring cell concentration, average diameter and photosynthetic quantum yield of photosystem II (Fv/Fm). At every transfer, $3.5x10^5$ cells were transferred into new flasks with fresh media to assure exponential growth. Cultures were always transferred before they reached stationary phase.

Cell concentrations before each transfer were measured with a Beckman Coulter counter by transfering 1 mL per replicate into 19 mL salt water (33 PSU) in a coulter vial. The vial was inverted several times before measurement. Cell concentration measurements with the coulter counter provided average diameter measurements at the same time which were used for diameter trends over the course of the experiment.

The photosynthetic quantum yield of photosystem II (Fv/Fm) was measured by transfering 5 mL of each dark adapted replicate into the quartz glass cuvette (WATER-K) provided with the WATER-PAM (Heinz Walz GmbH, Effeltrich, Germany). A modulated light source, too low to drive photosynthesis, was then used to measure the minimum fluorescence (Fo) followed by a short saturating light pulse, closing all reaction centres. This allows measurement of the maximum fluorescence (Fm). Fv/Fm is then calculated by the WATER-PAM software based on

the following formula:

$$Fv/Fm = \frac{(Fm - Fo}{Fm}) \tag{3}$$

An Fv/Fm in the range of 0.6-0.7 is considered healthy for *Thalassiosira pseudonana*, while a lower value is indicative of cell stress.

3.2.3 Sampling and extractions

Sampling for particulate carbon and nitrogen (POC/N), particulate phosphorus (POP), biogenic silica (bSi) and pigments was conducted for all cultures every 50th generation for all cultures. Due to instrument failure all samples of POC/N and POP, except for the last sampling point (300 generations), were destroyed. Replicate cultures were inoculated into 2 L Erlenmeyer flasks and grown to a density of $5{\cdot}10^5$ cells·mL^{-1}, providing enough material for extractions and analysis. Furthermore, samples of the selection lines were taken for cryopreservation as well as fixated with formaldehyde. For each of the five biological replicate, three technical replicates were taken.

3.2.3.1 Particulate organic carbon and nitrogen (POC/N)

For POC/N measurements, 50 mL of each replicate were filtered under low pressure (<150 mm Hg) onto pre-weighed and combusted (4.5 hours at 450°C) glass fibre filters (GF/F) (25 mm diameter, Whatman, ME, USA) and rinsed once with distilled water to remove salts that would interfere with the later analysis. Filters were placed in 6-well plates with combusted aluminium weighing pans, dried at 60°C and kept in desiccators until analysis. For analysis, dry filters were folded into nylon caps and finally analysed via the combustion technique in an Exeter 440 elemental analyser. In order to calculate the carbon content of the samples, a calibration was done by weighing different amounts of standards (Acetanilide, Sigma-Aldrich, MO, USA) into tin capsules. The carbon and nitrogen content of sample filters was then calculated internally by the instruments software. The carbon and nitrogen content was then calculated per cell based on cells on filter. Additionally, carbon and nitrogen content were expressed as content per volume. Average volume calculations per cell were performed with average diameter and length

measurements from light microscopy (see section 3.2.3.6 for more detail).

3.2.3.2 Particulate organic phosphate (POP)

For POP measurement, 50 mL of each replicate were filtered under low pressure (<150 mm Hg) onto pre-weighed and combusted glass fibre filters (GF/F) (Whatman, 25 mm diameter) and rinsed once with distilled water to remove salts that would interfere with the later analysis. Filters were placed in 6-well plates with combusted aluminium weighing pans, dried at 60°C and kept in desiccators. Phosphate extraction was done using the Potassium Persulfate (K_2 S_2 O_8) digestion method of (Suzumura, 2008). Dried filters were placed in 50 mL acid washed Scott Duran bottles and 18 mL of 3% persulfate solution was added. Bottles were sealed as tight as possible and autoclaved for 30 min at 120°C. After samples had cooled down, 18 mL MilliQ water was added to reduce the persulfate concentration and to increase the pH >1.5. A final potassium persulfate concentration of more than 3% has been shown to interfere with the optical analysis of the samples (Suzumura, 2008). Phosphate contend per filter was measured with an Perkin Elmer Lambda 25 UV/Vis spectrophotometer and autosampler. These results were then expressed as particulate phosphate content per cell and volume. Average volume calculations per cell were performed with average diameter and length measurements from light microscopy (see section 3.2.3.6 for more detail).

3.2.3.3 Biogenic silica (bSi)

In order to analyse the bSi content of the cultures, 50 mL were filtered onto polycarbonate filters (25 mm diameter, 1.2 µm pore size, Millipore, MA, USA) under low pressure (<150 mm Hg) and dried at 60°C. Filters underwent a silica digestion before analysis by placing the filters in polymethylpentene (PMP) tubes, adding 4 mL of 0.2 M NaOH. The tubes were then placed in a water bath heated to 95°C for 1 h, neutralized with 1 M HCl and cooled before centrifugation for 15 min at 3000 rpm. The supernatant was transferred into new falcon tubes and stored in the fridge for subsequent dissolved silicic acid (DSi) analysis. DSi concentrations were determined with an Perkin Elmer Lambda 25 UV/Vis spectrophotometer and autosampler. DSi content was expressed as content per cell and surface. Average surface per cell calculations were performed with average

diameter and length measurements from light microscopy (see section 3.2.3.6 for more detail).

3.2.3.4 Chlorophyll a

For pigment analysis, 50 mL of each biological replicate were filtered onto GF/F filters under low pressure (<150 mm Hg) and snap frozen in liquid nitrogen. Samples were stored at -80°C until further analysis. Pigments were extracted by placing the frozen filters in brown glass scintillation vials, adding 10 mL of 90% Aceton and incubation in the dark and cold overnight. In order to measure chlorophyll a (chl a) concentrations, acidified and unacidified sample absorption was measured at 665 and 750 nm with a spectrophotometer (Perkin Elmer LS45 Spectrometer, Waltham, MA, USA). To calculate chl a concentrations, a standard curve of known chl a concentration was measured (Appendix Figure 6.2.2). These measurements were expressed as chl a content per cell and volume. Average volume calculations per cell were performed with average diameter and length measurements from light microscopy (see section 3.2.3.6 for more detail).

3.2.3.5 Statistical analysis of stoichiometric results

Nutrient (carbon, nitrogen, phosphate and silicate) content per cell and volume as well as ratios were analysed for significant differences with the statistical program SPSS Statistics (IBM SPSS Statistics, Version 22). Each dataset was tested for homogenity of variance and if not significant at the 0.05 level, a multiple group comparison (one-way ANOVA) with a Tukey's post-hoc test was used to test for significant differences between temperature treatments. If the test for homogenity of variances was significant, a multiple group comparison with a Games-Howell test was used. Were available, technical replicates for biological replicates were averaged. Replicate 5 of the 9°C treatment was seperated from the other replicates because of the different evolved phenotype. Results were visualized as boxplots generated in Sigmaplot (SysStat Software, Inc, Version 10). Small letters were used to indicate significant differences (p-value <0.05) between the boxes with different letters.

3.2.3.6 Light microscopy and scanning electron microscopy

In order to relate stoichiometric measurements to size or shape changes of the cells, 80 - 100 mL of culture were fixed in duplicate with 2% Formaldehyde (final concentration) in brown glass bottles and stored in the dark and at room temperature (about 18°C). Formaldehyde was chosen as the fixing agent as it preserves frustule structure and allows sample preparations for scanning electron microscopy (Bank Beszerti, personal communication).

Scanning electron microscopy (SEM) Culture samples (15 mL) for scanning electron microscopy were transferred to a 15 mL centrifugation tubes and centrifuged at 500 rpm for 5 min. The supernatant was carefully decanted and tubes filled up with MilliQ water and centrifuged again. This washing step is repeated five times. After the washing, samples are treated with a 1:5 bleach solution for 7 min and washed again. Samples are mounted onto round glass covers and dried over night. Stubs are sputtered with gold particles and imaged with a SEM at the John Innes Centre (Norwich Research Park).

Light microscopy Samples for microscopic measurements of cell size are centrifuged to concentrate the cells and 1 mL was transferred to a Sedgewick-Rafter cell S50 (Pyser-SGI, West Chester, PA, USA). Sedgewick-Rafter cells were analysed with a Zeiss AxioPlan 2ie widefield microscope (Zeiss, Jena, Germany) equipped with a AxioCam HRm CCD camera (Zeiss, Jena, Germany). Pictures were analysed with ImageJ and Tp's cell length and diameter was measured for at least 100 cells per sample. These measurements were used to calculate the average cell surface (equation 4) and volume (equation 5) with the formula:

$$S = 2 \times \pi \times r \times h + 2 \times \pi \times r^2 \tag{4}$$

$$V = \pi \times r^2 \times h \tag{5}$$

3.2.4 Reciprocal transplant assays

After 300 generations of temperature selection, reciprocal transplant assays were conducted for each selection temperature by placing triplicates of each biological replicate at all three selection temperatures (Figure 3.2.1). All experimental cultures were acclimated for 10 generations at the assayed temperatures prior to growth measurements. A whole growth curve (exponential to stationary) was measured for each assay. Growth rates were calculated for the exponential phase of the growth curve by selection of at least three data points and fitting of a trend line.

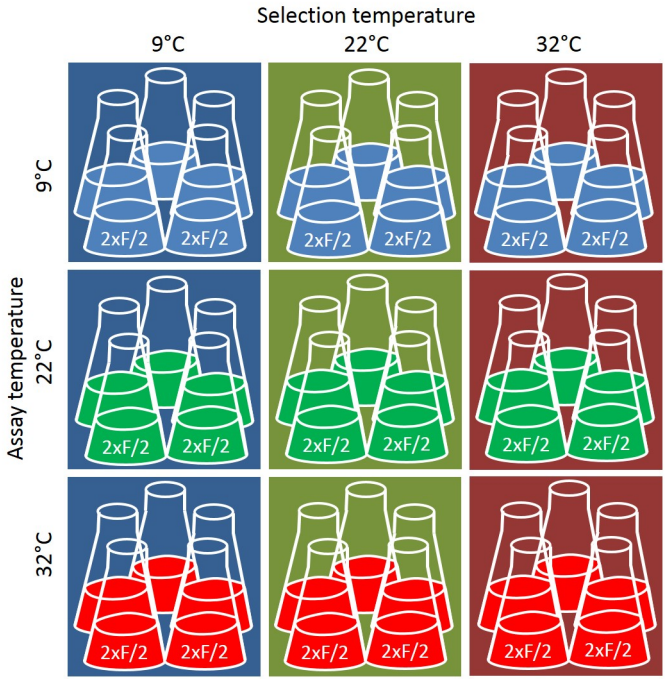

Figure 3.2.1: Experimental setup of reciprocal transplant assays. 9°C selection cultures are assayed at 9°C, 22°C and 32°C. The same is done for 22°C and 32°C selection cultures.

3.2.4.1 Statistical analysis of transplant assay

For the transplant assay, three technical replicates were measured for each of the five biological replicates. The technical replicates were averaged and all growth rates log10 transformed to produce a normal distributed dataset (Shapiro-Wilk normality test, p-value=0.06). Levene's test was used to test for homogeneity of variance (p-value=0.06). A two-way ANOVA (stats package in R) was conducted to examine the effect of selection temperature and assay temperature on growth rate. There was a statistically significant interaction between the effect of selection and

assay temperature on growth rate, (F(4,40)=470.8 (p-value <0.001)). A post-hoc Tukey HSD test was implemented to test for significant different growth rates between the selection lines at each assay temperature.

3.2.5 Temperature response curves

In order to identify shifts of the temperature optimum, temperature response curves were measured in a custom built Temperature Gradient Bar (by Eric Buitenhuis UEA) for one replicate of each selection temperature. Time restrictions did not allow for measuring all five replicates in the gradient bar, therefore, only one biological replicate was measured for each temperature treatment (22C Tp1, 32C Tp3, 9C Tp1 and Tp5). Cultures were acclimated for three transfers, approximately 10 generations, in the gradient bar with temperatures ranging from +4°C to +37°C. A preliminary response curve conducted under experimental light intensities (70 $\pm$ 5 µmol photons·m^{-2}·sec^{-1}) showed a wide plateau in growth rates, not allowing identification of a clear temperature maximum **(Appendix)**. Therefore, light intensities were set to 165 $\pm$ 15 µmol photons·m^{-2}·sec^{-1} to assure saturated photosynthesis and clear maximum growth rates.

3.2.5.1 Temperature response curves - Modelling

We selected three models to estimate temperature maximum (Tmax) and temperature optimum (Topt). Model (3) is a optimum equation based on Schoemann et al. (2005), the fourth model is an exponential equation after Norberg (2004) and the fifth is a beta distribution equation from Yan and Hunt (1999). The three models were evaluated with the temperature response curve for the control lines and model selection was based on Akaike's Information Criterion (AIC) (Anderson *et al.*, 1998) values. The model with the lowest AIC value, compared to other models, describes the data set best, however, it is no measure of how well a model fits overall. A significant difference between two models is found for differences in AIC values of more than 2.

$$\mu = \mu_{max} \times e^{\left[-\frac{(T-T_{opt})^2}{dT^2}\right]} \quad (6)$$

$$\mu = \left[1 - \left(\frac{T - Z}{w}\right)^2\right] \times 0.59e^{0.0633T} \tag{7}$$

$$\frac{\mu}{\mu_{max}} = \left(\frac{T_{max} - T}{T_{max} - T_{opt}}\right) \times \left(\frac{T}{T_{opt}}\right)^{\frac{T_{opt}}{T_{max} - T_{opt}}} \tag{8}$$

In model 6, growth rate μ is a function of the maximum growth rate μmax, temperature T, optimum temperature Topt and the temperature interval of growth dT (Schoemann *et al.*, 2005). Model 7 relates growth μ to temperature T, temperature niche width ω and the temperature optimum Z (Norberg, 2004). The fifth model (equation 8) relates normalized growth to temperature T, the optimum temperature Topt and the maximum temperature Tmax where growth occurs (Yan and Hunt, 1999).

3.2.5.2 Statistical analysis of temperature response curves

Growth data for all three lines (9 °C, 22 °C and 32 °C) was normalized to the maximum achieved growth rate. Nonlinear (weighted) least-squares (nls) estimates, based on the Yan et al. **?**Yan1999) model, (nls, part of the nlstools package) were used to estimate model parameters (Topt and Tmax) using the statistical computing software RStudio (Version 0.99.441) (**?**). Starting values for the nls function were estimated based on visual inspection of the normalized temperature response curve. Since the temperature response curves were only measured for one biological replicate per temperature treatment, Topt estimates (N=1 per temperature treatment) were tested for correlation (pearson) with selection temperature.

Temperature response curves of the smaller and larger cold adapted Tp cultures (9C Tp1 and Tp5, N=3 each) were tested for significant differences by comparing the non-linear curve fits with an ANOVA. For this, overall fits between the two data sets were compared by constructing a pooled data set containing growth rates for both phenotypes coded as factors. This then allows to fit different sets of values for the model parameters depending on the phenotype and test for significant difference between the pooled and phenotype fits with an ANOVA. If significant, parameter estimates for Topt and Tmax were tested for significant differences. T-scores based on a Welch's t-test were calculated for each parameter and used together with the degrees of

freedom to calculate the p-value with the ”P value from t score calculator” (http://www.socscistatistics.com/pvalues/tdistribution.aspx, last accessed 22nd of March 2016).

3.3 Results

3.3.1 Physiological responses to temperature changes

3.3.1.1 Growth and photosynthesis

Five replicates for each temperature treatment were grown at 22°C, 9°C and 32°C and are referred to as control, low and high temperature treatments, respectively. Initial growth rates, right after the transfer to the selection temperatures, are shown in Figure 3.3.1.

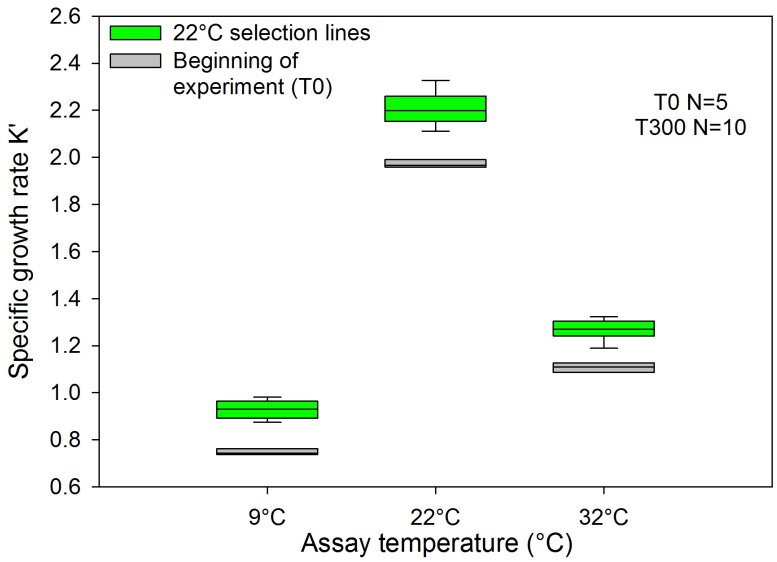

Figure 3.3.1: Boxplot of growth rates at the beginning and end of the experiment for control cultures at selection temperatures. Growth rates are based on five replicates at each assay temperature. Lines in the boxplot show the median growth rate. Gray boxes represent initial growth rates immediately after transfer to selection temperatures (T0). Green boxes show growth rates of the control cultures (22°C) after 300 generations measured at all three selection temperatures.

The initial temperature treatment at 9°C and 32°C resulted in reduced growth rates in comparison to the control temperature. The growth rate reduction was most pronounced at 9°C with an initial

growth rate of 0.78^{-d}. High temperature cultures showed an initial growth rate of 1.1^{-d} while the control cultures at 22°C showed the highest growth rate with 1.9^{-d}.

Before each transfer, cell concentrations and photosynthetic efficiency were measured. Number of generations completed between transfers were calculated based on cell yield after three days. Generations completed between each transfer increased for all three temperature treatments for about 50 generations and remained stable until the end of the experiment (Figure 3.3.2). Control replicates produced seven generations, high and low temperature replicates six and four generations on average, respectively.

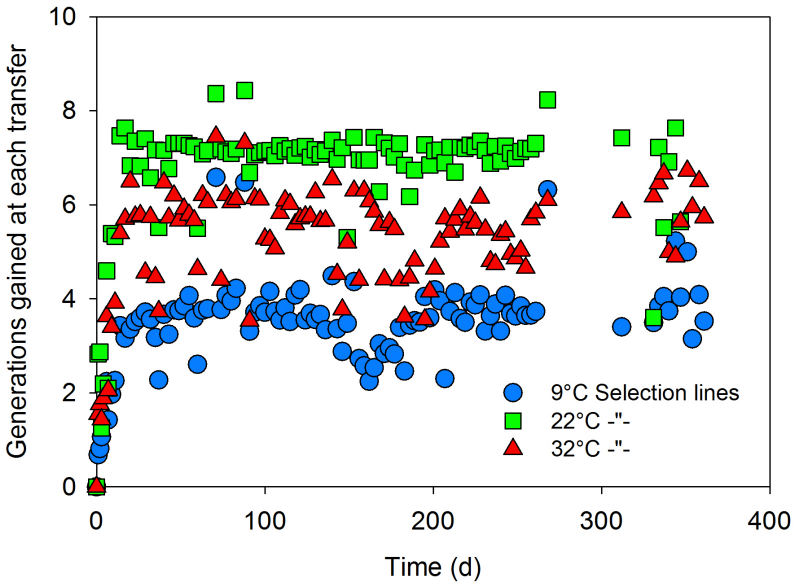

Figure 3.3.2: Generations produced of selection lines between each transfer over the course of the experiment. Control cultures (22°C) in green, high temperature cultures (32°C) in red and low temperature treatment (9°C) in blue.

Photosynthetic quantum yield (Fv/Fm) was monitored over the course of the experiment and measured before every culture transfer (Figure 3.3.3). High and low temperature treatment resulted in reduced Fv/Fm in the high and low temperature replicates while the control replicates were unaffected. Initial Fv/Fm dropped from 0.66 to 0.57 for the high temperature treatment and to 0.42 for the low temperature treatment. For replicates grown at control temperature (22°C), Fv/Fm remained stable at 0.66 over the course of the experiment. Interestingly, the low temperature cultures showed a continuous significant (p-value <0.001, $R^2=0.52$) increase in their

photosynthetic efficiency until the end of the experiment (Figure 3.3.3, blue).

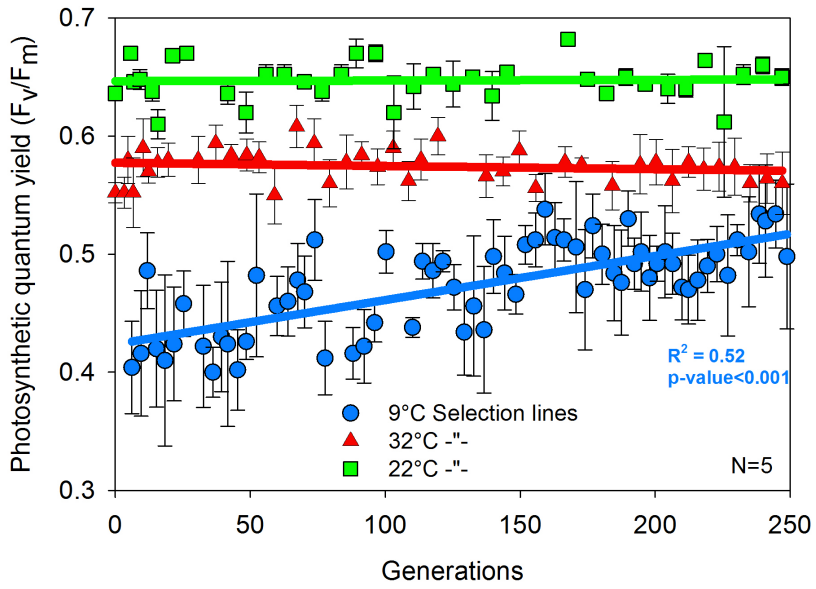

Figure 3.3.3: Photosynthetic quantum yield of PSII (Fv/Fm) over the course of the experiment. Control cultures (22°C) in green, high temperature cultures (32°C) in red and low temperature treatment (9°C) in blue. Control and high temperature cultures showed no change in their Fv/Fm, while low temperature Fv/Fm increased significantly (blue line, $p<0.001$) until the end of the experiment (R^2 0.52).

The average cell diameter was measured at every transfer and is shown in Figure 3.3.4. Control (green) and high temperature (red) selection cultures showed increases in their average cell diameter for the first 100 generations and remained stable at 4.8 µm and 5.25 µm, respectively, until the end of the experiment. The average diameter for the low temperature (light and dark blue) selection lines increased over 200 generations and stabilized at 5.9 µm for four (dark blue) out of five replicates.

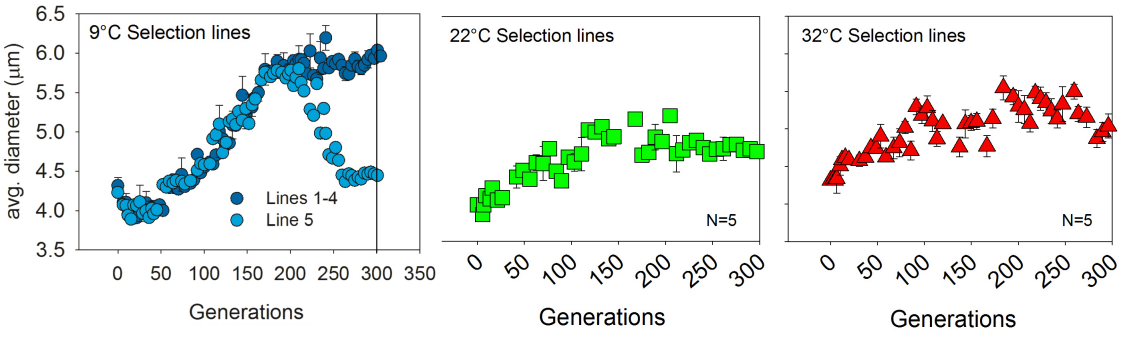

Figure 3.3.4: Average diameter trends over the course of the experiment for all three temperature treatments based on coulter counter measurements. Low temperature cultures (9°C, left), control cultures (22°C, middle) and high temperature cultures (32°C, right). Each measurement shows the average of five replicates (four replicates for the low temperature cultures 1-4 (darkblue), one for replicate 5 (light blue)) and the corresponding error bars.

One replicate of the cold temperature cultures showed a decline in the average cell diameter, to 4.5 µm, after the initial increase over 200 generations (Figure 3.3.4, light blue) compared to the other four replicates (Figure 3.3.4, dark blue).

In order to test for adaptation to the selection temperatures, reciprocal transplant assays were conducted after 300 generations for all replicates of each selection experiment. All selection cultures showed the highest growth rates at their selection temperature and reduced growth under ancestral conditions (Figure 3.3.5). There was a statistically significant interaction between the effect of selection and assay temperature on growth rate, (Two-way ANOVA, $F(4,40)=470.8$ (p-value <0.001)).

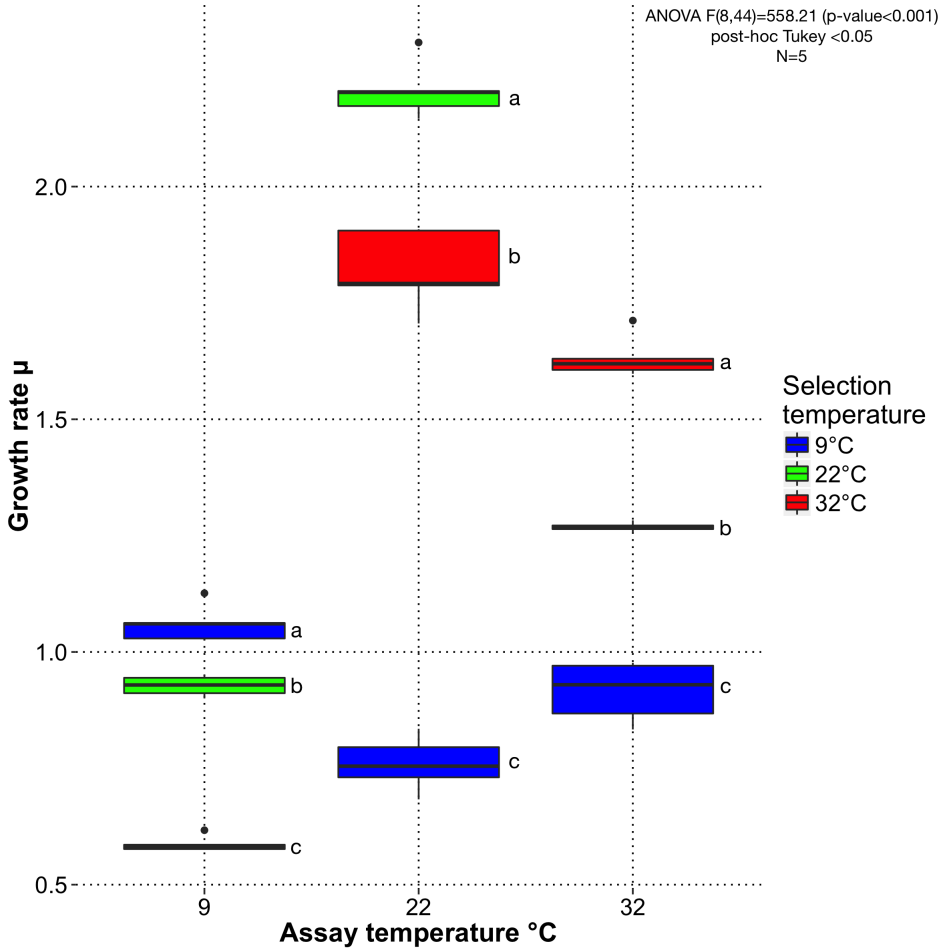

Figure 3.3.5: Boxplot of growth rates (μ/μ_{max}) of each temperature selection line (9°C (blue), 22°C (green), 32°C (red)), after 300 generations, at assay temperatures. A two-way ANOVA showed that there was a statistically significant interaction between the effect of selection and assay temperature on growth rate, (F(4,40)=470.8 (p-value <0.001)) (Table 3.3.1). A post-hoc Tukey HSD test was used to test for significant differences (p-value <0.05) between growth rates of each selection line at the assay temperatures. Small letters indicate significant differences between boxes. Error bars indicate one standard deviation.

Table 3.3.1: Results of the two-way ANOVA with Selection Temp and Assay Temp interaction term implemented to analyse the reciprocal transplant assay. All terms were significant at the 0.05 level.

Source	DF	F value	p-value
Selection Temp	2	357.0	<0.001
Assay Temp	2	628.7	<0.001
Selection Temp : Assay Temp	4	470.8	<0.001
Residuals	36		

3.3.1.2 Temperature response curves (TRC)

To evaluate whether temperature niche and width changed under the three selection temperature, we evaluated three temperature response models (see Methods section 3.2.5.1) with our control dataset (22°C) (Figure 3.3.6). The Schoemann equation (model 6, Figure 3.3.6, dark blue) overestimates growth responses at the higher temperature range while the Norberg equation (model 6, Figure 3.3.6, orange) overestimates growth responses at the lower temperature ranges. The Yan equation (model 8, Figure 3.3.6, black) fits the dataset smoothly which is also represented in the AIC values (below).

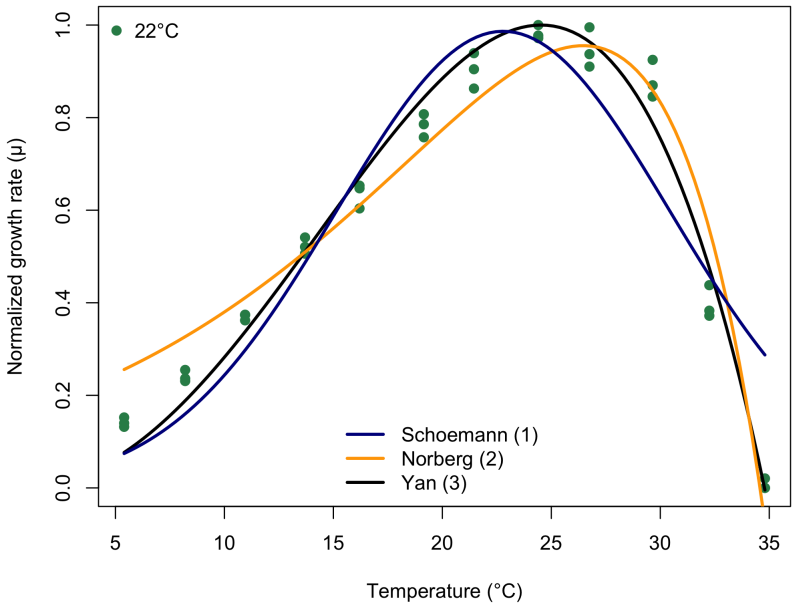

Figure 3.3.6: Visualization of tested models fitted to the temperature response curve of control cultures (22°C, dark green circles). Schoeman model (6) in dark blue, Norberg model (7) in orange and Yan (8) in black. The Yan equation showed the smoothest fit to the data.

The relative fit of the three tested equations was calculated with AIC values 3.3.2. AIC values are a measure for how well a model fits to a given dataset. The model with the smallest AIC value represents the best fit, however, it is no measure for the goodness of the fit. Based on the AIC value calculations and visual inspection, model 8 (Yan and Hunt, 1999) (arrow) represents the best fit to our dataset and was chosen for modelling the high and low temperature TRCs.

Table 3.3.2: AIC values of fitted models to the temperature response curve of control lines (22°C).

Model	AIC
Yan (8)	-101.31204 ←
Schoemann (6)	-41.39021
Norberg (7)	-74.28916

Model 8 (Yan and Hunt, 1999) was fitted to all three temperature response curves of the selection lines and parameter estimates for Topt and Tmax calculated. Low temperature selection line (9°C Tp1, blue) showed a decrease in their optimal temperature from 24.94°C ±0.23°C to 22.82°C ±0.22°C. Furthermore, the 9°C selection lines showed a reduction in their temperature maximum from 35.57°C ±0.17°C to 33.94°C ±0.24°C. In contrast, the high temperature selection line (32°C, red) showed the opposite trend, namely an increase in their temperature optimum to 25.8°C ±0.25°C. The temperature maximum showed no change. Changes in optimum temperature were not significantly correlated to the selection temperature (p-value=0.1) (Figure 3.3.7).

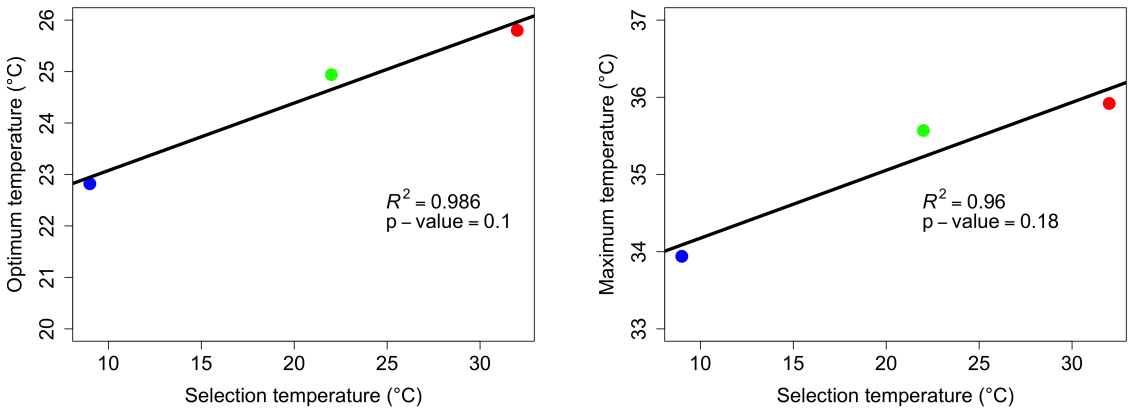

Figure 3.3.7: Correlation between temperature optimum (A) and maximum (B) estimates based on Yan model (Yan and Hunt, 1999) and selection temperature. Low temperature optimum (9°C) estimate is shown in blue, optimum temperature optimum (22°C) in green and high temperature optimum (32°C) in red. The correlation was not significant (p-value 0.1).

Low temperature replicate 5 (light blue) showed a significant different temperature response curve in comparison to replicates 1-4 (blue) and control cultures (green) (Figure 3.3.8, right). Replicate 5 showed a optimum temperature of 23.81°C ±0.29°C ($p<0.001$) and a temperature maximum of

36.09°C ±0.46°C (p<0.001).

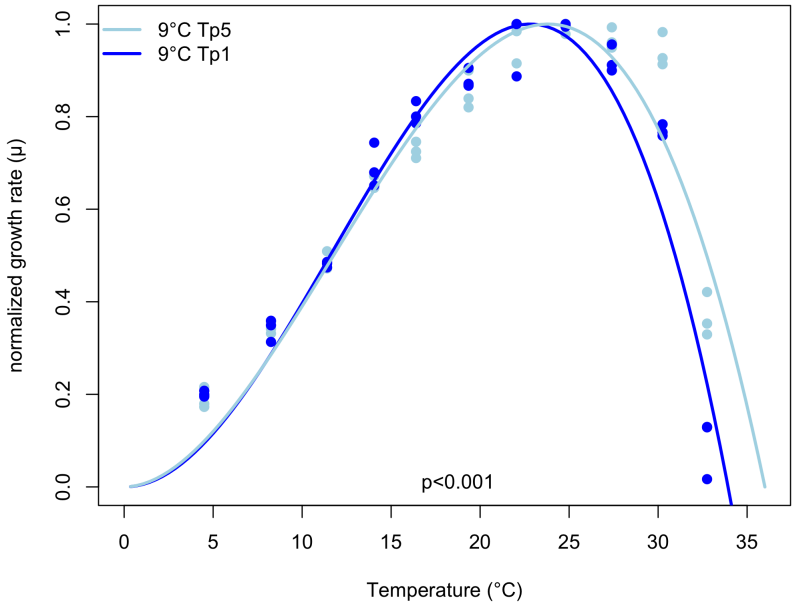

Figure 3.3.8: TRC models for low temperature selection lines 1-4 (larger phenotype, darkblue) and line 5 (smaller phenotype, lightblue) after 300 generations of temperature selection.

3.3.2 Cellular composition of *Thalasssiosira pseudonana* under high and low temperature selection

Temperature selection had an impact on the cellular stoichiometry of Tp (Figure 3.3.9). Carbon content per cell (Figure 3.3.9) increased significantly (p<0.05) with temperature from 8.1 ±2.2 pg at 9°C (blue) to 18.2 ±4.1 pg at 32°C (red), while control cultures at 22°C (green) showed 11.7±1.2 pg carbon per cell. Biogenic silica content (pg) per cell (Figure 3.3.9) showed significant (p 0.05) increases for both low (blue) and high (red) temperature lines in comparison to the control cultures (green). The biogenic silica was lowest for the control with 4.4 ±0.42 pg and increased to 5.9 ±0.39 pg and 6.6 ±1.3 pg for low and high temperature cultures. Cellular nitrogen content per cell (Figure 3.3.9) was significantly (p 0.05) reduced for low temperature cultures (blue) at 1.25 ±0.39 pg while control (2.02 ±0.13 pg) and high temperature (2.14 ±0.39 pg) lines showed no significant difference in their cellular nitrogen content. Furthermore, temperature had an impact on the cellular chlorophyll a (chl a) content (Figure 3.3.9) which

increased with increasing temperature. Low temperature cultures (blue) had the lowest chl a content per cell with 0.22 ±0.02 pg. High temperature cultures (red) showed the highest chl a content with 0.74 ±0.15 pg while control cultures (green) had 0.44 ±0.03 pg chl a per cell. Concentrations of intracellular phosphate did not vary significantly between the selection lines at low and high temperature ($p > 0.05$). Low temperature lines (Tp 1-4) showed the lowest phosphate content with 0.46 ±0.05 µg per cell. Low temperature replicate 5 had a significant lower phosphate content than the other low temperature replicates ($p < 0.05$).

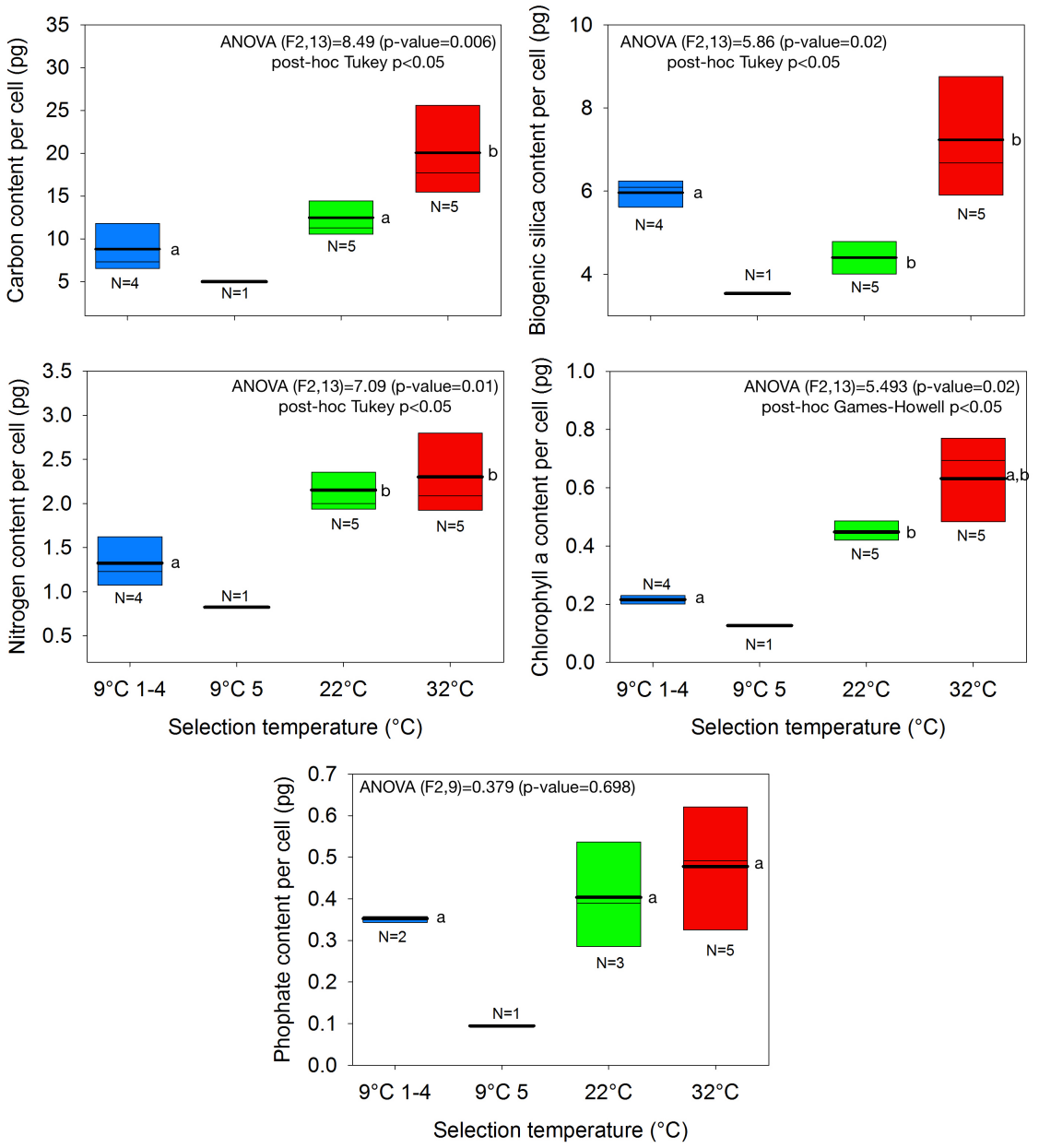

Figure 3.3.9: Cellular carbon, nitrogen, biogenic silica and chlorophyll a content per cell for all three selection lines. Control cultures in green, high temperature and low temperature treatments in red and blue respectively. Thick black lines in the box plot represent mean while thin black lines respond to the median. Small letters next to the box indicate significant differences to other boxes. Boxes sharing the same letter(s) are not significantly different, while different letters indicate significant differences. Values were significantly different at a p-value cut-off of 0.05.

Because the average cell volume changed under the three different selection temperatures (see Figure 3.3.4), cellular content of carbon, biogenic silica, nitrogen and chlorophyll a was related to cell volume, and cell surface in case of biogenic silica (Figure 3.3.10). Carbon content per cell volume (Figure 3.3.10) increased significantly from 9°C to 22°C from 0.07 μm^{-3} ±0.02 pg to 0.17

μm^{-3} ±0.02 pg ($p<0.05$). 32°C cultures (0.21 μm^{-3} ±0.05 pg) showed an insignificantly increased volume related carbon content in comparison to the 22°cultures. When bSi content was related to cell surface area (Figure 3.3.10), low temperature culture (blue) concentrations decreased (0.04 μm^{-2} ±0.003 pg) while high temperature culture (red) concentrations increased significantly (0.06 μm^{-2} ±0.01 pg) in comparison to the control cultures (green) with 0.05 μm^{-2} ±0.005 pg ($p<0.05$). Low and high temperature cultures had the same bSi content per cell (Figure 3.3.9) however, in case of the low temperature cells, distributed over a larger surface area resulting in the decreased bSi content per cell surface area. Low temperature replicate 5 showed the lowest nitrogen content per volume with 0.006 μm^{-3} ±0.002 pg (Figure 3.3.10) ($p<0.05$). Low temperature replicates 1-4 had a higher nitrogen content per volume 0.01 μm^{-3} ±0.002 pg but significantly lower than control and high temperature cultures ($p<0.05$). The highest nitrogen per volume content was reported for control cultures with 0.03 μm^{-3} ±0.002 pg while nitrogen per volume content was slightly lower for the high temperature cultures (red) at 0.03 μm^{-3} ±0.005 pg ($p>0.05$). Chlorophyll a content per cell volume (Figure 3.3.10) increased significantly from 9°C to 22°C ($p<0.05$). The low temperature selection lines showed a significant decrease in their chlorophyll a content per volume in comparison to the control lines (Figure 3.3.10, blue and green respectively) while the high temperature selection lines (red) showed a slight increase in their chlorophyll a ($p>0.05$).

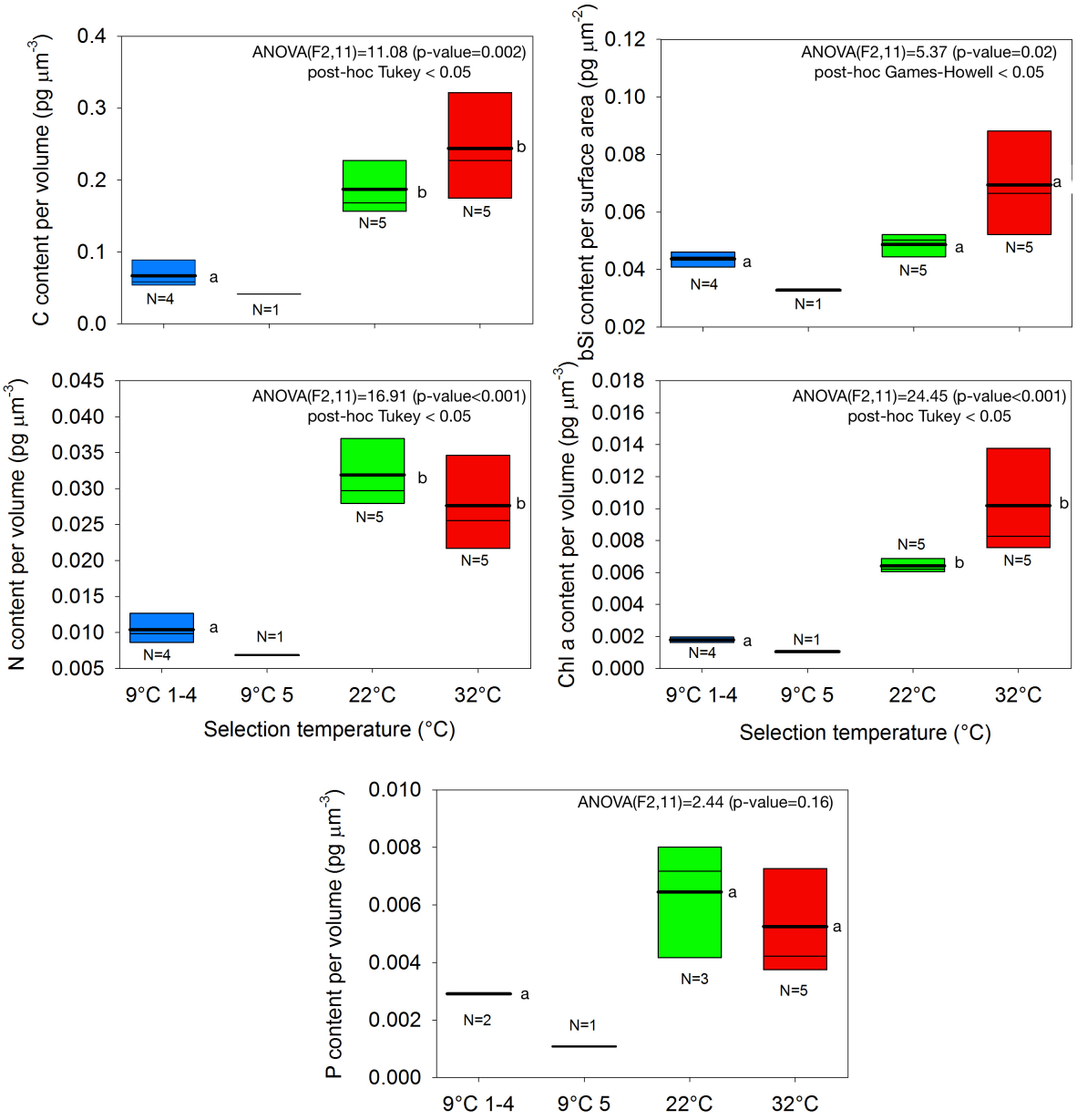

Figure 3.3.10: Cellular carbon, nitrogen, biogenic silica and chlorophyll a content per cellvolume for all three selection lines. Control cultures in green, high temperature and low temperature treatments in red and blue respectively. Thick black lines in the box plot represent mean while thin black lines respond to the median. Small letters next to the box indicate significant differences to other boxes. Boxes sharing the same letter(s) are not significantly different, while different letters indicate significant differences. Values were significantly different at a p-value cut-off of 0.05.

Furthermore, C:N, C:bSi, C:chl a and N:P ratios were calculated for the three selection lines (Figure 3.3.11). The general Redfield ratio of carbon to nitrogen (C:N) has been reported as 6.6 carbon per nitrogen (Redfield, 1958). This corresponds well with the low temperature cultures (blue) with a C:N of 6.4 ±0.58. Both, low (1-4, blue) and high (red) temperature selection lines

showed a significantly higher C:N ratio in comparison to the control culture ($p<0.05$). High temperature cultures (red) showed an C:N of 8.7 ±0.7 while the control cultures (green) showed a slightly lower C:N of 5.8 ±0.4 (Figure 3.3.11). The C:bSi ratio (Figure 3.3.11) increased significantly with increasing temperature from 9°C to 22°C and showed no further significant increase to 32°C. Low temperature cultures (blue) showed the lowest C:bSi ratio with 1.36 ±0.34, while 22°C and 32°C had ratios of 2.86 ±0.64 and 2.79 ±0.27, respectively. Nitrogen to phosphate ratios showed no significant difference between control and low and high selection lines and varied between 2.61 ±0.27, 4.61 ±1.68 and 5.84 pg N to pg P ±3.13 pg N to pg P for low temperature, control and high temperature lines, respectively. However, the small replicate (5, light blue) of the low temperature lines showed a significantly higher N:P ration with 13.1 pg N to pg P ±2.62 ($p<0.05$). The C:chl a ratio (Figure 3.3.10) showed no significant changes with selection temperature ($p<0.05$) however, low temperature cultures (blue) had the highest C:chl a with a ratio of 40.8 ±5.97. High (red) and control (green) cultures C:chl a ratios were reduced by almost 50% in comparison to low temperature cultures and reported as 24.6 ±2.04 and 24.2 ±2.13 respectively.

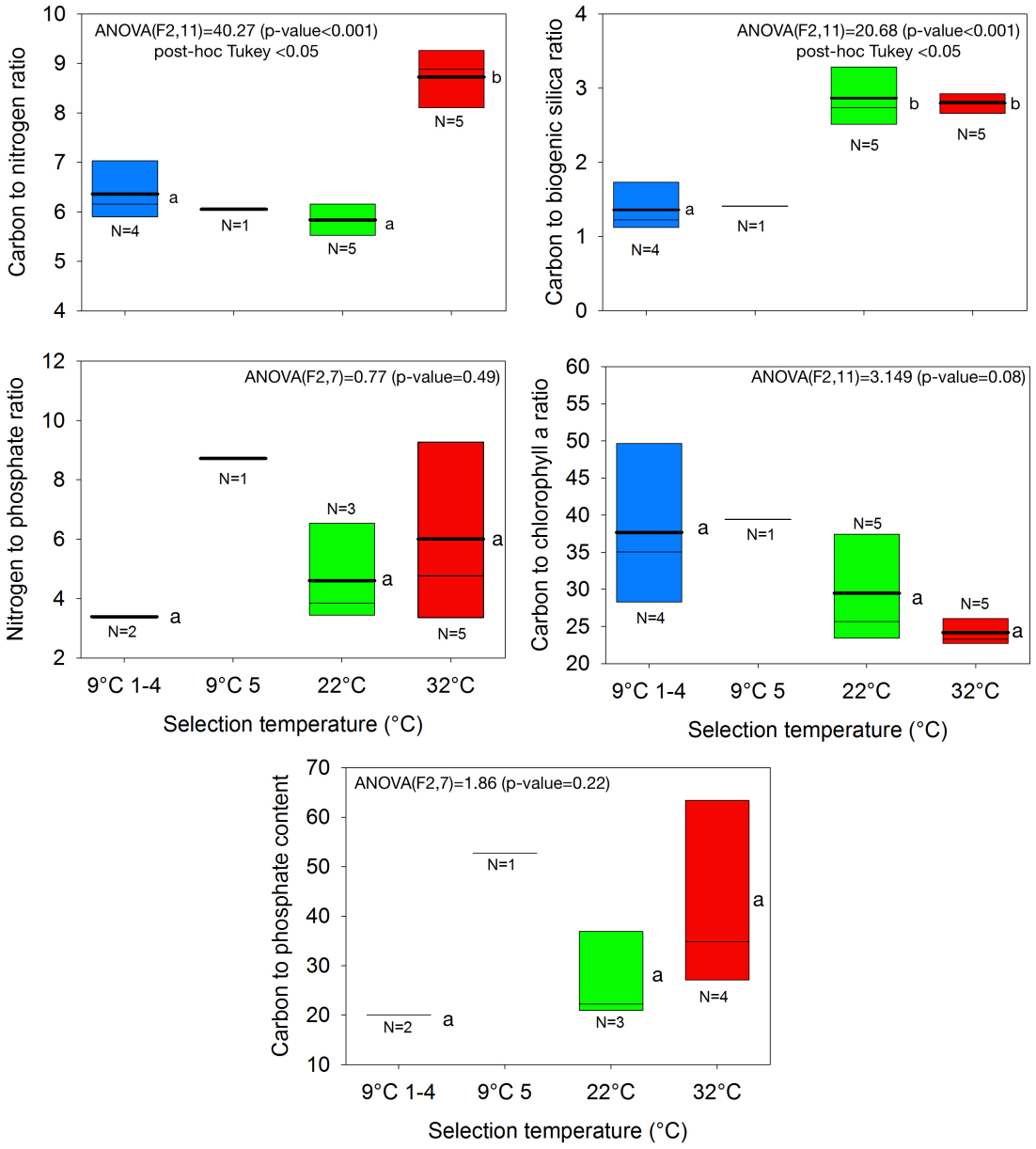

Figure 3.3.11: Carbon to nitrogen, carbon to biogenic silica, carbon to chlorophyll a and nitrogen to phosphate ratios for all three temperature treatments after 300 generations of temperature selection. Control cultures in green, high temperature and low temperature treatments in red and blue, respectively. Thick black lines in the box plot represent median. Small letters next to the box indicate significant differences to other boxes. Boxes sharing the same letter(s) are not significantly different, while different letters indicate significant differences. Values were significantly different at a p-value cut-off of 0.05.

3.4 Discussion

3.4.1 Physiological responses to high and low temperature selection

Experimental evolution approaches are a newly adopted method to study the effects of climate change on species adaptation. The main focus was, until recently, the effect of increasing CO_2 concentrations on marine phytoplankton. However, recent studies have shown the overarching effect of temperature on phytoplankton physiology is playing a more important role than nutrient availability or CO_2 concentration (Boyd *et al.*, 2015*a*; Schlüter *et al.*, 2014). However, not much is known about long-term and adaptive responses of marine phytoplankton to temperature changes. In order to study the effects of ocean warming on diatoms, we subjected the model diatom *Thalassiosira pseudonana* to 300 generations of low (9°C) and high (32°C) temperature selection under nutrient replete, semi-continuous culture conditions.

Evolved thermal tolerance has been shown to evolve within 100 generations in the green algae *Chlorella vulgaris* (Padfield *et al.*, 2016), meaning that growth rates under the selection temperature increased over time. This agrees with results obtained for this thesis. A study with *E. hux* reported similar trends for high temperature adaptation (Schlüter *et al.*, 2014). *E. hux* was selected at its upper thermal limit and showed increased fitness under the selection temperature compared to the ancestral conditions, which the authors argue suggest a shift in *E. hux's* thermal reaction norm. Furthermore, similar trends have been observed for yeast (Caspeta and Nielsen, 2015) and flagellates (a shift from 15°C to 70°C) (Dallinger, 1887) under high temperature selection. We found that the temperature optimum of selection lines potentially showed flexibility and responded to the selection pressure by shifting optimum (Figure 3.3.7, left), however, this trend was not significant due to the lack of replication.

High temperature replicates showed a greater variability than low temperature replicates (with one exception, see below). This might suggest population differences between replicates of the high temperature lines. Since we did not start the experiment from a single cell (genotype), it is likely that the replicate cultures consisted of populations with pre-existing genetic variability. Slight differences in fitness between the different genotypes (novel or pre-existing) in the

replicate populations might have resulted in different genotypes being selected in each replicate (Bell, 1997*a*). This was even more prominent in the low temperature selection lines which produced a distinct phenotype in one of the five replicates. Interestingly, no such divergence has been reported for *Emiliania huxleyi* or *Chlorella vulgaris* (Padfield *et al.*, 2016) under high temperature selection (Schlüter *et al.*, 2014).

Strikingly, the average cell diameter of one replicate, of the low temperature lines decreased after the initial increase to a value similar to the control lines (22°C) (Figure 3.3.4). This sudden reduction coincided with a drop in the cell concentration reached between transfers for all low temperature lines, however only replicate 5 reduced its average diameter afterwards. Two different processes can be identified to cause such a divergence between replicates. The starting population of the low temperature lines was likely to comprise of cells with a range of cell diameter. Low temperature selected for cells at the bigger diameter spectrum and these cells gradually rose to dominance in all replicate populations. However, some cells with the smaller average diameter might have still been present in the population after the diameter increase. The speculated drop in fitness, which coincided with the diameter decrease, could have given the smaller cells a fitness advantage allowing them to reach a different fitness peak, that required small cell size, and a high N to P ratio (Figure 3.3.11), compared to the bigger cells. However, the diameter reduction occurred over a much quicker time scale than the increase. This might suggest that a mutation, with a strong fitness benefit, occurred in the smaller cells which were then able to out compete the bigger cells. Furthermore, if some environmental variability (technical fault of incubator, different batch of medium etc.) caused this drop in fitness, a return to stable conditions would have again select against the smaller cells.

The diameter of the smaller, low temperature line stabilized shortly before the end of the experiment. Surprisingly, we were not able to find a significant different growth rate between the two phenotypes but the smaller lines were able to reach a higher yield in stationary phase (Appendix, Figure 6.2.1). However, temperature response curves differed significantly between the two phenotypes, with the smaller type showing a higher optimum and maximum temperature than the bigger type (Figure 3.3.8). This might suggest that low temperature, and thus metabolic rate was growth limiting as a smaller cell size is generally linked to a higher growth rate (see

Finkel (2010).

Changes in cell volume over the course of the experiment support the ”temperature size rule” (TSR) (Atkinson, 1994), which suggests a negative correlation between size and temperature, about 2.5% reduction in cell volume for every °C increase (Atkinson *et al.*, 2003). However, not all diatoms seem to follow this rule (Adams *et al.*, 2013). Low temperature Tp replicates 1-4 showed bigger cell volume compared to control conditions (22°C). Lowered metabolic rate, due to low temperature, might result in delayed life cycle completion and a higher build up of biomass (Heinle, 2013). In contrast, increased metabolic rates, under higher temperatures, accelerate life cycle completion at the expense of cell size (Atkinson *et al.*, 2003). Additionally, lowered metabolic rate can be compensated by increasing ribosome concentrations. This is a general response to a decline in temperature and has been suggested for other organisms (Loladze and Elser, 2011). Interestingly, increased ribosome concentrations have been shown to be linked to duplication of rDNA and length of their intergenic spacer (Elser *et al.*, 2000). Thus, this might underline the link between genome size, rDNA copy number and cell volume which has been shown to vary between and within populations (Shuter *et al.*, 1983; Zhu *et al.*, 2005; Von Dassow *et al.*, 2008; Vargas *et al.*, 2015; Elser *et al.*, 2000).

High temperature selection resulted in slightly bigger cells, and was in agreement with Atkinson (2003) who showed that under sub-lethal high temperatures, cell size can increase. This might be an adaptive strategy to increase viability (Gallagher, 1983; Montagnes and Franklin, 2001). Indeed, high temperature selection (32°C) was carried out at sub-lethal temperatures close to the upper temperature limit (35°C). A similar trend of cell size increase under high temperature was found for the zooxanthelle *Symbidium sp.* (Kirralee Baker, personal communication) and the coccolithophore *P. carterae* (Heinle, 2013).

The photosynthetic quantum yield of PS II (Fv/Fm) of the low and high temperature lines reduced significantly in comparison with the control lines after the start of the experiment. Interestingly, the Fv/Fm of the low temperature cultures increased continuously over the course of the experiment, while it remained stable for the high temperature lines (Figure 3.3.3). A study by Mock et al. (2005) on long-term low temperature acclimation in the polar diatom *Fragilariopsis cylindrus* (Fc) found full recovery of Fv/Fm to the level measured in their control culture. Low temperature treatment interrupts the photosystems efficiency to distribute absorbed

energy to metabolic sinks (Ensminger *et al.*, 2006), for example, by slowing the electron transport across thylakoid membranes due to lowered membrane viscosity (Morgan-Kiss *et al.*, 2002). This increases the minimal fluoresces (F0) as more energy is quenched as fluorescence and results in a lowered Fv/Fm indicative of plant stress (Horvath *et al.*, 1987; Morgan-Kiss *et al.*, 2002). Low temperature treatment exacerbates photodamage of the photosystems D1 protein (Huner *et al.*, 1993). In cyanobacteria it has been shown that the repair of D1 proteins, namely the reassembly of the protein into the membrane, is inhibited by decreased membrane fluidity (Nishida and Murata, 1996). Thus, increasing unsaturated fatty acids in thylakoid membranes might allow recovery of the Fv/Fm (Falcone *et al.*, 2004; Morgan-Kiss *et al.*, 2002), due to elevated dissipation of excitation energy and increased repair rates of D1. Additionally, *Synechococcus* has been shown to possess two different forms of the D1 protein (D1 and D1:2) with D1:2 having a higher resistance to low temperature induced photoinhibtion (Campbell *et al.*, 1995). This might suggest that differential expression of D1 genes (psbA) aided Fv/Fm recovery (Mock and Hoch, 2005; Campbell *et al.*, 1995).

Unfortunately, the experiment presented here was terminated before the Fv/Fm stabilized, therefore it is unclear whether the low temperature lines would have reached the Fv/Fm of the control. However, the findings suggest that Tp, allowing enough time, is able to adjust its photosynthesis under low temperature conditions. Nevertheless, adaptation under low temperature selection might take longer than 300 generations.

3.4.2 Elemental and pigment changes under temperature selection

Cellular stoichiometric composition can be attributed to cell biochemistry with proteins and nucleic acid contributing most to the organic N pool, while ribosomal RNA and phospholipids are associated with the organic P pool. Additionally, N is found in chlorophyll a, b and c, and amino acids (Elser *et al.*, 1996; Geider and La Roche, 2002). In general, the increasing trends in cellular carbon, nitrogen and chl a content with temperature increasing from low to optimum are in accordance with other studies (Thompson, 1999; Berges *et al.*, 2002; Stramski *et al.*, 2002). Carbon to nitrogen ratios remained stable from 9°C to 22°C but increased significantly at 32°C. A similar response was found for *E. hux* to high temperature selection (Schlüter *et al.*, 2014), but stands in contrast to studies by Berges et al. (2002) and Thompson et al. (1999) where C:N ratios

of Tp were significantly lower at low temperature. Neither study investigated temperatures above Tps optimum temperature.

The increased C:N ratio at 32°C appeared to be driven by decreased cellular nitrogen as well as increased carbon content. Decreases in the different N pools (proteins, nucleic acid, amino acids or chl a) could underline this trend. However, studies on *Cylindrotheca fusiformis* and green algae found increases in amino acid concentrations with increasing temperature (Bermúdez *et al.*, 2015; James *et al.*, 1989). Decreases in protein concentrations (and presumably N) have been reported for several marine diatoms in response to temperatures raised to the optimum (Thompson, 1999), however, in this study, decreased N content was found for Tp cultivated above its optimum temperature. This is in accordance with *Nannochloropsis* and *Skeletonema costatum* where declining concentrations of amino acids and proteins (and presumably nitrogen) were found in response to temperature increases above optimum (James *et al.*, 1989; G. Falkowski, 1977). Enzyme reaction speeds can be temperature dependent and increase with temperature, enhancing biochemical reactions (Gillooly *et al.*, 2001). Increased biochemical reactions might be complemented by more protein and nucleic acids, potentially increasing the cellular nitrogen pool (Thompson *et al.*, 1992*a*). These arguments stand in contrast to my findings, and additional research would be necessary to clarify the effect of decreased nitrogen content of high temperature cultures resulting in an increased carbon to nitrogen ratio. The increase in carbon content might be related to the increase in chlorophyll a (Figure 3.3.10) as it contains much more carbon than nitrogen (Geider and La Roche, 2002). Thus, an increase in chlorophyll a content might have a bigger effect on the cellular carbon content than on the nitrogen content. Additionally, even though the C:N ratio was statistical significantly higher for 32°C evolved cultures, the values were in the range of ratios (5.1 to 13.3) observed for Tp and other diatom species (Ho *et al.*, 2003; Leonardos and Geider, 2004; Brzezinski, 1985).

Several models (Toseland *et al.*, 2013; Daines *et al.*, 2014; Yvon-durocher *et al.*, 2015) predict an increased N:P with increasing temperature due to reduced synthesis of ribosomes under high temperature in contrast to increased synthesis of ribosomes and their P-rich rRNA under low temperatures. This agrees with results obtained for this thesis, where N:P ratios increased with increasing temperature (Figure 3.3.11). However, this seems to be mostly driven by a decline in nitrogen (Figure 5).

Surprisingly, low temperature replicate 5 had an five times higher N:P ratio (Figure 3.3.11) than the other low temperature replicates (1-4). This appeared to be driven by a reduction in phosphate content per cell (Figure 3.3.9). This stands in contrast to low temperature adaptations found in other phytoplankton species where P-rich ribosome concentrations were increased to counteract the reduced reaction rates at cooler temperatures (Loladze and Elser, 2011). However, a study on *Skeletonema cf.* reported increased protein (and presumably nitrogen) concentrations under low temperature (Jørgensen, 1968). Ribosomes are costly in P to synthesize, compared to N-rich proteins, assuming cheap P phosphorylation (Flynn *et al.*, 2010). Thus, the small low temperature phenotype might have employed a strategy of increasing protein abundance, or functional protein modification, without increasing ribosome abundance. Raven and Geider (1988) argue that both strategies are not mutually exclusive. Either strategy must have provided a strong fitness benefit, as they were able to out compete the bigger cells. Additionally, if they did not duplicate their rDNA genes, as suggested for the bigger cells, this might also explain the smaller cell size, as genome and cell size are suggested to be linked (Vargas *et al.*, 2015; Gregory, 2001). The combined differences (cell volume, temperature response curve and N:P ratio) of replicate 5 compared to the other low temperature replicates seem to suggest physiological and possibly genotypic differences, however, gene expression patterns between low temperature replicates were very similar (see chapter 4). Nonetheless, the high N:P ratio could have also been caused by a methodical error, such as harvesting cells at the end of exponential phase where potential P-limitation could increase the N:P ratio. However, cells, for all replicates, were harvested in mid-exponential phase (about 5x10^5 cells mL^{-1}).

It would have been interesting to see whether the elemental compositions changed initially due to the temperature stress or if the endpoint measurements were due to gradual changes over time. Sampling was indeed conducted to include more time points, however, due to instrument failure, additional time points were unfortunately lost.

Temperature selection resulted in chlorophyll a concentrations changes (Figure 3.3.9), where chl a increased with increasing temperature and decreased under low temperature. A change to a lower temperature under constant light, as in this experiment, mimics high light conditions resulting in an increased excitation of the photosystem as biochemical reactions are slowed were as photochemical reaction are not. This can cause damage to the photosynthetic apparatus as

cells might be unable to dissipate excess excitation energy and results in low temperature chlorosis (Geider *et al.*, 1998). This effect potentially results in the decrease of chl a and light harvesting pigments (Geider, 1987; Davison, 1991). Another mechanism to counteract low temperature photoinhibition (Davison, 1991) is the production of carotenoids. They act as photoprotective pigments by scavenging reactive oxygen species and blocking excess light and can thus counteract photoinhibition (**?**). Similar responses have been found for cyanobacteria (Tang and Vincent, 1999; Várkonyi *et al.*, 2002), diatoms (Anning *et al.*, 2001; Spilling *et al.*, 2015) and green algae (Krol *et al.*, 1997). Additionally, conversion of light harvesting pigments to quenching pigments can reduce the energy absorption by chl a, as they are unable to pass their energy towards chl a (Horton *et al.*, 1999; Niyogi *et al.*, 2005). However, we did not measure pigments other than chl a and can thus only speculate.

3.5 Summary and conclusion

Thermal tolerance has been shown above to evolve rapidly within 300 generations of low and high temperature selection providing first insights into phenotypic outcomes of temperature adaptation in diatoms. Phenotypes seemed to reflect adaptive processes inside the cell resulting in distinct phenotypes of low and high temperature lines. The low temperature phenotype of replicates 1-4 was characterized by low chl a content, lowest element content, big average cell diameter, and lowered optimum and maximum temperature. In contrast, low temperature replicate 5 showed an even lower element content but the highest nitrogen to phosphate ratio as well as a small average diameter and a higher maximum temperature compared to the other low temperature replicates. The divergent evolution of low temperature replicates was unexpected and highlights intraspecific variability in evolutionary outcomes. Elemental composition varied between temperature treatments and were likely related to intrinsic adjustments of resource allocation and assembly machinery. Thus, diatoms possess cellular plasticity in terms of cell size, resource usage and allocation to physiologically adapt to temperature changes. This capacity will certainly play a role in phytoplankton responses to ocean change and have implications for biogeochemical cycles.

Chapter 4

Molecular underpinnings of temperature adaptation in *Thalassiosira pseudonana*

4.1 Introduction

Experimental evolution studies with microbes are a useful method elucidating adaptive pheno- and genotypic modifications to a multitude of environmental conditions. Evolutionary phenotypic outcomes to a variety of selection pressures (CO_2, temperature, nutrients, salt) have been investigated in *Escherichia coli* (Lenski and Bennett, 1993), *Saccharomyces cerevisiae* (Dhar *et al.*, 2011), *Chlamydomonas reinhardtii* (Perrineau *et al.*, 2013), *Emiliania huxleyi* (Lohbeck *et al.*, 2012), *Trichodesmium* (Hutchins *et al.*, 2015) and *Skeletonema marinoi* (Scheinin *et al.*, 2015), however, only a limited number made use of genomics-enabled tools to identify the molecular underpinnings of the evolved phenotype (for example (Benner *et al.*, 2013; Ying *et al.*, 2015; Helliwell *et al.*, 2015; Hutchins *et al.*, 2015; Blount *et al.*, 2012)). Two main questions that can be answered by using an genomics-enabled approach are: (i) How does the environment influence the organismal transcriptome and (ii) How are gene expression differences translated into a phenotype?

One avenue to identify such underpinnings is the analysis of gene expression changes to novel

environments, generated by RNA sequencing (RNA-Seq), as the transcriptome, underpinning the evolved phenotype, is orchestrated by the combination of environment (E, plasticity), organismal genotype (G, evolution) and the resulting genotype by environment interaction (G x E) (Debiasse and Kelly, 2016; Rockman, 2008; Zhou and Stephens, 2012). Thus, plastic responses can be disentangled from evolutionary responses (**?**). RNA-Seq provides two advantages over the commonly used microarray technology: (i) it provides sequence information of the entire complement of mRNA in a organism and (ii) abundance information of million of reads (De Wit *et al.*, 2012). This approach allows identification or characterisation of genes responsive to environmental stimuli. For example, *Emiliania huxleyii's* (E. hux) response to increased CO_2 levels provided great detail into the genetic regulation of carbon concentration mechanisms (CCM) through up-regulation of putative CCMs (**?**Benner *et al.*, 2013). Additionally, comparative transcriptomics can be used to identify adaptive and plastic responses. For example, a study with the intertidal snail *Chlorostoma funebralis* showed that temperature sensitive populations differentially expressed a larger set of genes under heat stress than the control population (Gleason and Burton, 2015). On the other hand, thermotolerant populations of seagrass showed a greater transcriptome response than thermosensitive populations (Franssen *et al.*, 2014). These two responses could be interpreted as either an adaptive response, or that an increased differential gene expression indicates a greater stress response (Debiasse and Kelly, 2016). In addition to gene regulation, expression differences might be a result of epigenetic mechanisms underpinning phenotypes before genotypic changes arise (West-Eberhard, 2005).

One potential issue of whole-genome transcriptomics is the generation of long lists of differentially expressed genes with known and unknown function. The typical annotation rate of genes is about 50% limiting our understanding of transcriptome responses as many of the unannotated genes exhibit differential expression (Logan and Buckley, 2015). Lists of annotated differentially expressed genes can be grouped by function allowing identification of expression differences between groups. One commonly used source of gene groups is the Gene Ontology (GO) (Ashburner *et al.*, 2000). Here, genes are grouped into three ontologies: Biological process, molecular function and cellular component. Differential expressed genes can be assigned to different cellular processes, falling into the three ontologies, which then allows testing for

enrichment of certain ontologies or GO-terms, as one gene or gene product can perform functions in multiple processes (Blake *et al.*, 2015). Thus, a long list of differentially expressed genes can be condensed into biological knowledge and potentially provide glues for further investigation.

Another database for functional annotation of gene or protein sequences is InterPro. With InterPro, sequences are analysed for their functional sites and domains and grouped into protein families. This provides a more detailed functional annotation than GO terms and can be a useful tool to gain more functional insights of differentially expressed genes (Apweiler *et al.*, 2001; Hunter *et al.*, 2009; Finn *et al.*, 2016).

Combining transcriptomics and genomics has helped in some instances to elucidate adaptive genotypic changes in response to novel environments. For example, in a study on experimental evolution of *E. coli* to different oxygen levels, transcriptome profiles guided the search for beneficial mutations (Puentes-téllez *et al.*, 2014). Whole-genome resequencing of evolved phenotypes has shown that interactions, for example potentiation, between mutations can occur. In case of a key innovation in *E. coli*, the genetic background had to acquire other mutations in order for the new function to arise through mutation (Blount *et al.*, 2012). Identifying potential beneficial mutations is the first step in linking evolved geno- and phenotype. Allelic replacement experiments, where identified mutations are brought into the ancestor background provide the causal link between *de novo* mutation and observed phenotype (Conrad *et al.*, 2011). Bantinaki et al. 2007, for example, used allelic replacement experiments to identify the mutation causing a Wrinkly Spreader phenotype in *Pseudomonas fluorescens*. However, not only mutations contribute to organism adaptation, transposable elements and DNA methylation, among others, can also contribute to phenotypic changes under novel conditions.

Epigenetics refers to the study of phenotypic changes that are not underlined by changes in the genotype. Genome methylation, a epigentic mechanism, is the process of adding methylation marks to cytosine and/or adenine, however cytosine marks have been studied in more detail. These methylation marks have been shown to activate, repress or disable gene expression and can be inherited (Maumus, 2009). Natural variation in DNA methylation marks between ecotypes of *Arabidopsis* has been shown to relate to diverse phenotypes under varying environmental

conditions (Bossdorf *et al.*, 2010; Riddle and Richards, 2002). Additionally, Hutchins et al. 2015 suggested that DNA methylation changes in *Trichodesmium* underpin adaptation to high CO_2 concentrations. Thus, environmentally induced methylation patterns can have similar effects on a phenotype as DNA mutations (Biémont and Vieira, 2006), however, it is currently unclear how much methylation contributes to phytoplankton adaptation. An important regulatory function of DNA methylation is the activation and silencing of transposable elements (TEs).

TEs are DNA sequences that can be spread through an organisms genome in a copy-and-paste or cut-and-paste manner. Thus, TEs are likely to impact adaptation of organisms through their ability to generate mutations, altering gene function or expression (Maumus *et al.*, 2009). Additionally, TEs have been shown to be response to changes in environmental conditions (Casacuberta and González, 2013). Based on their mode of transposition, two classes of TEs have been established: Retrotransposons (Class I) that replicate through an RNA intermediate and DNA transposons (Class II) that lack the RNA intermediate. The most abundant class of transposons in diatoms are retrotransposons (Maumus *et al.*, 2011) and as they replicate through an RNA intermediate, whole-genome transcriptomics can be used to identify TE activity. For example, Pt showed increased expression levels of a retrotransposon under nitrate starvation which might have aided adaptation to the stress (Maumus *et al.*, 2009). Thus, whole-genome transcriptomics can provide first insights into epigenetic mechanisms, such as DNA methylation and transposons, employed by the organism to adapt to novel environmental conditions. Furthermore, whole genome resequencing can be implemented to search for novel transposition events and their potential effects on affected genes.

Here, I report the first genome-wide expression study using RNA-Seq to elucidate the connection between transcriptome and phenotype after low (9°C) and high (32°C) temperature selection over 300 generations in the marine diatom *Thalassiosira pseudonana*. Whole-genome resequencing of a high temperature replicate provided further insight into the genetic basis of temperature adaptation. To the best of my knowledge, this is the first experimental evolution study with a diploid marine diatom implementing genomic approaches to elucidate molecular underpinnings of temperature adaptation.

The availability of a genome sequence makes Tp a good candidate for genomic-enabled experimental evolution studies. However, using a non-model organism can be problematic in

terms of gene function due to the high number of unannotated genes compared to established model systems. About 50% of Tp genes could be annotated (Armbrust *et al.*, 2004) leaving half of Tp genes with no known function. Additionally, a lot of the genes are not well characterized in terms of their place in pathways or regulatory functions, hindering predictions about their importance under novel conditions. Tp's fast generation time and the ease of maintaining large population sizes as well as the possibility to store a "fossil" record (cryopreservation of ancestral populations) make them well suited for the use in experimental evolution studies. Gene expression pattern after 50 and 300 generations of temperature selection were compared to the control and furthermore, analysed for their biological function. Evolved phenotypes from chapter 3 guided selection of potential relevant pathways for the evolution of thermal tolerance. For example, biosynthesis of unsaturated fatty acids and their desaturation were selected based on the recovery of the photosynthetic efficiency of low temperature replicates and that membrane fluidity plays a key part in low temperature adaptation. Expression changes of the synthesis pathway of photosynthesis relevant pigments (chlorophyll a and photoprotective pigments) were investigated based on the changes in chlorophyll a content under both selection temperatures. Variations in phosphate content of selections lines was argued to be underpinned by expression changes of ribosomal genes, thus, the synthesis of ribsomes was selected. Additionally, we examined the role of epigenetic processes in aiding the evolution of thermal tolerance through expression of transposable elements and DNA-methyltransferases. One high temperature selection replicate was selected for genome resequencing to gain preliminary insights into the role of mutations for temperature adaptation.

The following hypothesis were tested with the above described approach:

- Phenotypic responses to selection temperatures in chapter 3 were underpinned by transcriptomic responses.
  - Low temperature adaptation is mediated by increases in the translation machinery.
  - Adjustments to membrane fluidity aided low temperature adaptation of the photosystem.
- Short-term and long-term responses of the cellular transcriptome differ.

- Transposon activity aids temperature adaptation.

4.2 Material and Methods

4.2.1 Transcriptome analysis

4.2.2 RNA sampling and extraction

Three replicates of 100 mL each were sampled per biological replicate for RNA and DNA extractions. Filtrations were done onto 25 mm diameter, polycarbonate filters (PC) (1.2 µm, Whatman, ME, USA) under low pressure (<150 mbar Hg). Filters were quickly rinsed with 0.2 µm filtered artificial seawater, folded into half, snap frozen in liquid nitrogen and stored at -80°C until extractions. Total RNA was extracted for all replicates of each temperature treatment using the Direct-zol^{TM} RNA MiniPrep (Zymo Research, Irvine, CA, USA) followed by an DNA digest with DNAse I (Zymo Research, Irvine, CA, USA) for 1 hour at 37°C. Purity of RNA was measured with a NanoDrop (Thermo Fisher Scientific, Waltham, MA, USA). Furthermore, integrity was checked with a 2% denaturating formaldehyde gel.

4.2.2.1 Library preparation and sequencing

To compare transcriptomes of the cultures under temperature selection, all 5 replicates were sequenced at the start and after 50 and 300 generations of the experiment. Total RNA was sent to The Genome Analysis Centre (TGAC, Norwich, UK) for library preparations and sequencing. Sequencing libraries were prepared for each replicate using TrueSeq RNA kits (Illumina). Libraries were reverse transcribed with random hexamers. 15 and 16 samples were multiplexed and each sequenced on a single lane of an Illumina HiSeq 2000 platform (100 bp paired end reads).

4.2.2.2 Sequence analysis

RNASeq analysis was carried out with support from Dr. Andrew Toseland (University of East Anglia, School for Computational Science) and Dr. Jan Strauss (EMBL Hamburg, Germany).

The work flow of the transcriptome analysis is shown in figure 4.2.1. The quality of raw reads was checked with FastQC (v0.11.3) (Andrews, 2010) and quality trimming and sequencing adapter trimming was performed with trimgalore (v.0.4.0) and Python package cutadapt (v.1.8.1) with a minimum sequence return length of 50 bp. Trimmed sequences had an average length of 101 bp and were mapped to the *Thalassiosira pseudonana* reference genome (Joint Genome Institute, `http://genome.jgi-psf.org/Thaps3/Thaps3.download.ftp.html/Thaps3_chromosomes_assembly_chromosomes.fasta.gz`) using TopHat2 (v2.0.14) (Kim *et al.*, 2013) in conjunction with Bowtie2 (v2.2.2) (Langmead and Salzberg, 2012), which allows splice junction sensitive mapping. Mapping was done using default settings and allowing for a maximum of 2 mismatches. Only reads that mapped uniquely and concordantly for both reads of a pair were retained. Mapped reads were sorted with samtools (v1.2) by name for consecutive generation of read count tables. The Python package HTSeq-count (v0.6.0) (Anders *et al.*, 2014) was used to count unique reads mapping to genomic features, guided by a annotation file from JGI (`http://genome.jgi-psf.org/Thaps3/Thaps3.download.ftp.html/Thaps3_chromosomes_geneModels_FilteredModels2.gff`), in mode ”union”. Only reads that did not overlap/match to more than one genomic feature were counted. Read pairs that match partly or completely to more than one feature were reported as ambiguous.

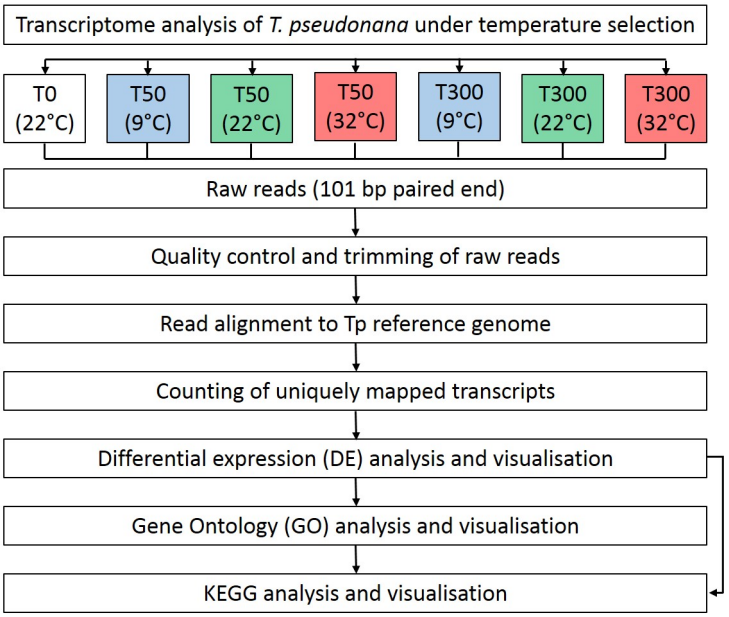

Figure 4.2.1: Work flow of bioinformatic analysis of transcriptome profiles.

4.2.2.3 Differential gene expression analysis

Differential gene expression analysis was performed by feeding the HTSeq generated count tables into a R script (courtesy of Dr. Jan Strauss) that utilizes the R Bioconductor package edgeR (Robinson *et al.*, 2009). Replicates were grouped according to their treatment and weakly expressed genes were removed from the dataset. Counts were normalized according to the "trimmed mean of M values" (TMM) in order to avoid biases towards highly abundant transcripts (Robinson and Oshlack, 2010). EdgeR's generalized linear model was used to identify differentially expressed genes for each treatment group. Differentially expressed genes were then filtered based on a two fold change and a adjusted p-value of <0.05. Each temperature treatment was compared at 50 and 300 generations to T0 (baseline) and was reported as T0vsT1 and T0vsTE.

4.2.2.4 Multidimensional scaling of transcriptomes

Transcriptomes of the selection and control lines were used for visualising differences in the gene expression patterns between time points and selection temperature. RPKM (Reads Per Kilobase Million) values were calculated based on the count table generated by HTSeq using the EdgeR package in the R environment which normalizes read counts to library size and gene length (Robinson *et al.*, 2009). The generated RPKM values were log transformed and used for multidimensional scaling to visualise differences in the gene expression pattern between selection lines and time points. Distances between points are based on the Manhattan index which calculates absolute differences between vectors. RPKM values take all expressed genes into account while differential gene expression data is based on differences of expression. For example, a gene that shows a high expression (high RPKM value) at one time point can have a similar high expression at another time point and would therefore not be differentially expressed as the relative difference between the time points would be close to 0.

4.2.2.5 Gene ontology (GO) analysis of differentially expressed genes

Based on the differentially expressed genes, gene ontology analysis was performed with the R Bioconductor package goseq (Young *et al.*, 2010). GO analysis is used to identify enriched GO-terms based on gene annotations and associated GO-terms. These GO-terms are grouped into three domains:

- Cellular component
- Molecular function
- Biological process

GO-term analysis was corrected for gene length which was provided by Andrew Toseland from the *T. pseudonana* filtered gene model annotation file. The gene length and differentially expressed genes were given to goseq (Young *et al.*, 2010) to test for enriched GO-terms. GO-term analysis was performed for different sets using up-regulated and down-regulated genes (GLM likelihood ratio test, $p<0.05$). Significantly overrepresented GO-terms (Wallenius approximation, Benjamini-Hochberg adjusted $p<0.05$) were visualized in semantic similarity-based scatterplots and treemaps with the web based tool ReVIGO (`http://revigo.irb.hr/`, last accessed 27th of April 2016) (Supek *et al.*, 2011).

4.2.2.6 Pathway analysis

Insights gained from the phenotypic responses reported in chapter 3 were used to select pathways that were potentially involved in the evolution of thermal tolerance. The KEGG pathway database (`http://www.kegg.jp/kegg/pathway.html`, last accessed 1st of June 2016) was utilized to search for relevant genes involved in the selected Tp pathways. These genes were then matched with their respective RPKM value and used for visualization as heatmaps with gene-e (`http://www.broadinstitute.org/cancer/software/GENE-E/index.html`, last accessed 1st of June 2016).

4.2.3 Genome resequencing of high temperature replicate

To gain initial insights into potential adaptive genomic changes, one replicate of the evolved high temperature selection lines was selected for genome resequencing together with the T0 sample. Due to financial restrictions we were only able to sequence one replicate and the starting population (T0).

4.2.3.1 DNA sampling and extraction

Three technical replicates of 100 mL each were sampled to extract DNA from a high temperature selection replicate. Filtrations were done onto 25 mm diameter, polycarbonate filters (PC) (1.2 µm, Whatman, ME, USA) under low pressure (<150 mbar Hg). Filters were quickly rinsed with 0.2 µm filtered artificial seawater, folded into half, snap frozen in liquid nitrogen and stored at -80°C until extractions. Total DNA was extracted using the EasyDNA kit (Invitrogen, Carlsbad, CA, USA) with some adjustments. Cells were washed off the filter with pre-heated (65°C) Solution A and the supernatant was transferred into a new tube with one small spoon of glass beads (425-600 µm, acid washed) (Sigma-Aldrich, St. Louis, MO, USA). The samples were then vortexed three times in intervals of three seconds to break the cells. RNase A was added to the samples and incubated for 30 min at 65°C. The supernatant was transferred into a new tube and Solution B was added followed by a chloroform phase separation and an ethanol precipitation. DNA was pelleted by centrifugation and washed several times with Isopropanol, air dried and suspended in 100 µL TE buffer. DNA concentration was measured with a Nanodrop (Thermo Fisher Scientific, Waltham, MA, USA), tubes snap frozen in liquid nitrogen and stored at -80°C until sequencing. Furthermore, DNA integrity was checked with agarose gel electrophoresis (0.9% agarose gel).

4.2.3.2 Library preparation and sequencing

High molecular weight DNA was sent to The Genome Analysis Centre (TGAC, Norwich, UK) for library preparations and consecutive sequencing. Whole genome sequencing libraries were prepared for each replicate using TrueSeq DNA kits (Illumina). The two samples were multiplexed and sequenced on a single lane of a Illumina HiSeq 2000 platform (100 bp PE).

4.2.3.3 Single nucleotide polymorphism (SNP) analysis of high temperature genotype

Identification of SNPs in the high temperature replicate after 300 generations in comparison to the control (T0) was conducted by Andrew Toseland (School of Computational Sciences, UEA). For this 82 and 87 million, 101 bp paired reads, for T0 and TE 32°C respectively, were quality screened with fastQC (Andrews, 2010). Reads were then filtered by removing of duplicates and reads with unidentified bases resulting N characters. Bases with a quality score below 30 were trimmed from the 3'end and Illumina adapter sequences were trimmed from the 5'end using Prinseq-lite (Schmieder and Edwards, 2011). Only reads with a length of at least 50 bp were retained. The remaining reads were mapped to the reference genome with STAR and indexed with samtools. The GATK (McKenna *et al.*, 2010) best practise guidelines were performed consisting of local realignments around indels and base recalibration based on an initial set of SNPs and indels. SNPs were called using GATK's HaplotypeCaller module (phred quality $\geq$ 30). SNP calling was based on the consensus of three callers: GATK's HaplotypeCaller; samtools and VarScan (Koboldt *et al.*, 2009). After initial calling, SNPs and indels were separated from the call files and filtered (min quality 40) before being merged into a single SNP/indel file. We then selected the intersection of each file with GATK's SelectVariants module. SNPs only present in the temperature adapted data set (combine the two sets then take the difference) were extracted and summarised with SNPeff (Cingolani *et al.*, 2012) using optimised gene models for Tp (Gruber *et al.*, 2015). SNPs within 100 bases of a predicted gene were considered up/downstream of that gene. The results from SNPeff were then used for further analysis by myself.

4.2.3.4 GO analysis of genes with SNPs

Genes with SNPs were used for GO-term analysis using the goseq package (Young *et al.*, 2010) in the R environment (**?**). Terms with a p-value$<$0.05 were used for visualization with REVIGO (`http://revigo.irb.hr/`) (Supek *et al.*, 2011). See section 4.2.2.5 for more details.

4.3 Results

4.3.1 Transcriptome sequence analysis

Sequencing on the Illumina HiSeq 2000 platform produced 9.9 to 19 million paired-end sequences (on average 13.9 million) except for one sample (T1 9C Tp1) which only produced 3093 reads and was excluded from further analysis due to poor quality reads (Table 4.3.2). Raw reads were on average 101 bases long and of high quality (average base quality score of 37). Quality control of raw reads, which included adapter trimming, returned on average 13.3 million paired reads of which 85.8% could be uniquely aligned to the *T. pseudonana* reference genome. Out of these 92.0% could be counted as unique genomic features (e.g. exons, introns and intergenic regions) and were used for gene expression analysis. 17.8% on average, could not be assigned to predicted genomic features, indicating novel transcriptional active genomic regions (e.g. transcription factors, non-coding RNAs, novel splice variants). Interestingly, the amount of reads mapping to predicted non-coding regions was lowest under control conditions and increased under high and low temperature selection.

4.3.2 Differential gene expression of low and high temperature cultures

Transcriptome analysis of the control, high and low temperature experimental cultures showed a total of 4558, 4031 and 4909 differentially expressed genes (out of 11242 genes in Tp genome), respectively, compared to the starting culture. After 50 generations (T1) low temperature selection lines showed 1844 and 2033 genes up- and down-regulated, respectively (Table 4.3.1, a). However, the number of differentially expressed genes in the low temperature lines was significantly reduced after more than 300 generations (TE). The total amount of up- and down-regulated genes for the high temperature lines increased further after 50 generations (T1) in comparison to 300 generations (TE) (Table 4.3.1, b).

Table 4.3.2: Transcriptome read output for each sample (Sample ID) and the output through the analysis pipeline (before and after quality control, mapping and counting of mapped reads).

Sample ID	**before quality control**			**after quality control**			**Mapping**	**Read counting**		
	read pairs	avg length	avg quality	avg length	avg quality	paired reads	concordant pairs	counted pairs	no feature	ambiguous
T0 Tp	15206512	101	37	100	37	14609411	12198858	11499331	1960208	289413
T1 9C Tp1	3093	discarded due to poor quality								
T1 9C Tp2	13592412	101	37	100	37	12991851	11393853	10764912	1582283	170503
T1 9C Tp3	17766818	101	37	100	37	17146957	15089322	14113416	2204831	182495
T1 9C Tp4	12268455	101	37	100	37	11661473	10250434	9660765	1421668	125708
T1 9C Tp5	13684702	101	37	100	37	12821050	11231239	10606158	1597268	142333
T1 22C Tp1	13061566	101	37	100	37	12565195	10592459	9923966	1590007	184130
T1 22C Tp2	15052101	101	37	100	37	14547269	12292442	11415966	1955494	218769
T1 22C Tp3	19348476	101	37	100	37	18054874	15003600	14339009	2176242	326919
T1 22C Tp4	19070931	101	37	100	37	18273711	15532654	14627940	2207104	271479
T1 22C Tp5	14887780	101	37	100	37	14012596	12036819	11184705	1781529	203750
T1 32C Tp1	16447497	101	37	100	37	15737955	13613331	12454815	2434041	181052
T1 32C Tp2	17493149	101	37	100	37	16478611	14188084	13380487	2230080	238679
T1 32C Tp3	14454783	101	37	100	37	13957167	12128778	11004779	2218161	153426

Continued on next page

Continued from previous page

Sample ID	before quality control			after quality control			Mapping	Read counting		
	read pairs	avg length	avg quality	avg lenght	avg quality	paired reads	concordant pairs	counted pairs	no feature	ambigous
T1 32C Tp4	10164252	101	37	100	37	9779340	8468908	7814213	1447099	116050
T1 32C Tp5	12858761	101	37	100	37	12337828	10684559	9786306	1900474	149286
TE 9C Tp1	15400646	101	37	100	37	14831216	12428559	11586117	2079202	237404
TE 9C Tp2	15813957	101	37	100	37	15335803	12928081	11782574	2385268	184541
TE 9C Tp3	13393010	101	37	100	37	12883859	10964164	10007251	1940778	170495
TE 9C Tp4	10063302	101	37	100	37	9697789	8233422	7545716	1451941	133284
TE 9C Tp5	9987215	101	37	100	37	9543308	8169071	7560643	1339898	124565
TE 22C Tp1	11978884	101	37	100	37	11386896	9473897	8857199	1494656	196398
TE 22C Tp2	13186392	101	37	100	37	12743815	10692060	9266895	1722228	193758
TE 22C Tp3	13464908	101	37	100	37	12993784	11135672	9918853	1777349	210558
TE 22C Tp4	12270789	101	37	100	37	11885617	10105324	9266895	1535547	213062
TE 22C Tp5	12705400	101	37	100	37	12268391	10379058	9516143	1616255	203129
TE 32C Tp1	11627419	101	37	100	37	10994712	9613981	8662191	1704812	126570
TE 32C Tp2	10473135	101	37	100	37	9972688	8711542	7941987	1451095	124940
TE 32C Tp3	13056467	101	37	100	37	12460804	10953515	9937441	1834520	153336

Continued on next page

Continued from previous page

Sample ID	**before quality control**			**after quality control**			**Mapping**	**Read counting**		
	read pairs	avg length	avg quality	avg lenght	avg quality	paired reads	concordant pairs	counted pairs	no feature	ambigous
TE 32C Tp4	11710638	101	37	100	37	11333925	9807222	8290498	2325837	129611
TE 32C Tp5	18678984	101	37	100	37	17681706	15446591	14037844	2639516	197424

Table 4.3.1: Overall numbers of up- and down-regulated genes (BH adjusted p-value<0.05, >2-fold increase) for each experimental setup and after 50 (T1) and 300 (TE) generations.

	a) 9°C		b) 32°C	
	T1	TE	T1	TE
Up	1844	438	870	1356
Down	2033	594	1163	1586

Multidimensional scaling of transcriptomes (RPKM values) for all replicates is shown in Figure 4.3.1. Replicate samples for each temperature treatment cluster well together as shown by the distances between replicates. A clear difference was seen for low temperature time points where the transcriptome after the first 50 generations (T1_9C_1-4 blue, T1_9C_5 light blue) was very different from the Tend transcriptome (TE_9_1-4 dark blue, TE_9_5 blue) that was more related to the start of the experiment (T0, green). High temperature transcriptomes (light and dark red) of both sampling time points were very similar as evident by the clustering of all red dots.

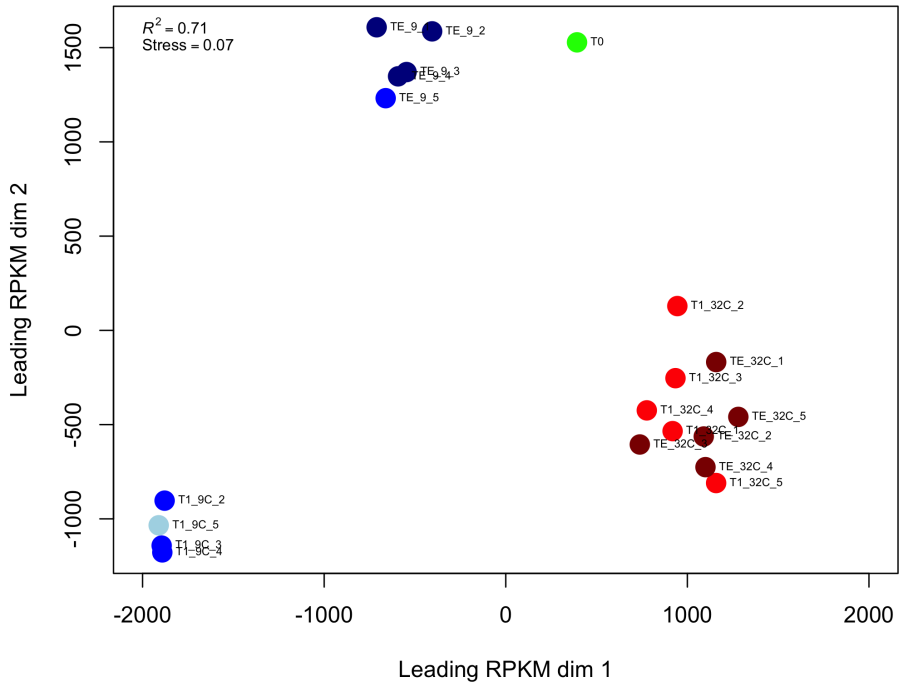

Figure 4.3.1: Multidimensional scaling of temperature related transcriptomes based on log-transformed RPKM values. Manhatten index was used for calculating distances between timepoints and samples. Low temperature replicates are shown in light blue (T1) and dark blue (Tend). High temperature replicates are shown in light (T1) and dark red (Tend). T0 is shown in green.

Differential gene expression (log_2-fold change) between the low temperature selection lines compared to the control, and at two time points (T0vsT1 and T0vsTE) was visualized with a heatmap and hierarchical clustering of genes (Figure 4.3.2). Gene expression pattern showed clear differences after 50 (T0vsT1) and 300 (T0vsTE) generations of low temperature selection. Over the first 50 generations 3877 genes were differentially expressed while this was reduced by two-thirds after 300 generations.

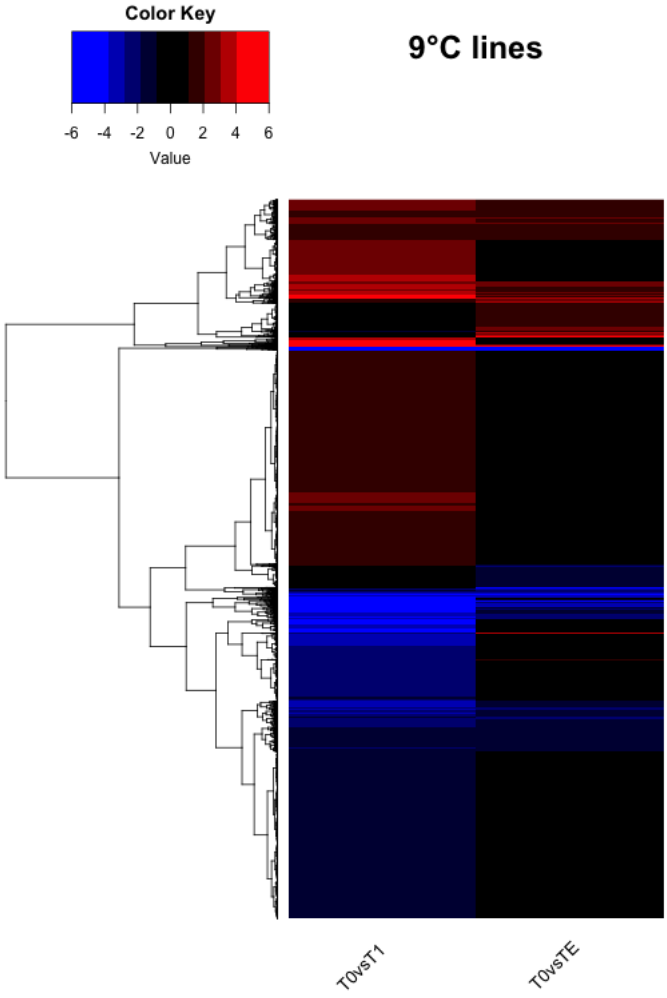

Figure 4.3.2: Transcriptome profile heatmap of low temperature selection lines after 50 (T1) and 300 (TE) generations in relation to the starting culture. Genes that were up-regulated are colour coded from 0 to 6 log_2-fold change with dark to light red while down-regulated genes are colour coded from 0 to -6 log_2-fold change with dark blue to light blue. Genes were clustered hierarchical.

In contrast, differential gene expression pattern (log_2-fold change) of the high temperature selection lines compared to the control, and at two time points (T0vsT1 and T0vsTE) was very

similar (Figure 4.3.3).

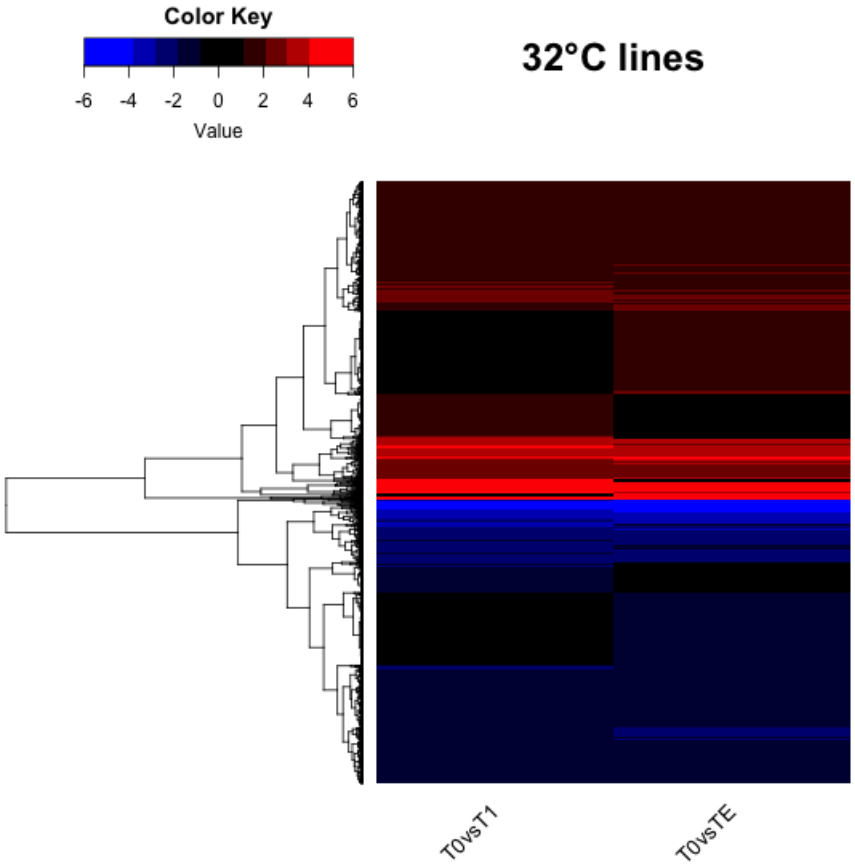

Figure 4.3.3: Transcriptome profile heatmap of high temperature cultures after 50 (T1) and 300 (TE) generations in relation to the starting culture. Genes that were up-regulated are colour coded from 0 to 6 $\log_2$-fold change with dark to light red while down-regulated genes are colour coded from 0 to -6 $\log_2$-fold change with dark blue to light blue. Genes were clustered hierarchical.

4.3.2.1 GO analysis

Differentially expressed genes of low and high temperature selection cultures were functionally analysed for significantly overrepresented gene ontology (GO) term annotations ($\log_2 > 1$, BH adjusted $p < 0.05$). Lists of overrepresented GO-terms were visualised using semantic similarity scatter plots (ReVIGO scatter plots) by which semantically related GO-terms were plotted closer together. Furthermore, GO-terms are represented as bubbles and their sizes relates to the frequency of the GO-term in the Gene Ontology Annotation database, which means that larger

bubbles correspond to general or more frequent GO-terms than more specific GO-terms as smaller bubbles.

Figure 4.3.4 shows GO-term annotations that were overrepresented in the set of up- (a) and down-regulated (b) genes for the GO-term categories biological process of low temperature cultures after 50 generations. Overrepresented GO-terms of biological functions in up-regulated genes were porphyrin-containing compound biosynthetic process (GO:0006779), phosphorylation (GO:0016310), RNA processing (GO:0006396) as well as nucleoside metabolic process (GO:0009116). Low temperature selection had a profound impact on cellular metabolism as represented by down-regulated GO-terms including carbon related meta- and catabolic processes (chitin metabolic process (GO:0006030), chitin catabolic process (GO:0006032), carbohydrate metabolic process (GO:0005975), cell wall macromolecule catabolic process (GO:0016998), oxygen and reactive oxygen species metabolic process (GO:0006800)) as well as response to other organism (GO:0051707), regulation of cell cycle (GO:0051726), proteolysis (GO:0006508) and regulation of transcription, DNA-templated (GO:0006355).

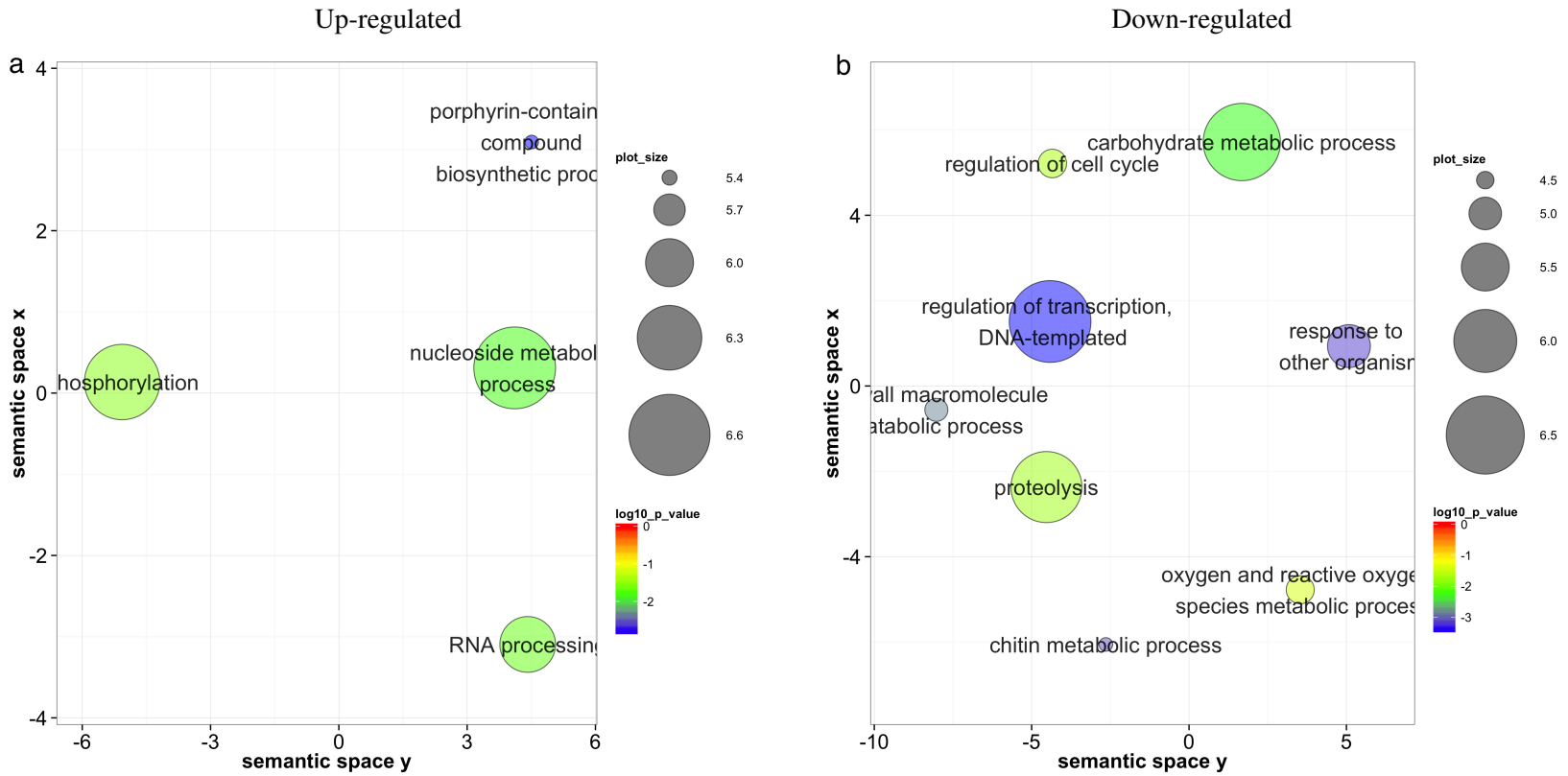

Figure 4.3.4: ReVIGO GO-term scatterplots for low temperature cultures after 50 generations of temperature selection for biological processes. Overrepresented GO-terms (BH adjusted p-value <0.05) of up-regulated genes are shown on the left, down-regulated genes on the right.

After 300 generations of low temperature selection overrepresented GO-terms of up-regulated genes (Figure 4.3.5) were dominated by photobiological related GO-terms (cytochrome complex assembly (GO:0017004), signal transduction (GO:0007165), photosystem II stabilization (GO:0042549)) but also included mRNA metabolic process (GO:0016071), unsaturated fatty acid biosynthesis process (GO:0006636) as well as nucleobase-containing compound metabolic process (GO:0006139). Down-regulated enriched GO-terms involved a majority of protein cycle processes (protein secretion (GO:0009306), protein metabolic process (GO:0019538), proteolysis (GO:0006508) and regulation of transcription, DNA-templated (GO:0006355)). Furthermore, cell-matrix adhesion (GO:0007160), thiamine biosynthetic process (GO:0009228), mismatch repair (GO:0006298) and cell adhesion (GO:0007155) were also found to be down-regulated GO-terms after 300 generations of high temperature selection (Figure 4.3.5).

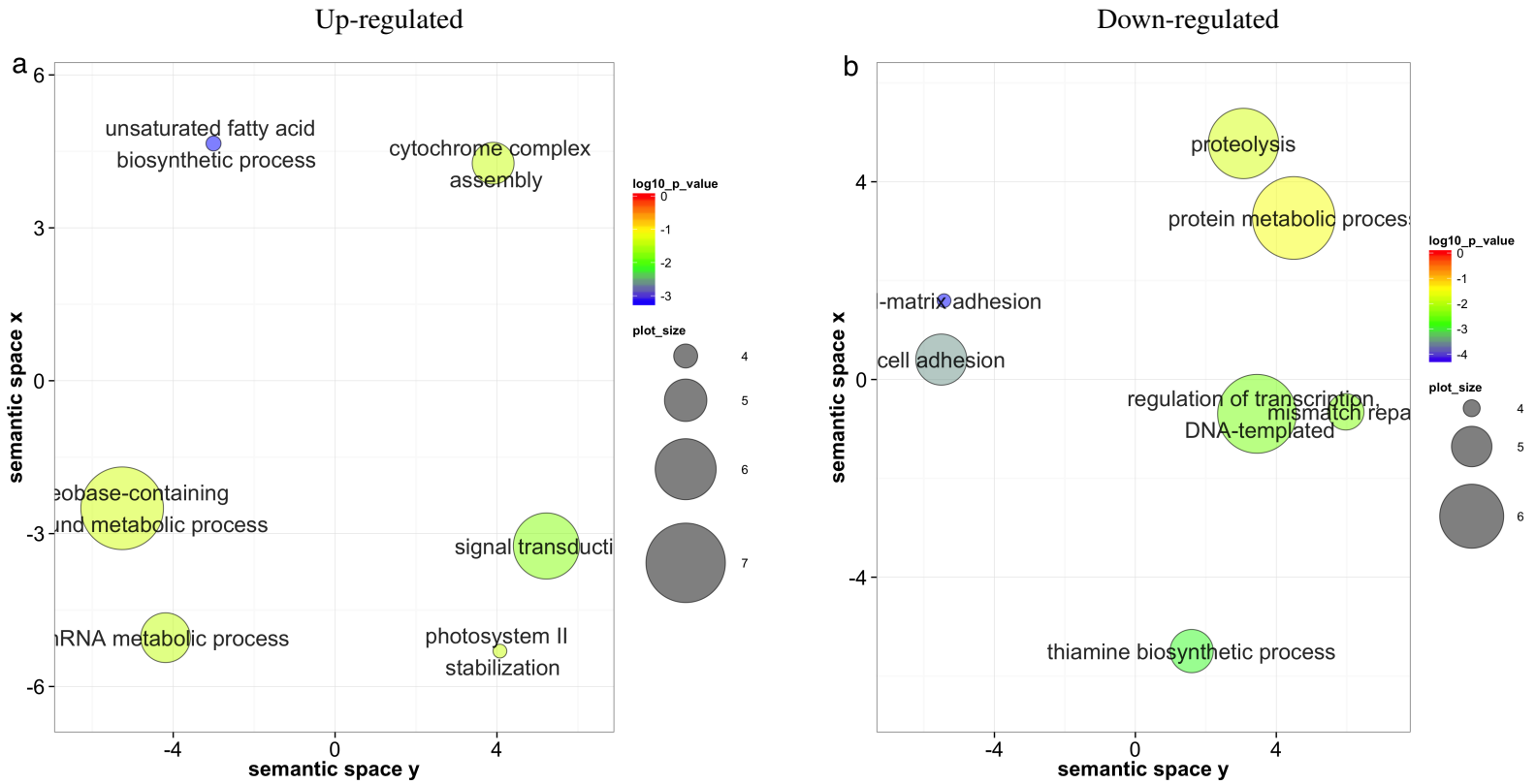

Figure 4.3.5: REVIGO GO-term scatterplots for low temperature cultures after 300 generations of temperature selection. Overrepresented GO-terms (BH adjusted p-value <0.05) of up-regulated genes are shown on the left, down-regulated genes on the right.

High temperature selection, over 300 generations, resulted in the overrepresentation of GO-terms related to biosynthetic processes (Figure 4.3.6, a) which included biosynthetic process (GO:0009058), phospholipid biosynthetic process (GO:0008654), phytochromobilin biosynthetic process (GO:0010024). Furthermore, photosynthesis (GO:0015979) and one-carbon metabolic process (GO:0006730) were identified as significantly upregulated GO-terms (Figure 4.3.6).
Downregulated enriched GO-terms were identified as translation (GO:0006412), response to other organism (GO:0051707), regulation of cell cycle (GO:0051726), ER to Golgi vesicle-mediated transport (GO:0006888), carbohydrate metabolic process (GO:0005975), chitin metabolic process (GO:0006030), chitin catabolic process (GO:0006032), cell wall macromolecule catabolic process (GO:0016998), mRNA processing (GO:0006397) as well as mismatch repair (GO:0006298) and regulation of transcription, DNA-templated (GO:0006355).

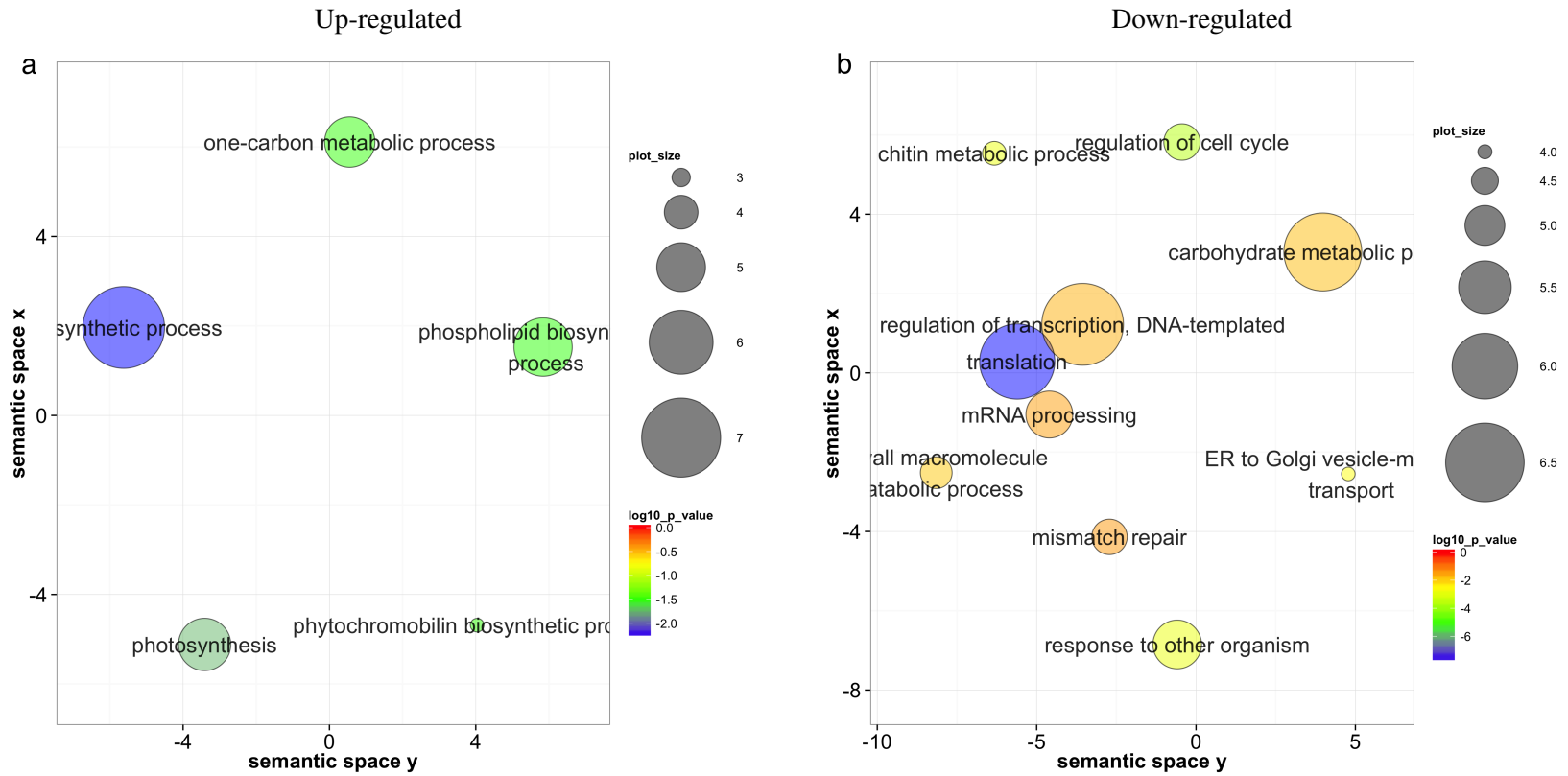

Figure 4.3.6: REVIGO GO-term scatterplots for high temperature cultures after 300 generations of temperature selection. Overrepresented GO-terms (BH adjusted p-value <0.05) of up-regulated genes are shown on the left, down-regulated genes on the right.

4.3.2.2 InterPro Domain analysis - Overrepresented domains

Differentially expressed genes of low and high temperature selection cultures were analysed for significantly overrepresented InterPro domains (BH adjusted $p<0.05$).

Overrepresented InterPro annotations of down-regulated genes after 50 generations of low temperature selection in comparison to control conditions (T0, 22°C) are shown in Table 6.4.1. Highly overrepresented domains belonged to heat shock factors as well as chinases and chitin binding domains. Additionally, a cyclin domain was significantly overrepresented in down-regulated genes. The full list of overrepresented InterPro domains can be found in the appendix (Table).

Table 4.3.3: Overrepresented InterPro domains of down-regulated genes at T1 9°C. P-values were adjusted after Benjamin-Hochberg with a cut-off of 0.05

InterPro ID	**Overrepresented p-value**	**InterPro description**	**Domain description**
IPR002341	8.03E-09	HSF/ETS, DNA-binding	HSF ETS
IPR000232	1.13E-08	Heat shock factor (HSF)-type, DNA-binding	sp HSF1 ARATH P41151
IPR011583	0.001	Chitinase II	Glyco 18
IPR006671	0.015	Cyclin, N-terminal	Cyclin N

Table 6.4.2 provides an overview of significantly overrepresented InterPro domains of up-regulated genes after 50 generations of low temperature selection. The most significantly overrepresented term was related to the twin-arginine translocation pathway. Additionally, several terms related to translation were identified as overrepresented in up-regulated genes such as elongation factor Tu (IPR004161), Phosphoribosyltransferase (IPR000836) and Aminoacyl-tRNA synthetase (IPR002300). The full list of overrepresented InterPro domains can be found in the appendix (Table).

Table 4.3.4: Overrepresented InterPro domains of up-regulated genes after 50 generations at 9°C. P-values were adjusted after Benjamin-Hochberg with a cut-off of 0.05.

InterPro ID	**Overrepresented p-value**	**InterPro description**	**Domain description**
IPR006311	0.003	Twin-arginine translocation pathway signal	TAT signal seq
IPR004161	0.008	Elongation factor Tu, domain 2	GTP EFTU D2
IPR000836	0.014	Phosphoribosyltransferase	Pribosyltran
IPR002300	0.014	Aminoacyl-tRNA synthetase, class Ia	tRNA-synt 1

After 300 generations of low temperature selection, overrepresented InterPro domains of down-regulated genes showed similarities with domains after 50 generations (Table 6.4.3). Several heat shock factors (IPR000232, IPR002341) and proteins (IPR002068, IPR001023) as well as a chaparone (IPR002939) were identified as overrepresented in down-regulated genes. Additionally, a domain related to the MutS2 homologue (Smr protein/MutS2 C-terminal, IPR002625) of the mismatch repair pathway was included in overrepresented down-regulated genes. The full list of overrepresented InterPro domains can be found in the appendix (Table).

Table 4.3.5: Overrepresented InterPro domains of down-regulated genes at TE 9°C. P-values were adjusted after Benjamin-Hochberg with a cut-off of 0.05.

InterPro ID	**Overrepresented p-value**	**InterPro description**	**Domain description**
IPR000232	2.93E-06	Heat shock factor (HSF)-type, DNA-binding	sp HSF1 ARATH P41151
IPR002341	2.93E-06	HSF/ETS, DNA-binding	HSF ETS
IPR002068	8.88E-05	Heat shock protein Hsp20	HSP20
IPR001023	0.040	Heat shock protein Hsp70	sp DNAK ODOSI P49463
IPR002939	0.032	Chaperone DnaJ, C-terminal	DnaJ C
IPR002625	0.004	Smr protein/ MutS2 C-terminal	Smr

Cultures selected at 9°C for 300 generations showed highly significant overrepresented fatty acid desaturases (IPR005804, IPR010257). Additional overrepresented InterPro IDs corresponded to photosystem related proteins (Cytochrome b5, Photosystem II 12 kDa extrinsic and Cytochrome c biogenesis protein) and light sensing (Phytochrome). The full list of overrepresented InterPro domains can be found in the appendix (Table).

Table 4.3.6: Overrepresented InterPro domains of up-regulated genes at TE 9°C. P-values were adjusted after Benjamin-Hochberg with a cut-off of 0.05.

InterPro ID	Overrepresented p-value	InterPro description	Domain description
IPR005804	3.35E-04	Fatty acid desaturase	FA desaturase
IPR010257	7.32E-04	Fatty acid desaturase subdomain	sp Q84KI8 PHATR Q84KI8
IPR001199	0.012	Cytochrome b5	Cyt-b5
IPR010527	0.039	Photosystem II 12 kDa extrinsic	PsbU
IPR003834	0.041	Cytochrome c biogenesis protein, transmembrane region	DsbD

4.3.2.3 Pathway analysis

Genes for biosynthesis of saturated fatty acids showed increased expression under low temperature conditions after 50 generations and under high temperature conditions at both 50 and 300 generations in comparison to T0 (22°C) (Figure 4.3.7). The activation of malonyl CoA for reaction with acetyl-ACP by Malonyl CoA:ACP transacylase (MCT1, 5219) showed increased expression under T1 9°C replicates 1-4 and 5 as well as TE 9°C replicate 5 and TE 32°C. The produced malonyl-ACP is then used to extend the fatty acid chain by priming with acetyl-ACP via beta-ketoacyl-ACP synthase (KAS2, 262879). This step showed highest expression at both time points for 32°C cultures. Once a 16:0 fatty acid (Hexadecanoic acid) is formed, several long-chain-fatty-acid CoA ligases facilitate the reaction to hexadecanoyl CoA which can then be used in the glycerolipid and glycersophospholipid metabolism, or for fatty acid elongation. This final step of fatty acid synthesis showed high expression at 9°C replicates 1-4 and 5 (ACS1, 13156) at both time points, as well as 32°C (ACS1 13156; ACS2_1, 29867; FCL1, 263105) at both time points in comparison to the control (T0).

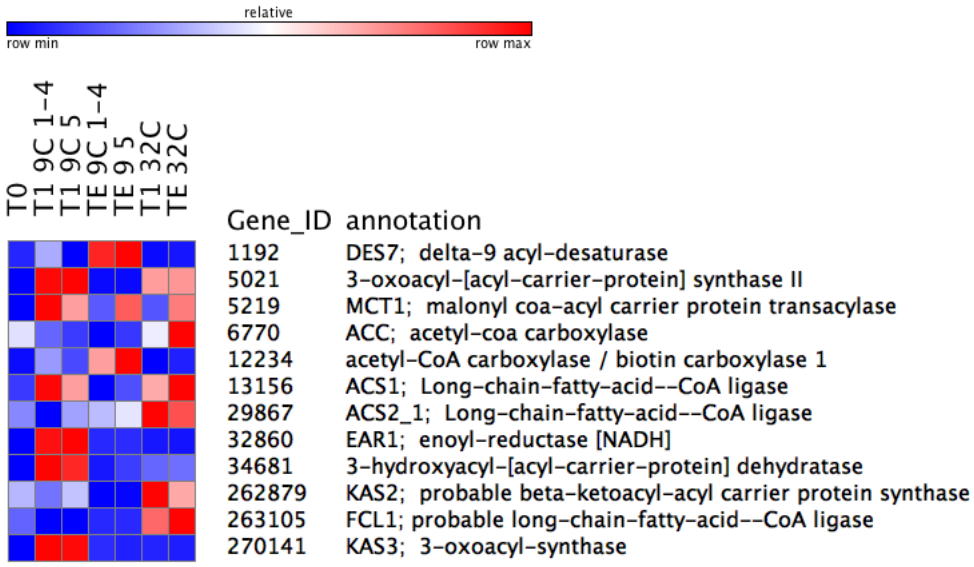

Figure 4.3.7: Expression of genes related to fatty acid biosynthesis in Tp during temperature selection. Heatmap is scaled to RPKM values and color coded by row minimum (blue) and maximum (red).

Desaturation of fatty acids is facilitated by desaturases. Expression of genes related to this pathway is shown in Figure 4.3.7. Two delta-9 desaturases (1527, 22511) showed high RPKM values under both high temperature (32°C) time points. Additionally, long chain fatty acid modifiers (reductase, dehydrogenase and thiolase) showed expression under low (33911, 22650) and high temperature (22650) conditions. Beta-oxidation of fatty acids facilitated by a 3-ketoacyl-CoA thiolase (KCT2, 36742), acyl-CoA oxidase (AOX1_1, 263878) and enoyl-CoA hydratase (26365) showed high RPKM values under low and high temperature conditions for each time point in comparison to T0.

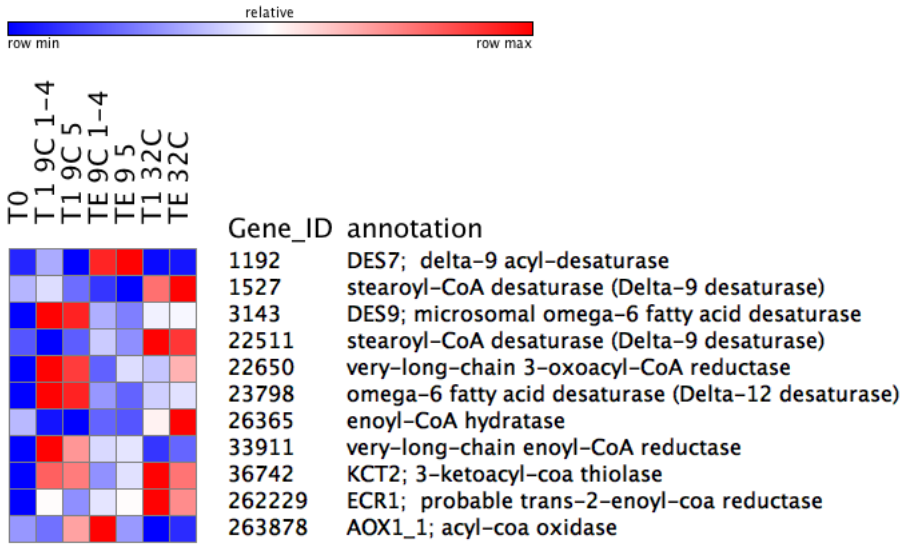

Figure 4.3.8: Heatmap of expressed genes involved in desaturation fatty acids. Heatmap is scaled to RPKM values and color coded by row minimum (blue) and maximum (red).

Genes involved in carotenoid synthesis in Tp were taken from Coesel (2008) and their expression in terms of RPKM values is shown in Figure 4.3.9. Seven out of eight identified genes showed high RPKM values after 50 generations of low temperature selection in replicates 1-4 and 5 in comparison to the control (T0). After 300 generations at 9°C RPKM values were reduced. High temperature cultures showed high expression of genes involved in the synthesis of xanthophylls (zeaxanthin, antheraxanthin and violaxanthin). Beta-carotene hydroxylase (BCH, 263437) introduces hydroxyl groups into beta-carotene forming zeaxanthin and showed highest expression under both high temperature time points. Further epoxidation (ZEP1, 269147; ZEP2, 261390) of zeaxanthin into antheraxanthin into violaxanthin showed moderate expression under high temperature in comparison to the control (T0).

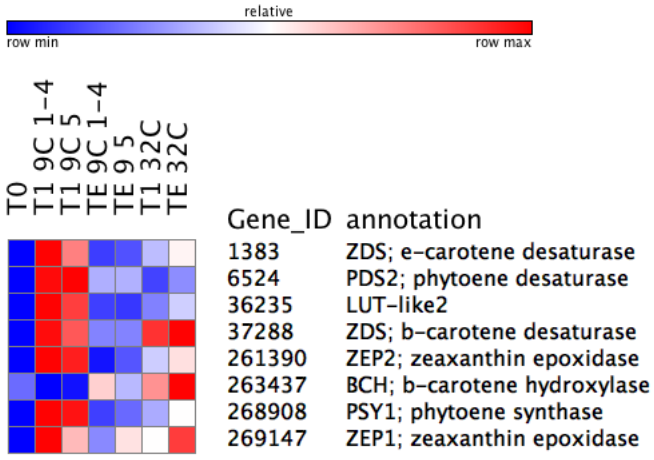

Figure 4.3.9: Heatmap of expressed genes part of the carotenoid synthesis pathway. Heatmap is scaled to RPKM values and color coded by row minimum (blue) and maximum (red).

Expression of genes involved in the porphyrin and chlorophyll metabolism are presented in Figure 4.3.10. Most identified genes showed highest expression under low temperature conditions within in the first 50 generations (T1 9C 1-4 and 5) and moderate expression after 300 generations (TE 9C 1-4 and 5). High temperature cultures showed increased expression after 300 generations (TE 32C). Chlorophyll(ide) b reductase (8063) is one of the last steps in chlorophyll a synthesis and showed high expression after 50 generations of low temperature selection and moderate expression after 300 generations of high temperature selection.

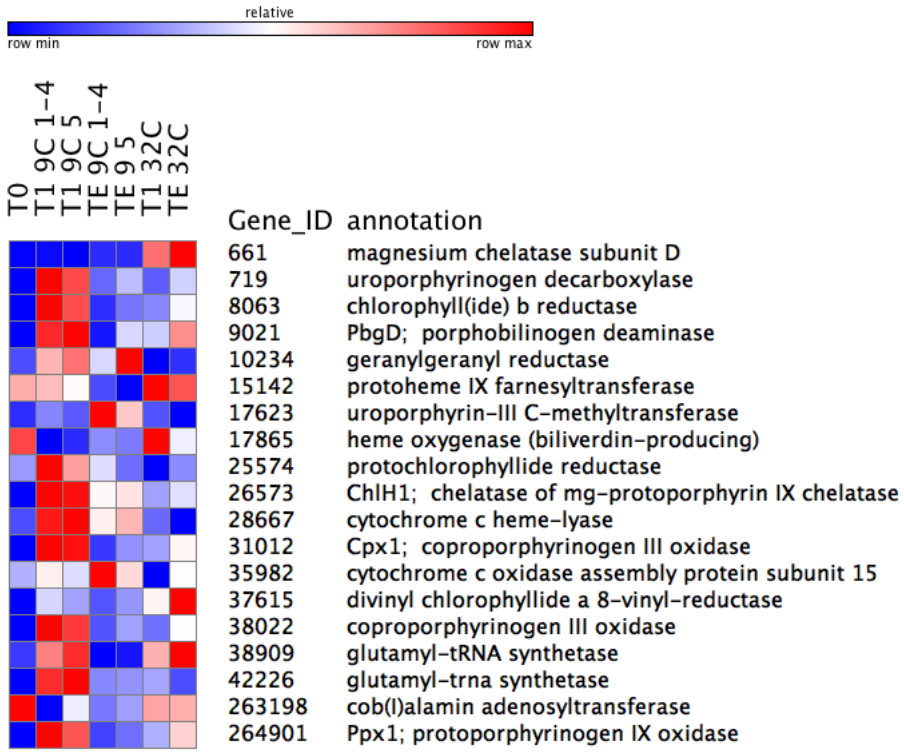

Figure 4.3.10: Heatmap of expressed genes part of porphyrin and chlorophyll metabolism. Heatmap is scaled to RPKM values and color coded by row minimum (blue) and maximum (red).

Figure 4.3.11 shows the expression of ribosomal protein genes in terms of RPKM values for control cultures (T0), low temperature conditions (9°C) and high temperature conditions (32°C). Almost all ribosomal protein genes showed high expression within the first 50 generations (T1 9C 1-4 and 5) and moderate expression after 300 generations (TE 9C 1-4 and 5) under low temperature selection. In contrast, high temperature cultures showed minimal expression of ribosomal protein genes at both time points (T1 32C and TE 32C).

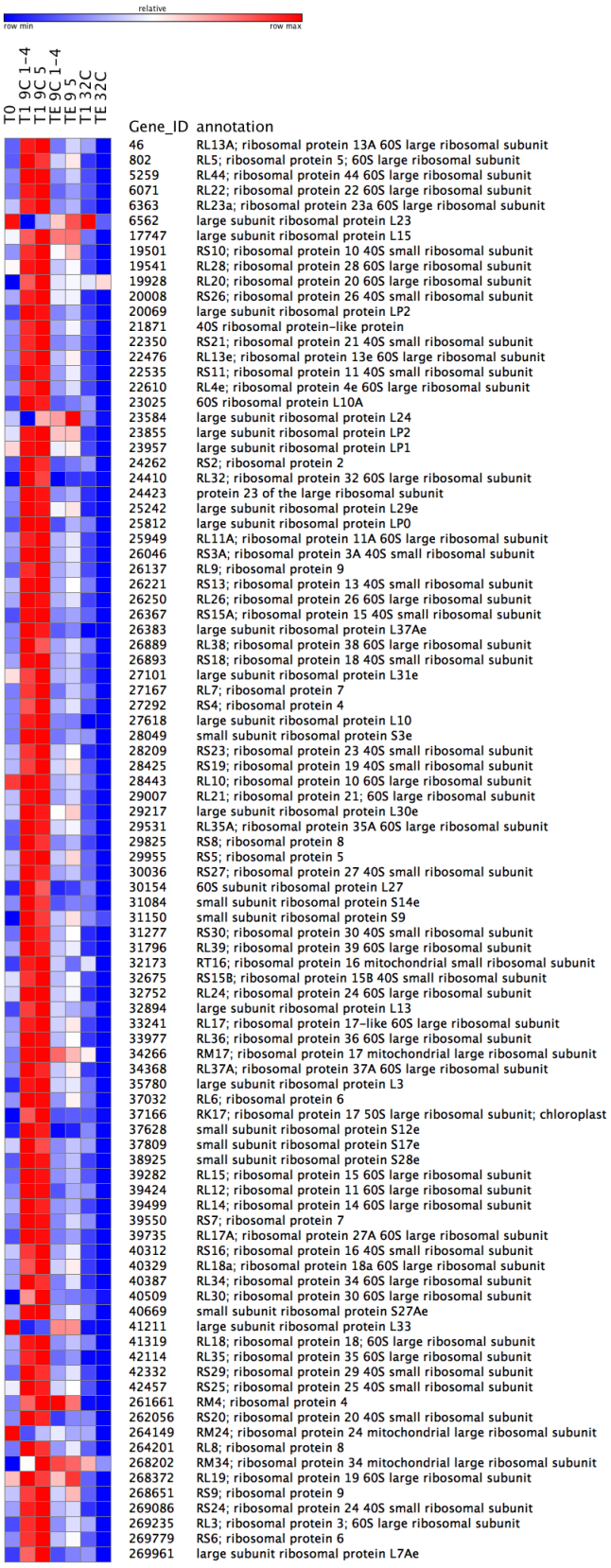

Figure 4.3.11: Heatmap of expressed ribosomal genes. Heatmap is scaled to RPKM values and color coded by row minimum (blue) and maximum (red).

Genes coding for proteins and enzymes involved in the mismatch repair showed high expression under low temperature conditions within the first 50 generations (Figure 4.3.12) based on RPKM values. After 300 generations, expression seemed to be reduced except for two mismatch repair proteins (MLH1, 263509 and PMS2, 264783). High temperature cultures showed only high expression of the mismatch repair protein PMS2 (264783) after 50 (T1 32C) and 300 (TE 32C) generations.

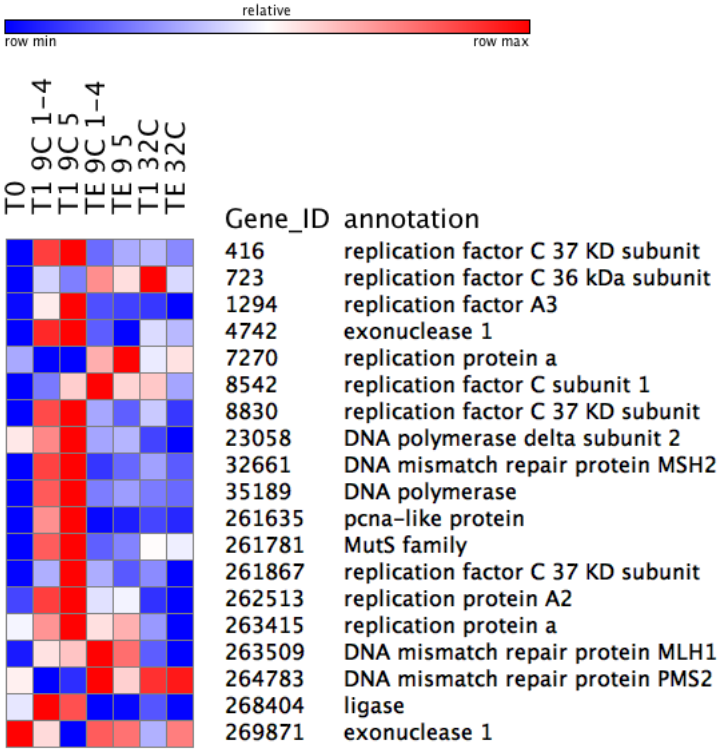

Figure 4.3.12: Heatmap of expressed genes part of the mismatch repair pathway. Heatmap is scaled to RPKM values and color coded by row minimum (blue) and maximum (red).

4.3.2.4 Transposable elements (TEs)

To investigate potential genetic changes that might have occurred under temperature selection, the activity of known transposable elements in *Thalassiosira pseudonana* was investigated. Furthermore, expression of enzymes relating to DNA-(de-)methyltransferases was examined.

To date a total of 77 TEs have been identified in the Tp genome (Maumus *et al.*, 2009), out of which six were expressed in the first 50 generations of low temperature selection and five after 300 generations (Table **??**). The identified TEs active under low temperature selection, belonged

to Copia LTR retrotransposons and one to LINE transposons. Temperature specific regulation of several TEs was found (bold entries) and only differentially expressed under low temperature selection.

Table 4.3.7: Differentially expressed TEs (BH adjusted p-value <0.05, $\log_2$-fold change) in low, high and control selection lines. Bold entries were temperature specific.

Transposon	Protein ID	9°C T1	9°C Tend	32°C T1	32°C Tend	control T1	control Tend
LTR RT (Copia)	7301	-4.61	-4.61	-3.35	0	0	0
LTR RT (Copia)	8235	2.25	3.52	-3.40	-3.29	0	0
LTR RT (Copia)	10585	1.67	0	3.35	3.16	0	0
LTR RT (Copia)	11724	0	0	6.05	0	0	0
LINE (Ambal)	11865	0	0	**4.95**	**4.77**	0	0
LTR RT (Copia)	12122	0	**2.39**	0	0	0	0
LTR RT (Copia)	12209	**-5.66**	**-5.20**	0	0	2.11	0
LINE (Ambal)	25567	**-2.87**	0	0	0	0	0
LTR RT (Copia)	25771	**1.40**	**2.49**	0	0	0	0
LTR RT (Gypsy)	263201	0	0	**10.92**	**12.09**	0	0
LTR RT (Copia)	263838	0	0	**5.12**	0	0	0
LTR RT (Copia)	264303	0	0	**5.43**	**6.72**	0	0

Six DNA specific (de-)methyltransferases are annotated in the Tp genome, out of which five were differentially expressed in the selection cultures (Table 4.3.8). However, they did not show a temperature specific expression pattern.

Table 4.3.8: Differentially expressed DNA-specific methyltransferases (BH adjusted p-value <0.05, $\log_2$-fold change) in low, high and control temperature selection lines and at two time points (T1, 50 generations and Tend, 300 generations).

Methyltransferase	Protein ID	9°C T1	9°C Tend	32°C T1	32°C Tend	control T1	control Tend
cytosine specific DNA methyltransferases	2049	-3.24	0	-1.75	-2.02	-1.60	-1.02
cytosine specific DNA methyltransferases	11011	0	0	0	0	0	-1.34
DNA methyltransferases	21517	0	0	-1.15	-1.43	0	0
cytosine specific DNA methyltransferases	22139	1.69	0	0	0	0	0
adenine specific DNA methyltransferases	34058	0	0	0	0	-1.02	0
cytosine specific DNA methyltransferases	269660	0	0	0	0	0	-1.04

4.3.2.5 Genome resequencing - SNP analysis of high temperature replicate

SNP analysis revealed a total of 6,374 mutations over the whole genome (Table 4.3.9) after 300 generations of high temperature selection. This calculates to a mutation rate of 7.98 x 10^{-7} bp generation^{-1}. Notably, chromosome 14 showed the highest number of mutations (1,690 or 1 every 590 bp) followed by chromosome 18 with 837 mutations (1 every 988 bp). 66% of genes on chromosome 14 showed mutations with on average 5.2 SNPs per gene. Additionally, of the top 229 genes with more than 4 mutations, 39.7% belonged to chromosome 14. In contrast, genes on chromosome 18 had on average 10.6 mutations with the top genes making up 11.4% of genes with more than 4 mutations.

Table 4.3.9: Summary of variants generated by SNPs per chromosome and variant rate (SNP per base).

Chromosome	Length (bp)	Variants	Variants rate (SNP per base)
1	3,042,585	237	12,837
2	2,707,195	130	20,824
3	2,440,052	182	13,406
4	2,402,323	223	10,772
5	2,305,972	386	5,974
6	2,071,480	202	10,254
7	1,992,434	157	12,690
8	1,267,198	314	4,035
9	1,191,060	94	12,670
10	1,105,668	176	6,282
11a	806,142	91	8,858
11b	82,843	42	1,972
12	1,128,382	198	5,698
13	1,052,196	148	7,109
14	**998,643**	**1,690**	**590**
15	931,268	152	6,126
16a	459,776	142	3,237
16b	169,376	32	5,293
17	659,924	98	6,733
18	**827,053**	**837**	**988**
19a_19	607,239	84	7,229
19b_31	151,677	42	3,611
19c_29	291,194	338	861
20	800,234	83	9,641
22	1,057,565	91	11,621
23	454,954	73	6,232
24	297,349	132	2,252
Total	31,301,782	**6,374**	**4,910**

SPNs were grouped by region and are summarized in Figure **??**. Interestingly, intergenic and exons showed similar amounts of SNPs (40.68% and 35.68% respectively). In contrast, up- and downstream, 5'UTR and 3'UTR, as well as introns and splice variants received between 0.09% and 5% of SNPs.

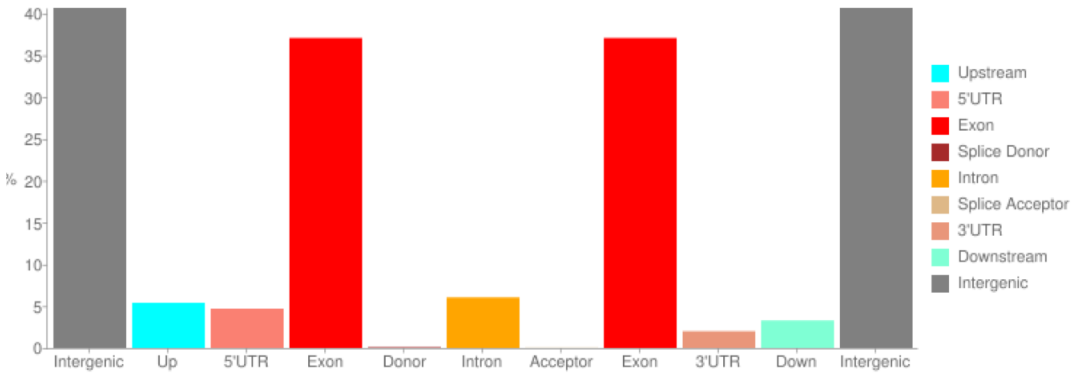

Figure 4.3.13: Percentage of SNPs per genomic region.

SNPs occurring in exons resulted in different effects (Table 4.3.10). 1,534 (58.31%) mutations caused missense amino acids (AA) while 47 (1.79%) caused a new stop codon. In contrast 1,050 (39.91%) caused no changes in AAs (silent mutation).

Table 4.3.10: Overall occurrence of missense, nonsense and silent SNPs.

Type	Count	Percent
MISSENSE	1,534	58.31%
NONSENSE	47	1.79%
SILENT	1,050	39.91%

GO analysis of genes with SNPs The identified genes with SNPs were used for GO-term analysis related to biological processes and are shown in Figure 4.3.14. Most significant GO-terms related to nucleosome assembly and included chromosome organization, regulation of transcription and transcription initiation. Additionally, genes related to the cell cycle and intracellular protein transport were overrepresented.

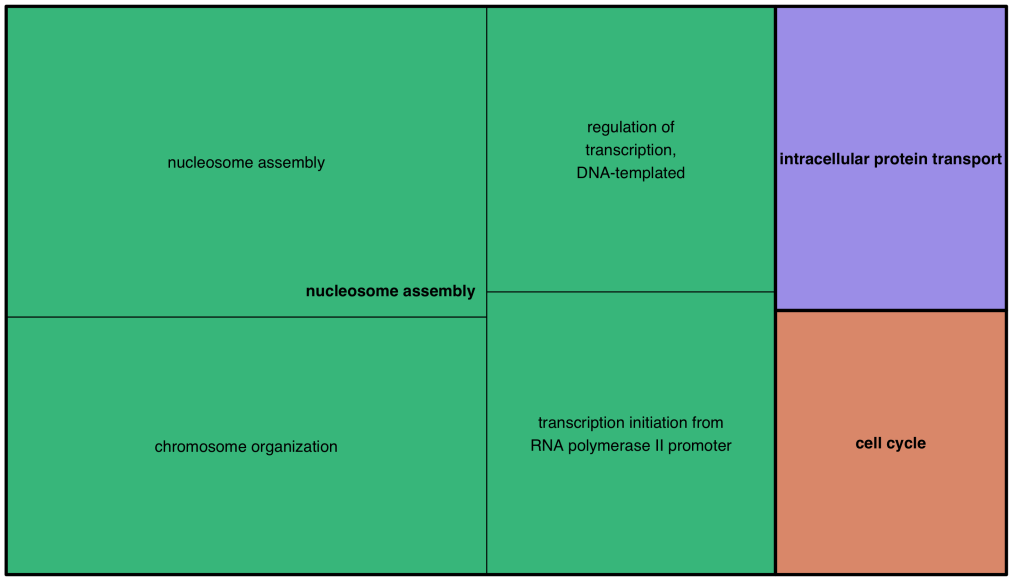

Figure 4.3.14: Treemap of overrepresented (BH adjusted p-value<0.05) GO-terms for all genes with SNPs.

Genes with mutations on chromosome 14 and 18 were further used for GO-term analysis of biological processes. Chromosome 18 showed one overrepresented GO-term related to translational termination (GO:00006415). Most significant GO-terms of genes on chromosome 14 related to chromosome organization and DNA replication initiation (Figure 4.3.15).

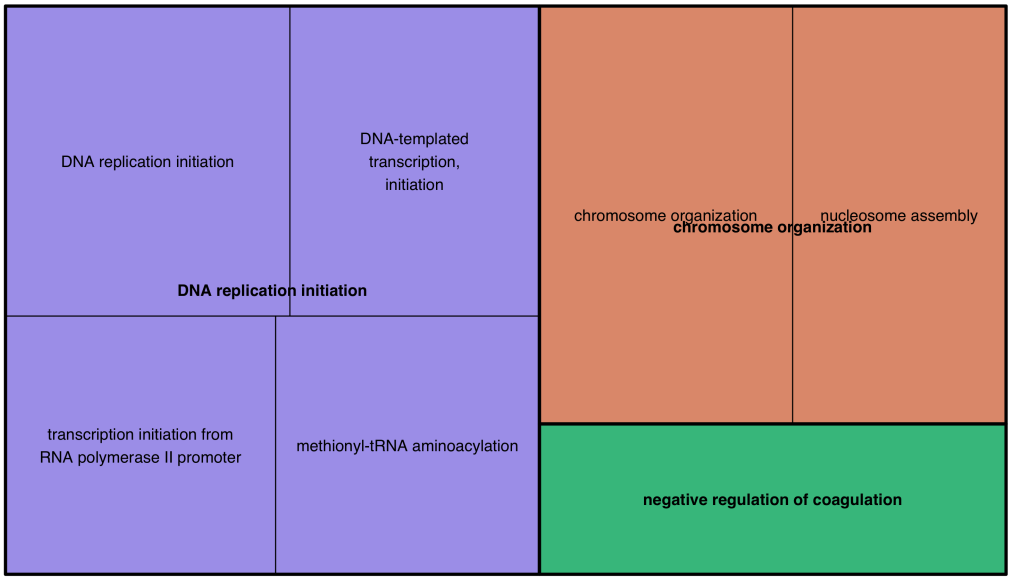

Figure 4.3.15: Treemap of overrepresented (BH adjusted p-value<0.05) GO-terms for all genes with SNPs.

Amino acid (AA) changes A total of 1,582 AA changes were caused by missense mutations. Out of the 20 known AA, Arginine, Serine and Valine were changed most often (121, 149 and 108 times, respectively). Furthermore, AA were most often changed into Arginine, Serine and Threonine (147, 149 and 125 times, respectively). Amino acid changes caused 109 charge changes from negative to positive or neutral and 182 times from positive to negative or neutral.

	*	?	A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y
*	4			1		2		1			2	3				10	13	1			6	3
?		1																				
A			102		5	4		4							2			20	41	27		
C				25			4	2									9	6			4	8
D			5		40	15		16	4					24						8		7
E			8		25	35		14			16					2				5		
F				4			15			5		13						11		4		4
G	4		7	4	16	10		93									21	13		5	1	
H					4				17			2		5	3	8	15					18
I							5			43	4	11	7	5			1	1	21	17		
K	2					17				2	37		4	18		9	18		15			
L	4						14		3	14		123	4		13	4	11	15		6	5	
M										8	7	12					4		13	3		
N					18				4	10	17			33				16	10			3
P			2						3			20			73	6	8	25	5			
Q	8					6			16		15	4			9	28	19					
R	14			8				17	11	3	6	9	2		3	22	67	14	1		11	
S	6		14	10			10	12		7		16		17	20		12	121	19		1	5
T			32							23	10		7	6	3		4	17	105			
V			37		6	3	4	10		28		10	10							65		
W	7			2				2				1					12	1				
Y	2			12	6		9		8					5				9				23

Figure 4.3.16: Amino acid changes in the high temperature replicate caused by the identified SNPs. Changes are colour coded, with grey representing synonymous changes, colours of green show missense changes. Amino acid code: G - Glycine (Gly), P - Proline (Pro), A - Alanine (Ala), V - Valine (Val), L - Leucine (Leu), I - Isoleucine (Ile), M - Methionine (Met), C - Cysteine (Cys), F - Phenylalanine (Phe), Y - Tyrosine (Tyr), W - Tryptophan (Trp), H - Histidine (His), K - Lysine (Lys), R - Arginine (Arg), Q - Glutamine (Gln), N - Asparagine (Asn), E - Glutamic Acid (Glu), D - Aspartic Acid (Asp), S - Serine (Ser), T - Threonine (Thr), $*$ - Stop codon, ? - unknown

4.4 Discussion

In order to identify genetic processes that enable evolution of thermal tolerance in phytoplankton, we analysed the transcriptome of *Thalassiosira pseudonana* (Tp) after 50 and 300 generations of high and low temperature selection. On average 13.9 million paired reads were generated per sample of which 85.8% mapped to the reference genome and were further used for gene expression analysis. This is a somewhat higher percentage than has been found previously (75%, uniquely mapping and no mismatches) (Dyhrman *et al.*, 2012) and is potentially due to more

relaxed mapping settings in this study (uniquely mapping and allowing for 2 mismatches). Sequenced reads may not map because of single nucleotide polymorphisms (biological or introduced by sequencing error), because of incorrectly assigned intron/exon boundaries or reads mapping to UTRs that were not included in the predicted gene models. About 17% of expressed genes could not be assigned to genomic features and these unassigned features increased under both temperature selections. This suggest, since the unassigned features were induced by temperature, that there are aspects of thermal adaptation yet to be resolved in this diatom.

Patterns of differentially expressed genes varied between the temperature treatments. Tp showed a short-term response in the first 50 generations of temperature selection indicated by 3877 and 2033 differentially expressed genes ($\log_2$-fold change, BH adjusted $p<0.05$), under low and high temperatures respectively. Interestingly, the number of differentially expressed genes decreased by about 75% in the low temperature cultures while it slightly increased for the high temperature cultures. This might suggest in case of the low temperature cultures, they have overcome the acute stress phase and adapted to the low temperature condition. This is further supported by the multidimensional scaling of the transcriptomes (Figure 4.3.1) where the low temperature transcriptome at the end of the experiment (300 generations) is more similar to the control than the transcriptome after 50 generations. A similar response has been shown for *Escherichia coli* (*E. coli*) adapting to high temperature.

Acute environmental stress like temperature changes greatly affect diatom physiology and usually involves a general stress response which includes protection of cellular macromolecules, redistribution of metabolic resources towards stress responses and in the case of severe stress the reversible arrest of the cell cycle and ultimately programmed cell death through apoptosis (Logan and Buckley, 2015). Temperature treatment of Tp showed metabolic stress as GO-term analysis revealed the down-regulation of cell cycle regulatory genes and of transcription (DNA-templated) in both high (Figure 4.3.6) and low temperature (Figure 4.3.4) conditions. This is further supported by overrepresentation of cyclins in down-regulated genes after 50 generations of low temperature selection (Table 6.4.1) and 300 generations of high temperature selection (Table 6.4.5). Cyclins control the progression of the cell cycle by activating cyclin-dependent kinases (Galderisi *et al.*, 2003). Their down-regulation has been linked to cell cycle arrest in heat stressed

Arabidopsis (Kitsios and Doonan, 2011). Thus, their down-regulation might arrest or slow Tp's cell cycle as a response to temperature stress. This might be related to resource redistribution to processes related to temperature adaptation. However, it has to be noted, that down-regulation of genes corresponds even less well to protein abundance in comparison to the correlation between up-regulated gene and protein abundance (Vélez-Bermúdez and Schmidt, 2014).

After 300 generations of low temperature selection, GO-term analysis revealed cellular adaptation to the new condition. Processes related to photosynthesis stabilization were up-regulated which might have been linked to the high light effect of the light conditions under lower temperature (see Chapter 3). This effect has also been shown in Chapter 3, (Figure **??**) where chlorophyll a content was significantly reduced in low temperature cells. Additionally, genes involved in the production of carotenoids showed high expression after 50 generations of low temperature selection (Figure 4.3.9). As discussed in chapter 3, carotenoids act as photoprotective pigments as they can block excess light (**?**) and their conversion into quenching pigments can result in reduced energy absorption by chl a as they are unable to transfer their energy towards chl a (Horton *et al.*, 1999; Niyogi *et al.*, 2005). Thus the up-regulation of carotenoid synthesis might benefit adaptation under low temperature conditions, but it remains speculative as we did not measure other pigments than chl a.

Additionally, functions of up-regulated genes were annotated as the synthesis of unsaturated fatty acids and desaturation of fatty acids, a general mechanism, in Tp and other organisms, of adapting to low temperature (White and Somero, 1982; Thompson *et al.*, 1992*b*). Fatty acids are an important component of membranes and relate to membrane fluidity. A decrease in temperature hardens fatty acids, however, desaturation can increase membrane fluidity. Thus, an increased expression of genes involved in the synthesis of unsaturated fatty acids and desaturation might increase membrane fluidity. Membrane fluidity is especially important for thylakoid membranes in autotrophs as decreased fluidity could cause reduction in photosynthetic efficiency (see chapter 3 for more details). Thus, increasing unsaturated fatty acids in thylakoid membranes might allow recovery of the photosynthetic efficiency (Falcone *et al.*, 2004), as seen in chapter 3 (Figure Fv/Fm) and provides a link between phenotype and gene expression.

High temperature selection cultures seemed to experience chronic thermal stress as their differential gene expression was very similar after 50 and 300 generations. Interestingly, a relatively low number of genes was differentially expressed which was in contrast to other studies. For example, high temperature treatment of *E.coli* (Riehle *et al.*, 2003) and *F. cylindrus* (Strauss et al., unpublished data) resulted in 50% of genes differentially expressed, while for Tp only 25% genes were differentially expressed. The values for *E. coli* and *F. cylindrus* were gained after short-term temperature treatment and therefore might not reflect adaptive responses which could underline the differential expression in this study. Additionally, 50 generations after the initial temperature shock might be too long to find a heat shock related response as this is a rather short-term process. This is supported by findings by Ying et al. (2015) where the transcriptomes of high temperature evolved *E. coli* showed significant differences to short-term heat shock transcriptomes.

Additionally, a study by Sorensen (2007) on *Drosophila* temperature adaptation suggested that a heat shock response might not play a major role in evolutionary temperature adaptation, proposing separate mechanisms for dealing with short- and long-term thermal stress. Furthermore, the ecological temperature niche is likely to have an impact on temperature related stress responses. However, a 32°selection temperature should have posed a high enough selection pressure, as this corresponds to the highest seasonal temperature Tp experiences in Chesapeake Bay (Edgewater, MD, USA), where it regularly forms dense blooms (Sobrino and Neale, 2007; Marshall *et al.*, 2005). A study on sea grass comparing a northern and southern European habitat showed that HSP induction was absent but other stress related genes showed expression (Franssen *et al.*, 2014).

Interestingly, down-regulated genes involved in the mismatch repair were significantly overrepresented under both selection temperatures. In *E. coli* it has been shown that mismatch repair was down-regulated under nutrient limited conditions leading to a mutator phenotype. This resulted in a rapid diversification of *E. coli* populations (Harris *et al.*, 1997; Saint-Ruf and Matic, 2006). Thus, down-regulation of mismatch repair mechanisms might suggest a molecular

mechanism to produce genetic variation, by introducing single nucleotide polymorphisms (SNPs), on which selection can act (Saint-Ruf and Matic, 2006).

Another mechanism to generate potentially adaptive genotypes is the expression of transposable elements (TEs). TE are known to respond to environmental stimuli and have been shown to respond to nitrate starvation in *Phaeodactylum tricornutum* (Maumus *et al.*, 2011). Additionally, a "cut and paste" transposon has been identified to cause genetic fixation of a B12-dependent phenotype in *C. reinhardtii* by causing a insertion that resulted in a premature stop codon (Helliwell *et al.*, 2015). Thus, TEs can, relatively quickly, cause genetic variation but also regulate gene expression depending on where they are inserted. All but two of the TEs identified here belong to the class of **l**ong-**t**erminal-**r**epeat **r**etro**t**ransposons (LTR RT) which are spread in a "copy and paste" manner via an RNA intermediate and have all necessary genes for transposition on board. Interestingly, six of the differentially expressed TEs seemed to be temperature responsive and specific. One TE in *Arabidopsis* has been shown to include promotors, related to heat shock transcription factors, which enhances the transcription of neighbouring genes under temperature stress (Cavrak *et al.*, 2014). Thus, TEs could play an important role in regulating stress-sensitive responses in diatoms. The highly up-regulated ($\log_2$ 10.92) TE in the high temperature treatment might suggest a similar activation mechanism. However, to induce transposition of a transposon, high amounts of GAG protein compared to LTR TE DNA is required, 1400:1 in tobacco (Takeda *et al.*, 2001), which might have been achieved by this high differential expression. Further analysis would be needed to identify transposition events in the selection lines and their impact on neighbouring genes.

Genome resequencing, after 300 generations, of one high temperature replicate identified over 6,000 single nucleotide polymorphisms. This resulted in a 10-100x higher mutation rate (10^{-7}) in comparison to *C. reinhardtii* (Perrineau *et al.*, 2014) and *E. coli* (Blount *et al.*, 2008), and might be linked to the down-regulation of mismatch repair mechanisms (Harris *et al.*, 1997). Additionally, mutation occurrence showed a bias towards chromosome 14. This is in contrast to *C. reinhardtii* where no such mutation heterogeneity across chromosomes was found (Ness *et al.*, 2015). It is currently unclear whether these mutations were neutral or provided a fitness

advantage.

Out of the 6,000 identified mutations, 1,582 caused amino acid substitutions potentially altering protein stability and structural flexibility (Somero, 2004). Especially the substitution of serine with alanine/threonine or lysine/glycine with arginine has been shown to increase the thermal stability of proteins (Vogt *et al.*, 1997; Nishio *et al.*, 2003). For example, a single amino acid change, for example, from a threonine to an alanine was beneficial for protein binding under increased temperature in Pacific damselfishes (Johns and Somero, 2004).

Several of the mentioned substitutions occurred in the high temperature replicate, however, we currently have no information about where in the protein the changes occurred and could only speculate about the impact on the final protein and the adaptive benefit. Further analysis of the impact of the identified AA changes would be needed to pinpoint adaptive modifications of proteins in Tp.

4.5 Summary and conclusion

The genome-wide RNA-Seq analysis of low and high temperature transcriptomes provided first insights into the molecular underpinnings of the evolution of thermal tolerance in *Thalassiosira pseudonana*. Low temperature cultures showed an overrepresentation of fatty acid desaturase genes suggesting homeoviscous adaptation of membranes. Temperature specific expression of TEs suggests a potential thermal stress induced mechanism to produce differences in gene regulation. Furthermore, down-regulation of mismatch repair mechanisms suggests means of generating genetic variation on which selection can act. Whole genome resequencing of one high temperature replicate showed increased accumulation of mutation on chromosome 14 potentially affecting DNA transcription regulation and chromosome organization. These findings provided first insights into transcriptional and genomic processes enabling thermal adaptation for a marine diatom.

Chapter 5

General discussion

5.1 Summary of main results

Diatom ecology in the Atlantic Ocean Diatom diversity (based on alpha diversity) along a longitudinal (Fram Strait) and latitudinal (Atlantic Ocean) transect was shown to be mostly governed by temperature. Temperature had a negative effect on diatom diversity in the Fram Strait transect with stations showing low diversity at higher temperatures. In contrast, diatom diversity in the Atlantic Ocean was negatively correlated to temperature with high temperature showing low diatom diversity. The most wide-spread and abundant diatom order were the Coscinodiscales which were present at almost all stations.

Phenotypic responses to temperature selection Low and high temperature conditions resulted in the selection of adaptive phenotypes. Cultures under low temperature selection showed reduced growth under ancestral conditions. Fv/Fm increased continuously for low temperature replicates while high temperature cultures showed no recovery. Additionally, four out of five low temperature replicates increased their cell volume over the course of the experiment. One replicate reduced its cell volume following the initial increase. Furthermore, the elemental composition (C:N, C:bSi, C:chla) of low temperature cultures was significantly different to control cultures.

Genetic responses to temperature selection Tp's transcriptome responded differently to low and high temperature selection. High temperature transcriptomes were very similar after 50 and 300 generations of selection. In contrast, low temperature transcriptomes after 50 generations showed a high number of differential expressed genes and a reduction by 75% after 300 generations. Additionally, we found no significant difference between the transcriptomes of the two different low temperature phenotypes. Genome resequencing of one high temperature replicate after 300 generations of selection showed over 6000 mutations with about a fourth located on chromosome 14.

5.2 Discussion

Diatoms are one of the most successful groups of phytoplankton and produce about 50% of the annual primary production (Nelson *et al.*, 1995). With ongoing climate change they experience multiple, simultaneous changes (CO_2 concentration, temperature, stratification, light availability, nutrient availability), in their environment (Boyd *et al.*, 2015*a*) with temperature having the strongest effect on diatoms (Boyd *et al.*, 2013; Tatters *et al.*, 2013*a*; Schlüter *et al.*, 2014). However, little is know about phenotypic outcomes and the molecular basis of temperature adaptation in diatoms.

Tp's adaptation to the implemented temperature conditions might have been accomplished by different means. Combining the transcriptome responses (chapter 4) and the phenotypic outcomes (chapter 3), it seems that low temperature and high temperature selection lines have adapted genetically.

The transcriptomic response of low temperature lines showed a decrease in differentially expressed genes after 300 generations in comparison to 50 generations. High numbers of differentially expressed genes in the first generations under novel temperature conditions have been reported for a variety of other species (e.g. *E. coli*) (Riehle *et al.*, 2003; Ying *et al.*, 2015) and *F. cylindrus* (Strauss, 2012) but has been shown to revert towards a baseline with adaptation

(Ying *et al.*, 2015; Riehle *et al.*, 2003; Fong *et al.*, 2005). This potential reversion was observed for low temperature Tp lines but not for high temperature lines. Low temperature impacts most physiological processes of a cell by slowing reactions, such as translation, and decreasing membrane viscosity. Decreased translation efficiency can be compensated by increased expression of ribosomes (Conrad *et al.*, 2011). Increased expression of ribosomal protein genes was found for the first 50 generations of low temperature selection and slightly reduced after 300 generations (Figure 4.3.11). Thus, as ribosomes consist of P-rich RNAs, the cellular N:P ratio decreases as shown in chapter 3. The decreased expression of ribosomal genes after 300 generations might explain the modest change in the N:P ratio. This potentially underlines the link of ribosome content and cellular P quotas (Loladze and Elser, 2011).

Protein synthesis is resource expensive (Conrad *et al.*, 2011), therefore, increased translation might not be feasible in the long term and would result in different transcriptomic states (Figure 4.3.1). Functional changes of proteins, beneficial under low temperature, might be more cost effective. This could be achieved by introduction of non-synonymous mutations causing, for example, amino acid charge changes resulting in conformation changes of the final protein. Additionally, transposition of transposable elements can disrupt or alter protein function. Functional changes of enzymes and proteins could result in poorer phenotypes under ancestral conditions and has for example been suggested in *E. coli* (Kishimoto *et al.*, 2010), yeast (Caspeta *et al.*, 2014) and *E. huxleyi* (**?**). The reduced fitness under ancestral conditions, as shown in chapter 3, for low temperature lines supports the hypothesis of mutation driven adaptation. However, this argumentation is purely speculative as we currently cannot pinpoint the genetic changes facilitating an evolutionary response of low temperature lines.

High temperature lines showed genetic adaptation based on the reciprocal transplant assay (Figure 3.3.5, additionally, their transcriptomic similarity after 50 and 300 generations might suggest genetic adaptation within the first 50 generations (Figure 4.3.1). High temperature has been shown to increase occurrence of oxygen radicals which can cause genetic mutations (Allen *et al.*, 2006). This is one means by which organisms can increase their genotypic diversity. Additionally, down-regulation of mismatch repair mechanisms has been linked to the generation of mutator phenotypes (Conrad *et al.*, 2011; Kishimoto *et al.*, 2010). Indeed, the mutation rate

(10^{-7}) was high, compared to *E. coli* (10^{-9}-10^{-13}) (Blount *et al.*, 2008), however, similar rates have been observed for *C. reinhardtii* (10^{-8}) under salinity selection (Perrineau *et al.*, 2014). Additionally, mutations were mostly found in genes related to chromosome organization and regulation of transcription and translation (Figure 4.3.14). This is in agreement with other studies reporting mutations in genes related to cell replication, transcription and cell cycle providing adaptive thermal tolerance (Caspeta *et al.*, 2014; Knies *et al.*, 2006; Chang *et al.*, 2013). Furthermore, novel transcriptome states have been reported for *E. coli* after high temperature selection suggesting temperature dependent reorganization of global gene expression patterns (Ying *et al.*, 2015; Kishimoto *et al.*, 2010). Ying et al. (2015) argue that temperature adaptation in *E. coli* was mediated through a novel steady state of the cellular transcriptome. This would imply that thermal adaptation required continuous optimization of cellular gene expression. Thus, transcriptomes measured at different timepoints should show distinct clusters. This was not the case for the high temperature transcriptomes (Figure 4.3.1) which might also suggest adaptation within the first 50 generations. It is, however, currently unclear if the transcriptomic changes correlate with genotype changes.

5.2.1 Broader impact of results

Several modelling and field studies have predicted and shown that marine species are shifting northwards due to ocean warming (Barton *et al.*, 2016; Marinov *et al.*, 2013; Steinacher *et al.*, 2010; Walther *et al.*, 2002). This might cause changes to phytoplankton communities and biogeochemical cycles.

Chust et al. (2013) have shown that environmental conditions were more important for diatom diversity than geographical distance. This is in agreement with the findings in this thesis were environmental conditions, specifically temperature, had a significant impact on diatom diversity. As the oceans are warming, this might lead to a turn over of diatom communities as has been suggested by several modelling studies that found a northwards distribution shift of phytoplankton, e.g. diatom, communities (Barton *et al.*, 2016; Marinov *et al.*, 2013; Steinacher *et al.*, 2010). As temperature has been identified as the most important driver of diatom abundance and diversity it can be expected that diatom species potentially track their optimum temperature, invading communities further north. This is likely to increase species richness but

also competition in the temperate regions of the North Atlantic (Barton *et al.*, 2016). Temporal studies have shown, for example, that diatom communities are increasing in the North Atlantic, outcompeting dinoflagellates (Hinder *et al.*, 2012; Barton *et al.*, 2016). In addition, Barton et al. (2016) suggested that not only do diatom communities shift in space, their composition is also changing. As diatoms are an important functional group, climate change driven alterations of their distribution might have drastic effects on elemental fluxes and food webs. We have shown that adaptation to high temperature increases the carbon to nitrogen ratio (C:N) (Chapter 3, Figure 3.3.11). This potentially decreases the nutritional value for zooplankton predators as they have to feed on more cells to satisfy their N requirement (**?**). However, this might increase the carbon draw down by the biological pump to the deep sea, as zooplankton excrete excess carbon.

Furthermore, high temperature adapted Tp had a higher biogenic silica content per surface area which might cause a greater ballasting effect in marine snow (Klaas and Archer, 2002) and fecal pellets (Ploug *et al.*, 2008) further enhancing the biological pump. However, ocean warming related increases in respiration might mitigate this effect (Ragueneau *et al.*, 2006) as the sinking particles are potentially recycled before they reach the deep sea.

Even though climate change is a global phenomenon, regional changes have been shown to be highly variable with, for example, much faster warming in the polar regions compared to the tropical or temperate regions (Graversen *et al.*, 2008; Marinov *et al.*, 2013; Philippart *et al.*, 2011).

Diatoms prefer colder nutrient-rich regions, such as the temperate and polar Atlantic (Barton *et al.*, 2015) and might retract further north or become less abundant in the absence of an evolutionary response. Several studies have investigated the adaptation potential and outcome of diatoms, coccolithophores, cyanobacteria and green algae, to increases in temperature and ocean acidification (Benner *et al.*, 2013; Collins and Bell, 2006; Tatters *et al.*, 2013*a*; Lohbeck *et al.*, 2012; Hutchins *et al.*, 2015). All studies, and this thesis, showed that phytoplankton is able to adapt within a year to these specific changes in their environment. However, their adaptation potential in the ocean might be hindered as climate change is affecting several environmental parameters at the same time. Nevertheless, phytoplankton communities in the ocean have a high genetic standing variation and turnover rates potentially providing genetic variants adaptive to future climate change conditions. Tp and maybe diatoms in general, posses mechanisms to

increase their genetic diversity through, for example, transposable elements or down-regulation of mismatch repair pathways (see chapter 4).

However, the presence of species with beneficial genotypes might hinder adaptation of other species (De Mazancourt *et al.*, 2008) as studies have shown that monocultures respond differently to selection than mixed communities (Collins, 2011).

5.3 Future work

Experimental evolution is a useful tool to investigate biological responses of phytoplankton communities to environmental change. Genetic and epigenetic processes of adaptation in phytoplankton are currently not well understood. However, with more studies in this area of research, we might be able to better relate the (epi)genotype of phytoplankton to their phenotypes, which is key for our understanding how the environment shapes the evolution of microbes and biogeochemical processes they drive.

Allele frequency analysis of all replicated selection lines can pinpoint mutations with high fitness benefits. Genome re-sequencing enabled us to identify genome-wide SNPs, however we have no information how genotype frequencies changed and how the genes variants are expressed. Additionally, allele analysis will provide evidence of novel mutation or genotype sorting, as we currently do not know whether we selected pre-existing variants or if our phenotypes are based on novel mutations.

Our data suggest temperature specific expression of different transposon families, potentially diversifying genotypes. Transposon expression levels reported in chapter 4 might have been high enough to cause transposition (1000:1 protein:transcript required). Identifying new insertion sites of transposons might allow correlation with expression of nearby genes or genes the TE inserted into. This might be key to understand the plasticity of diatoms and therefore their success in many different environments.

The experimental setup used in this thesis was used to gain mechanistic understanding of

temperature adaptation and has limited transferability to the natural environment. Nevertheless, the identified transcriptome profiles and mutations could be used to identify ongoing temperature stress and potential adaptation in the changing ocean environment. The Sea of Change project provides metatranscriptomes and -genomes that could be used to investigate temperature adaptation in the contemporary ocean.

Univariate laboratory experiments with the most influential drivers provide valuable insights, however, they do not reflect processes in the field. Multivariate experiments allow better representation of the natural environment, but also make it more difficult to identify the most important driver. Nevertheless, feedbacks between several environmental factors might cause different responses than univariate factors. Furthermore, including communities rather than single species/genotypes in these kind of studies might allow a better understanding of community responses which could be integrated into current models to allow better prediction of future phytoplankton communities in the ocean.

5.3.1 From genomes to biomes

Through the advancement of transcriptomics and metatranscriptomics it has become possible to utilize sequence data in experimental field studies. Gene expression pattern have, so far, been linked to denitrification in soil (Lennon and Houlton, 2016) and gene abundance in the English channel to metabolite turnover of microbes (Larsen *et al.*, 2015). In the medical field, gene expression differences between phenotypic similar tissue cells are used to identify certain types of cancer before they develop into distinct phenotypes (Jones *et al.*, 2005; Selman *et al.*, 2006). Thus, identifying the transcriptomic fingerprint of phytoplankton species adapting to environmental conditions could be used to identify adaptation in the ocean. Measurements of gene expression in natural phytoplankton communities, mesocosms or microcosms in the laboratory, for example, allows tracking genes responsive to environmental conditions. Thus, environmental perturbations could be identified as gene expression patterns before they may lead to changes in phytoplankton communities. However, our understanding of the plasticity of gene expression to environmental change and the phenotypic effects is limited, hindering our ability to predict phytoplankton species and community responses to climate change (Doney *et al.*, 2004). Experimental evolution experiments combined with omics approaches, as in this thesis, are ideal

for identifying phenotypic and transcriptomic responses to long-term environmental changes, as they allow measuring phenotypic trait changes (e.g. fitness, elemental composition, etc.) and the related transcriptomic fingerprint. Linking trait changes to transcriptome responses can help refine biogeochemical models and by including adaptive responses, predict potential future biogeochemical cycles and phytoplankton communities (Litchman *et al.*, 2015). For example, a trait-based community model approach allows incorporation of trait evolution to changing environmental conditions which would not be resolved by species-based models (Litchman *et al.*, 2009; Merico *et al.*, 2009). Kramer 2014 used such trait evolution to model the growth season onset of trees allowing adaptation to temperature changes. Measuring of phytoplankton traits and trait evolution to changing environmental conditions would be beneficial to improving trait-based community models. Gobet et al. (2012) suggested that about 320 species might be the main drivers of biogeochemical cycles and these would be ideal candidates for experimental evolution experiments. The parameterization of models based on traits and transcriptomic fingerprints will be vital in gaining a mechanistic understanding of ecosystem function in the contemporary and future ocean (Litchman *et al.*, 2015; Boyd *et al.*, 2015*b*; Delong, 2009).

Chapter 6

Appendix

6.1 Supplementary information for chapter 2

Temperature and salinity showed co-correlation and based on VIF, salinity was excluded from the environmental metadata for the Fram Strait transect. Because of the co-correlation, salinity and temperature are interchangeable and both need to be taken into account. Figure 6.1.1 shows the correlation between salinity and alpha diversity for the Fram Strait transect.

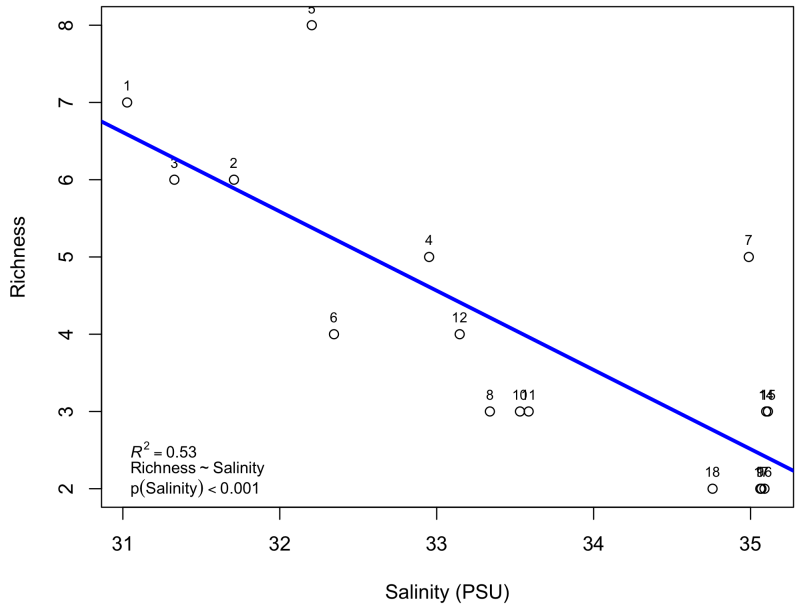

Figure 6.1.1: Correlation between richness and salinity for Fram Strait transect.

Figure 6.1.2 shows the outcome of RDA model selection when salinity was included instead of temperature.

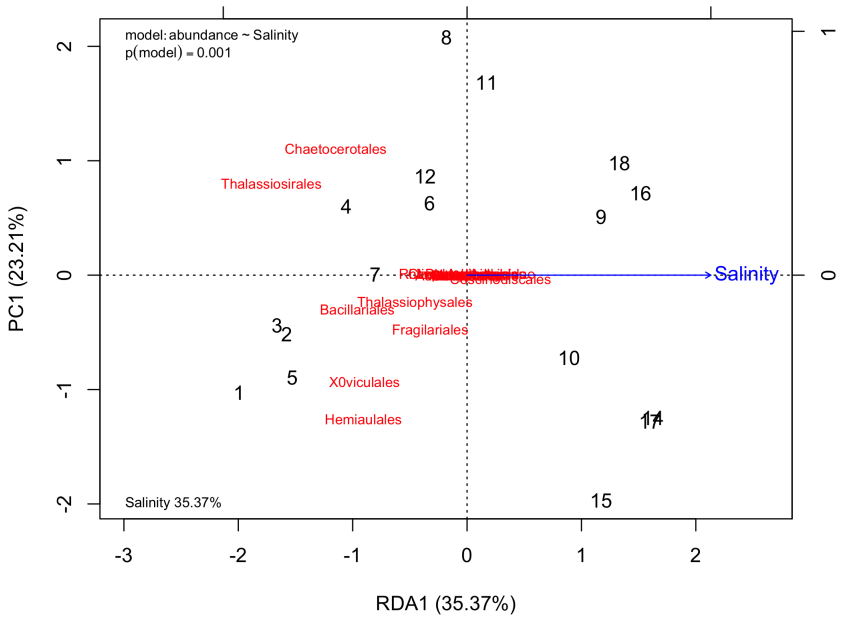

Figure 6.1.2: RDA for the Fram Strait transect including salinity instead of temperature

6.2 Supplementary information for chapter 3

6.2.1 SeaChem Marine SaltTM

Typical ion composition of SeaChem Marine SaltTM is summarized in Table 6.2.1. This mix was used as the basis of culture medium to which F-strength nutrients were added after Guillard (1979).

Table 6.2.1: Typical Ionic Concentrations of Marine SaltTM.

Ion	concentration (mg L^{-1})
Chloride	19,336
Sodium	10,752
Sulfate	2,657
Magnesium	1,317
Potassium	421
Calcium	380
Carbonate,Bicarbonate	142
Strontium	9.5
Boron	16
Bromide	56
Iodide	0.060
Lithium	0.3
Silicon	<0.1
Iron	0.0098
Copper	0.0003
Nickel	<0.015
Zinc	0.0107
Manganese	0.0023
Molybdenum	0.0098
Cobalt	0.0004
Vanadium	<0.015
Selenium	<0.019
Rubidium	0.118
Barium	0.04

6.2.2 Reciprocal transplant assay growth curves

Figure 6.2.1 shows growth curves of low and high temperature cultures during the reciprocal transplant assay. Low temperature replicate 5 (smaller phenotype) had a higher yield when

reaching stationary phase, however, growth rates showed no significant difference to the other low temperature replicates.

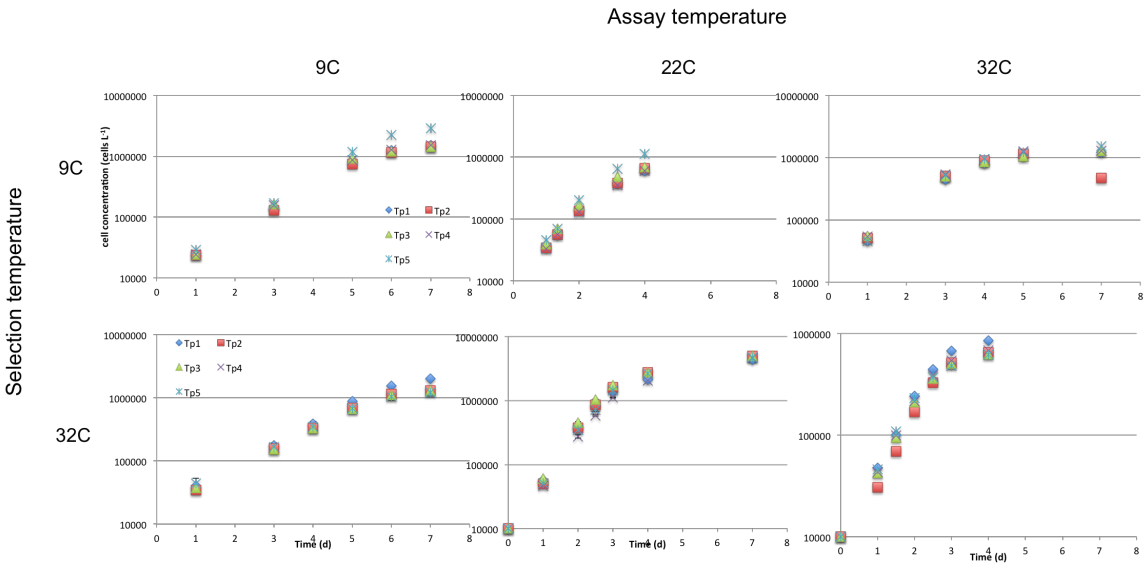

Figure 6.2.1: Growth curves of selection cultures during reciprocal transplant assay for each replicate (Tp1-5)

6.2.3 Chlorophyll a calibration curve

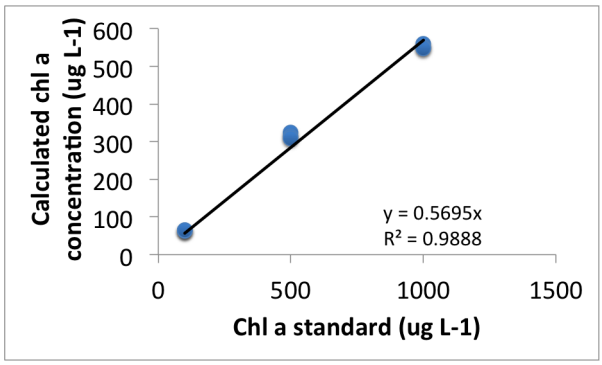

Figure 6.2.2: Calibration curve for chlorophyll a concentration calculations.

6.2.4 Temperature response curve of the 32°C selection lines at two different light intensities

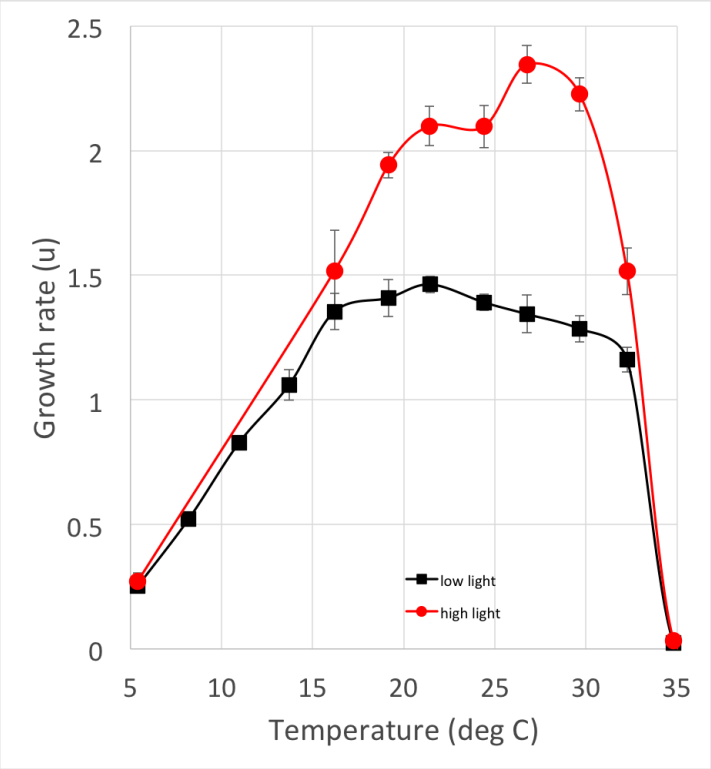

Figure 6.2.3: Temperature response curve of the high temperature replicate at two different light intensities.

6.3 Supplementary information for chapter 4

InterPro domain analysis T1 9C down reg

6.4 Publication output

6.4.1 A novel cost effective and high-throughput isolation and identification method for marine microalgae

This article was published in 2014 in Plant Methods. Martin T. Jahn established a quick and cost effective isolation and identification method of microalgae collected during the ARK27-1 field

campaign.

Table 6.4.1: Overrepresented InterPro domains of down-regulated genes at T1 9°C. P-values were adjusted after Benjamin-Hochberg with a cut-off of 0.05

InterPro ID	Overrepresented pvalue	InterPro description	Domain description
IPR002341	8.03E-09	HSF/ETS, DNA-binding	HSF ETS
IPR000232	1.13E-08	Heat shock factor (HSF)-type, DNA-binding	sp HSF1 ARATH P41151
IPR002557	0.001	Chitin binding Peritrophin-A	CBM 14
IPR001223	0.001	Glycoside hydrolase, family 18	Glyco hydro 18
IPR011583	0.001	Chitinase II	Glyco 18
IPR002110	0.002	Ankyrin	ANKYRIN
IPR000726	0.002	Glycoside hydrolase, family 19	sp O50152 STRGR O50152
IPR006970	0.002	PT repeat	PT
IPR001579	0.008	Glycoside hydrolase, chitinase active site	CHITINASE 18
IPR007197	0.009	Radical SAM	Radical SAM
IPR000668	0.010	Peptidase C1A, papain	sp Q9ST61 SOLTU Q9ST61
IPR006671	0.015	Cyclin, N-terminal	Cyclin N
IPR004843	0.016	Metallophosphoesterase	Metallophos
IPR003613	0.019	Zn-finger, modified RING	U-box
IPR004835	0.019	Fungal chitin synthase	Chitin synth 2

Continued on next page

Continued from previous page

InterPro ID	Overrepresented p-value	InterPro description	Domain description
IPR006025	0.021	Peptidase M, neutral zinc metallopeptidases, zinc-binding site	ZINC PROTEASE
IPR003409	0.021	MORN motif	MORN
IPR000583	0.037	Glutamine amidotransferase, class-II	GATase 2
IPR002085	0.037	Zinc-containing alcohol dehydrogenase superfamily	ADH zinc N
IPR000209	0.037	Peptidase S8 and S53, subtilisin, kexin, sedolisin	SUBTILASE ASP
IPR000210	0.037	BTB/POZ	BTB
IPR000182	0.037	GCN5-related N-acetyltransferase	Acetyltransf 1
IPR004184	0.042	Pyruvate formate-lyase, PFL	PFL
IPR006090	0.042	Acyl-CoA dehydrogenase, C-terminal	Acyl-CoA dh
IPR001989	0.042	Radical-activating enzyme	RADICAL ACTIVATING

Table 6.4.2: Overrepresented InterPro domains of up-regulated genes after 50 generations at 9°C. P-values were adjusted after Benjamin-Hochberg with a cut-off of 0.05.

InterPro ID	Overrepresented p-value	InterPro description	Domain description
IPR006311	0.003	Twin-arginine translocation pathway signal	TAT signal seq
IPR006073	0.004	GTP1/OBG	GTP1OBG
IPR004161	0.008	Elongation factor Tu, domain 2	GTP EFTU D2
IPR005289	0.010	GTP-binding	MG442
IPR000051	0.011	SAM (and some other nucleotide) binding motif	SAM BIND
IPR000836	0.014	Phosphoribosyltransferase	Pribosyltran
IPR002300	0.014	Aminoacyl-tRNA synthetase, class Ia	tRNA-synt 1
IPR000886	0.016	Endoplasmic reticulum targeting sequence	ER TARGET
IPR001179	0.019	Peptidylprolyl isomerase, FKBP-type	FKBP C
IPR002067	0.019	Mitochondrial carrier protein	MITOCARRIER
IPR001478	0.027	PDZ/DHR/GLGF	PDZ
IPR002125	0.028	Cytidine/deoxycytidylate deaminase, zinc-binding region	dCMP cyt deam
IPR008775	0.029	Phytanoyl-CoA dioxygenase	PhyH
IPR006204	0.034	GHMP kinase	GHMP kinases
IPR000357	0.035	HEAT	HEAT REPEAT

Continued on next page

Continued from previous page

InterPro ID	Overrepresented p-value	InterPro description	Domain description
IPR001241	0.035	DNA topoisomerase II	TPI2FAMILY
IPR001697	0.035	Pyruvate kinase	sp Q9A6N6 CAUCR Q9A6N6
IPR002205	0.035	DNA gyrase/topoisomerase IV, subunit A	sp TOP2 YEAST P06786
IPR005841	0.035	Phosphoglucomutase/phosphomannomutase	PGMPMM
IPR005844	0.035	Phosphoglucomutase/phosphomannomutase alpha/beta/alpha domain I	PGM PMM I
IPR005845	0.035	Phosphoglucomutase/phosphomannomutase alpha/beta/alpha domain II	PGM PMM II
IPR011558	0.035	DNA gyrase B	sp TOP2 ARATH P30182

Table 6.4.3: Overrepresented InterPro domains of down-regulated genes at TE 9°C. P-values were adjusted after Benjamin-Hochberg with a cut-off of 0.05.

InterPro ID	Overrepresented pvalue	InterPro description	Domain description
IPR000232	2.93E-06	Heat shock factor (HSF)-type, DNA-binding	sp HSF1 ARATH P41151
IPR002341	2.93E-06	HSF/ETS, DNA-binding	HSF ETS
IPR000413	5.40E-05	Integrins alpha chain	INTEGRIN ALPHA 2
IPR002068	8.88E-05	Heat shock protein Hsp20	HSP20
IPR006663	0.001	Thioredoxin domain 2	THIOREDOXIN 2
IPR006662	0.002	Thioredoxin-related	THIOREDOXIN
IPR000209	0.003	Peptidase S8 and S53, subtilisin, kexin, sedolisin	SUBTILASE ASP
IPR001107	0.004	Band 7 protein	Band 7
IPR002625	0.004	Smr protein/MutS2 C-terminal	Smr
IPR005123	0.006	2OG-Fe(II) oxygenase superfamily	2OG-FeII Oxy
IPR001140	0.008	ABC transporter, transmembrane region	ABC membrane
IPR000886	0.012	Endoplasmic reticulum targeting sequence	ER TARGET
IPR003998	0.013	Twin-arginine translocation protein TatB	TATBPROTEIN
IPR006970	0.014	PT repeat	PT
IPR003590	0.022	Leucine-rich repeat, ribonuclease inhibitor subtype	LRR RI

Continued on next page

Continued from previous page

InterPro ID	Overrepresented pvalue	InterPro description	Domain description
IPR006620	0.024	Prolyl 4-hydroxylase, alpha subunit	P4Hc
IPR008454	0.027	Cna B-type	Cna B
IPR006050	0.030	DNA photolyase, N-terminal	DNA photolyase
IPR001576	0.031	Phosphoglycerate kinase	PHGLYCKINASE
IPR002939	0.032	Chaperone DnaJ, C-terminal	DnaJ C
IPR011594	0.033	Thioredoxin-like	sp Q9ZR86 MAIZE Q9ZR86
IPR001023	0.040	Heat shock protein Hsp70	sp DNAK ODOSI P49463
IPR003613	0.041	Zn-finger, modified RING	U-box
IPR001849	0.041	Pleckstrin-like	PH
IPR000595	0.042	Cyclic nucleotide-binding	cNMP binding

Table 6.4.4: Overrepresented InterPro domains of up-regulated genes at TE 9°C. P-values were adjusted after Benjamin-Hochberg with a cut-off of 0.05.

InterPro ID	Overrepresented p-value	InterPro description	Domain description
IPR005804	3.35E-04	Fatty acid desaturase	FA desaturase
IPR010257	7.32E-04	Fatty acid desaturase subdomain	sp Q84KI8 PHATR Q84KI8
IPR008917	0.002	Eukaryotic transcription factor, DNA-binding	A DNA-binding domain in eukaryotic transcription factors
IPR002076	0.002	GNS1/SUR4 membrane protein	ELO
IPR001092	0.006	Basic helix-loop-helix dimerisation region bHLH	HLH
IPR002067	0.007	Mitochondrial carrier protein	MITOCARRIER
IPR000014	0.009	PAS	PAS
IPR005172	0.011	Tesmin/TSO1-like, CXC	CXC
IPR001199	0.012	Cytochrome b5	Cyt-b5
IPR006210	0.025	Type I EGF	EGF
IPR000804	0.039	Clathrin adaptor complex, small chain	Clat adaptor s
IPR003299	0.039	Flagellar calcium-binding protein (calflagin)	CALFLAGIN
IPR000837	0.039	Fos transforming protein	LEUZIPPRFOS
IPR010527	0.039	Photosystem II 12 kDa extrinsic	PsbU

Continued on next page

Continued from previous page

InterPro ID	Overrepresented p-value	InterPro description	Domain description
IPR003697	0.040	Maf-like protein	Maf
IPR009631	0.040	Protein of unknown function DUF1230	DUF1230
IPR001104	0.040	3-oxo-5-alpha-steroid 4-dehydrogenase, C-terminal	S5A REDUCTASE
IPR001899	0.040	Surface protein from Gram-positive cocci, anchor region	LPXTG anchor
IPR003042	0.041	Aromatic-ring hydroxylase	RNGMNOXGNASE
IPR005322	0.041	Peptidase U34, dipeptidase	Peptidase U34
IPR002606	0.041	Riboflavin kinase/FAD synthetase	sp Q9CQ95 MOUSE Q9CQ95
IPR000961	0.041	Protein kinase, C-terminal	S TK X
IPR003834	0.041	Cytochrome c biogenesis protein, transmembrane region	DsbD
IPR000385	0.041	MoaA/nifB/pqqE	MOAA NIFB PQQE
IPR010505	0.041	Molybdenum cofactor synthesis C	Mob synth C
IPR001087	0.041	Lipolytic enzyme, G-D-S-L	LIPASE GDSL
IPR004097	0.041	DHHA2	DHHA2
IPR000277	0.041	Cys/Met metabolism pyridoxal-phosphate-dependent enzymes	Cys Met Meta PP
IPR004803	0.041	Queuine tRNA-ribosyltransferase	Q tRNA tgt
IPR001313	0.041	Pumilio/Puf RNA-binding	PUF

Continued on next page

Continued from previous page

InterPro ID	Overrepresented p-value	InterPro description	Domain description
IPR006641	0.042	Ribonuclease RNase H fold containing	YqgFc
IPR010402	0.043	CCT	CCT
IPR003392	0.044	Patched	Patched
IPR000232	0.044	Heat shock factor (HSF)-type, DNA-binding	sp HSF1 ARATH P41151
IPR002790	0.044	Protein of unknown function DUF88	DUF88
IPR002049	0.045	Laminin-type EGF-like	EGFLAMININ
IPR000884	0.046	Thrombospondin, type I	TSP 1
IPR001294	0.046	Phytochrome	PHYTOCHROME
IPR003661	0.046	Histidine kinase A, N-terminal	HisKA
IPR011072	0.046	Protein kinase PKN/PRK1, effector	Effector domain of the protein kinase pkn/prk1
IPR001258	0.046	NHL repeat	NHL
IPR000086	0.049	NUDIX hydrolase	NUDIXFAMILY
IPR001993	0.049	Mitochondrial substrate carrier	Mito carr

Table 6.4.5: Overrepresented InterPro domains of down-regulated genes at TE 32°C. P-values were adjusted after Benjamin-Hochberg with a cut-off of 0.05.

InterPro ID	Overrepresented pvalue	InterPro description	Domain description
IPR000232	2.34E-06	Heat shock factor (HSF)-type, DNA-binding	sp HSF1 ARATH P41151
IPR002341	5.02E-06	HSF/ETS, DNA-binding	HSF ETS
IPR006671	2.37E-05	Cyclin, N-terminal	Cyclin N
IPR002557	0.001	Chitin binding Peritrophin-A	CBM 14
IPR001347	0.001	Sugar isomerase (SIS)	SIS
IPR004835	0.002	Fungal chitin synthase	Chitin synth 2
IPR000726	0.003	Glycoside hydrolase, family 19	sp O50152 STRGR O50152;
IPR001199	0.004	Cytochrome b5	Cyt-b5
IPR004367	0.008	Cyclin, C-terminal	Cyclin C
IPR003607	0.008	Metal-dependent phosphohydrolase, HD region	HDc
IPR006670	0.009	Cyclin	CYCLIN
IPR001107	0.010	Band 7 protein	Band 7
IPR001810	0.012	Cyclin-like F-box	FBOX
IPR001353	0.013	20S proteasome, A and B subunits	Proteasome
IPR006662	0.017	Thioredoxin-related	THIOREDOXIN

Continued on next page

Continued from previous page

InterPro ID	Overrepresented p-value	InterPro description	Domain description
IPR003961	0.018	Fibronectin, type III	fn3
IPR002625	0.019	Smr protein/MutS2 C-terminal	Smr
IPR006970	0.019	PT repeat	PT
IPR000192	0.020	Aminotransferase, class V	Aminotran 5
IPR001254	0.020	Peptidase S1, chymotrypsin	Trypsin
IPR001865	0.022	Ribosomal protein S2	RIBOSOMAL S2 1
IPR006663	0.022	Thioredoxin domain 2	THIOREDOXIN 2
IPR004843	0.027	Metallophosphoesterase	Metallophos
IPR007125	0.028	Histone core	Histone
IPR001579	0.033	Glycoside hydrolase, chitinase active site	CHITINASE 18
IPR000583	0.037	Glutamine amidotransferase, class-II	GATase 2
IPR007124	0.038	Histone-fold/TFIID-TAF/NF-Y	HIST TAF
IPR001223	0.040	Glycoside hydrolase, family 18	Glyco hydro 18
IPR011583	0.040	Chitinase II	Glyco 18

Table 6.4.6: Overrepresented InterPro domains of up-regulated genes at TE 32°C. P-values were adjusted after Benjamin-Hochberg with a cut-off of 0.05.

InterPro ID	Overrepresented pvalue	InterPro description	Domain description
IPR001283	3.07E-05	Allergen V5/Tpx-1 related	sp Q8S9B6 EEEEE Q8S9B6
IPR006970	0.002	PT repeat	PT
IPR005804	0.002	Fatty acid desaturase	FA desaturase
IPR001993	0.004	Mitochondrial substrate carrier	Mito carr
IPR001104	0.004	3-oxo-5-alpha-steroid 4-dehydrogenase, C-terminal	S5A REDUCTASE
IPR002068	0.005	Heat shock protein Hsp20	HSP20
IPR001952	0.009	Alkaline phosphatase	ALKALINE PHOSPHATASE
IPR001296	0.018	Glycosyl transferase, group 1	Glycos transf 1
IPR003672	0.027	CobN/magnesium chelatase	CobN-Mg chel
IPR006061	0.028	Bacterial extracellular solute-binding protein, family 1 domain	SBP BACTERIAL 1
IPR000523	0.028	Magnesium chelatase, ChlI subunit	Mg chelatase
IPR001341	0.029	Aspartate kinase	asp kinases
IPR003091	0.029	Voltage-dependent potassium channel	KCHANNEL
IPR002350	0.030	Proteinase inhibitor I1, Kazal	Kazal 1
IPR011497	0.030	Protease inhibitor, Kazal-type	Kazal 2

Continued on next page

Continued from previous page

InterPro ID	Overrepresented p-value	InterPro description	Domain description
IPR006667	0.031	MgtE integral membrane region	MgtE
IPR001834	0.031	NADH:cytochrome b5 reductase (CBR)	CYTB5RDTASE
IPR000740	0.032	GrpE protein	GRPEPROTEIN
IPR009249	0.032	Ferredoxin-dependent bilin reductase	Fe bilin red
IPR001148	0.032	Carbonic anhydrase, eukaryotic	sp Q9ZK30 HELPJ Q9ZK30
IPR006076	0.040	FAD dependent oxidoreductase	DAO

A novel cost effective and high-throughput isolation and identification method for marine microalgae

Martin T Jahn[1,2]
Email: martin.jahn@stud-mail.uni-wuerzburg.de

Katrin Schmidt[1]
Email: K.Schmidt@uea.ac.uk

Thomas Mock[1*]
[*] Corresponding author
Email: t.mock@uea.ac.uk

[1] School of Environmental Sciences, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK

[2] Current address: Department of Botany II, Julius-Maximilians University Würzburg, Julius-von-Sachs-Platz 3, 97082 Würzburg, Germany

Abstract

Background

Marine microalgae are of major ecologic and emerging economic importance. Biotechnological screening schemes of microalgae for specific traits and laboratory experiments to advance our knowledge on algal biology and evolution strongly benefit from culture collections reflecting a maximum of the natural inter- and intraspecific diversity. However, standard procedures for strain isolation and identification, namely DNA extraction, purification, amplification, sequencing and taxonomic identification still include considerable constraints increasing the time required to establish new cultures.

Results

In this study, we report a cost effective and high-throughput isolation and identification method for marine microalgae. The throughput was increased by applying strain isolation on plates and taxonomic identification by direct PCR (dPCR) of phylogenetic marker genes in combination with a novel sequencing electropherogram based screening method to assess the taxonomic diversity and identity of the isolated cultures. For validation of the effectiveness of this approach, we isolated and identified a range of unialgal cultures from natural phytoplankton communities sampled in the Arctic Ocean. These cultures include the isolate of a novel marine Chlorophyceae strain among several different diatoms.

Conclusions

We provide an efficient and effective approach leading from natural phytoplankton communities to isolated and taxonomically identified algal strains in only a few weeks. Validated with sensitive Arctic phytoplankton, this approach overcomes the constraints of standard molecular characterisation and establishment of unialgal cultures.

6.4.2 The role of phenotypic plasticity and epigenetics in experimental evolution with phytoplankton

This mini-review was published in 2016 in Perspectives in Phycology. Parts of it were included in the general introduction.

Perspectives in Phycology, Vol. 3 (2016), Issue 1, p. 29–36
Published online January 2016

Article

The role of phenotypic plasticity and epigenetics in experimental evolution with phytoplankton

Katrin Schmidt[1], Cock van Oosterhout[1], Sinéad Collins[2] & Thomas Mock[1*]

[1] School of Environmental Sciences, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ UK
[2] Institute of Evolutionary Biology, University of Edinburgh, Kings Building, Ashworth Laboratories, West Mains Road, Edinburgh, EH9 3FL UK
* Corresponding author: t.mock@uea.ac.uk

With 2 figures

Abstract: Phytoplankton experience multiple, simultaneous and interacting changes in their environment, which affect primary production, community composition and biogeochemical cycles. Many environmental fluctuations occur on short timescales, requiring plastic phenotypes for appropriate physiological responses. Although there is evidence for high phenotypic plasticity in phytoplankton, the underlying mechanisms and their role for evolutionary responses remain elusive but are essential for predicting how phytoplankton respond to climate change. Elucidating the molecular basis of sensing environmental changes and subsequent regulation of plasticity including epigenetics will provide new insights into reaction norms and therefore how species tolerate, persist and eventually adapt to changing environmental conditions. Quantifying the role of phenotypic plasticity in phytoplankton for adaptation is an increasingly important field of research in experimental evolution.

Keywords: Experimental evolution, phytoplankton, epigenetics, phenotypic plasticity, adaptation, microalgae, ocean

Introduction

Experimental evolution (EE) approaches are increasingly used to investigate evolutionary processes of organisms to controlled experimental conditions and can be applied to study several facets of adaptation and evolution (e.g. estimate mutation rates or evaluate evolutionary theories and concepts) (Kawecki et al. 2012). Examples of EE experiments include studies on organisms ranging all three domains of life: Bacteria, Archaea and Eukaryota (e.g. *Chlamydomonas reinhardtii*); (e.g. Bell 1997; Ratcliff et al. 2013; Collins & De Meaux 2009; Perrineau et al. 2013), *Escherichia coli* (e.g. Plucain et al. 2014; Lenski et al. 1991; Blount et al. 2012), *Saccharomyces cerevisiae* (e.g. Lewis et al. 1995; Dhar et al. 2011; Ratcliff et al. 2012), *Arabidopsis thaliana* (e.g. Bond & Baulcombe 2015; Rahavi 2011; Blödner et al. 2007) and *Drosophila melanogaster* (e.g. Azevedo et al. 2015; Burke et al. 2010) shading light on the general processes but also specific outcomes of evolution. More recently, researchers have started to use EE approaches to provide insight into phytoplankton responses to climate change.

Marine phytoplankton contributes about 50% of annual global carbon fixation and form the basis of most marine food webs (Falkowski et al. 2008; Field 1998). Furthermore, major functional groups of phytoplankton play important roles in elemental cycles as carbon fixers, nitrogen fixers, calcifyers and silicifyers (Tréguer et al. 1995). Under current climate change, phytoplankton experience multiple, simultaneous and interacting changes (Boyd et al. 2015) in their environment, which affect primary production, community composition and biogeochemical cycles. How phytoplankton respond to climate change depends on their phenotypic plasticity (the ability of a single genotype to produce multiple phenotypes) and adaptation (via novel genetic changes) of organisms within populations which ultimately results in species sorting that affects the composition of the phytoplankton community (Fig. 1). Quantifying the role each response plays in phytoplankton adaptation to climate change is an increasingly important field of research in experimental evolution.

The ability of phytoplankton to respond to climate change on the level of an individual cell can be largely considered by taking into account their genetic repertoire that underpins the initial response to a change in environmental conditions and therefore the degree of phenotypic plasticity. Larger genomes contain a higher diversity and redundancy of genes and therefore enable the expression of multiple phenotypes from a single genotype (high phenotypic diversity)

DOI: 10.1127/pip/2016/0036

6.4.3 A cell wall protein underlies cell size in centric diatoms

This article is currently under review at ISME. Parts of the results presented in chapter 3 and 4 are included in this paper.

A cell wall protein underlies cell size in centric diatoms

Amy R Kirkham[1], Patrick Richthammer[2,†], Katrin Schmidt[1], René Hedrich[2], Eike Brunner[2], Karl-Heinz van Pée[2], Angela Falciatore[3,4], Thomas Mock[1*]

[1] School of Environmental Sciences, University of East Anglia, Norwich Research Park, Norwich, UK

[2] Allgemeine Biochemie, TU Dresden, 01062 Dresden, Germany

[3] Université Pierre et Marie Curie, UMR7238, CNRS-UPMC, Paris, France

[4] CNRS, UMR7238, Laboratory of Computational and Quantitative Biology.

[†] Present address: Helmholtz Zentrum München, Comprehensive Molecular Analytics, Ingolstädter Landstraße 1, 85764 Neuherberg, Germany

* Corresponding author

T.Mock@uea.ac.uk; +44 (0)1603 59 2566

The authors declare no conflict of interest.

Abstract

The size of an organism is the consequence of evolutionary adaptation. However, mechanisms of cell size regulation are not well understood in any cell type, and most known mechanisms relate to the timing of cell division to determine cell length, leaving the control of cell diameter unexplained. Here, a combination of reverse genetics, experimental evolution and RNA-sequencing enabled us to identify that a cell wall protein (Silacidin) determines the cell diameter of diverse and ecologically significant centric diatoms. In phytoplankton, size underlies ecological functioning through its influence on nutrient uptake capability and sinking rates as a function of surface area to volume ratio. The regulation of this gene depends on nutrient availability and temperature in the model diatom *Thalassiosira pseudonana*, creating larger cell diameters by down-regulating the gene in selection experiments under low temperature and increased nutrients. The association of cell size with these physical parameters have long been predicted by paleo and comparative ecosystem studies of phytoplankton communities.

Chapter 7

References

Adams, G.L., Pichler, D.E., Cox, E.J., O'Gorman, E.J., Seeney, A., Woodward, G., and Reuman, D.C. (2013). Diatoms can be an important exception to temperature-size rules at species and community levels of organization. *Global Change Biology* **19(11)**: 3540–3552.

Allen, A.E., Dupont, C.L., Oborník, M., Horák, A., Nunes-Nesi, A., McCrow, J.P., Zheng, H., Johnson, D.A., Hu, H., Fernie, A.R., and Bowler, C. (2011). Evolution and metabolic significance of the urea cycle in photosynthetic diatoms. *Nature* **473**: 203–207.

Allen, A.E., Vardi, A., and Bowler, C. (2006). An ecological and evolutionary context for integrated nitrogen metabolism and related signaling pathways in marine diatoms. *Current Opinion in Plant Biology* **9(3)**: 264–273.

Allen, A.P. and Gillooly, J.F. (2009). Towards an integration of ecological stoichiometry and the metabolic theory of ecology to better understand nutrient cycling. *Ecology Letters* **12(5)**: 369–384.

Alverson, A.J., Beszteri, B., Julius, M.L., and Theriot, E.C. (2011). The model marine diatom Thalassiosira pseudonana likely descended from a freshwater ancestor in the genus Cyclotella. *BMC evolutionary biology* **11(1)**: 125.

Anders, S., Pyl, P.T., and Huber, W. (2014). HTSeqa Python framework to work with high-throughput sequencing data. *Bioinformatics* p. btu638.

Anderson, D.R., Burnham, K.P., and White, G.C. (1998). Comparison of Akaike information criterion and consistent Akaike information criterion for model selection and statistical inference from capture - recapture studies. *Journal of Applied Statistics* **25(2)**: 263–283.

Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. *Reference Source* .

Anning, T., Harris, G., and Geider, R. (2001). Thermal acclimation in the marine diatom Chaetoceros calcitrans (Bacillariophyceae). *European Journal of Phycology* **36(3)**: 233–241.

Apweiler, R., Attwood, T.K., Bairoch, A., Bateman, A., Birney, E., Biswas, M., Bucher, P., Cerutti, L., Corpet, F., Croning, M.D., Durbin, R., Falquet, L., Fleischmann, W., Gouzy, J., Hermjakob, H., Hulo, N., Jonassen, I., Kahn, D., Kanapin, A., Karavidopoulou, Y., Lopez, R., Marx, B., Mulder, N.J., Oinn, T.M., Pagni, M., Servant, F., Sigrist, C.J., and Zdobnov, E.M. (2001). The InterPro database, an integrated documentation resource for protein families, domains and functional sites. *Nucleic acids research* **29(1)**: 37–40.

Armbrust, E.V. (2009). The life of diatoms in the world's oceans. *Nature* **459(7244)**: 185–192.

Armbrust, E.V., Berges, J.a., Bowler, C., Green, B.R., Martinez, D., Putnam, N.H., Zhou, S., Allen, A.E., Apt, K.E., Bechner, M., Brzezinski, M.a., Chaal, B.K., Chiovitti, A., Davis, A.K., Demarest, M.S., Detter, J.C., Glavina, T., Goodstein, D., Hadi, M.Z., Hellsten, U., Hildebrand, M., Jenkins, B.D., Jurka, J., Kapitonov, V.V., Kröger, N., Lau, W.W.Y., Lane, T.W., Larimer, F.W., Lippmeier, J.C., Lucas, S., Medina, M., Montsant, A., Obornik, M., Parker, M.S., Palenik, B., Pazour, G.J., Richardson, P.M., Rynearson, T.a., Saito, M.a., Schwartz, D.C., Thamatrakoln, K., Valentin, K., Vardi, A., Wilkerson, F.P., and Rokhsar, D.S. (2004). The genome of the diatom Thalassiosira pseudonana: ecology, evolution, and metabolism. *Science (New York, N.Y.)* **306(5693)**: 79–86.

Armbrust, E.V. and Galindo, H.M. (2001). Rapid Evolution of a Sexual Reproduction Gene in Centric Diatoms of the Genus Thalassiosira. *Applied and Environmental Microbiology* **67(8)**: 3501–3513.

Armstrong, R.A., Lee, C., Hedges, J.I., Honjo, S., and Wakeham, S.G. (2002). A new, mechanistic model for organic carbon fluxes in the ocean based on the quantitative association of POC with ballast minerals. *Deep-Sea Research Part II: Topical Studies in Oceanography* **49(1-3)**: 219–236.

Arrigo, K.R., Mills, M.M., Kropuenske, L.R., van Dijken, G.L., Alderkamp, A.C., and Robinson, D.H. (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)**: 950–966.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M., Sherlock, G., and Consortium, G.O. (2000). Gene Ontology: Tool for The Unification of Biology. *Nature Genetics* **25(1)**: 25–29.

Atkinson, A., Siegel, V., Pakhomov, E., and Rothery, P. (2004). Long-term decline in krill stock and increase in salps within the Southern Ocean **10**: 1–6.

Atkinson, D. (1994). Temperature and organisms size - A biological law for ectotherms? *Advances in Ecological Research* **25**: 1–54.

Atkinson, D., Ciotti, B.J., and Montagnes, D.J.S. (2003). Protists decrease in size linearly with temperature: ca. 2.5% degrees C-1. *Proceedings. Biological sciences / The Royal Society* **270(1533)**: 2605–11.

Azevedo, R.B.R., French, V., Partridge, L., Azevedo, R.B.R., French, I.V., and Partridge, L. (2015). Thermal Evolution of Egg Size in Drosophila melanogaster. *Evolution* **50(6)**: 2338–2345.

Bantinaki, E., Kassen, R., Knight, C.G., Robinson, Z., Spiers, A.J., and Rainey, P.B. (2007). Adaptive divergence in experimental populations of Pseudomonas fluorescens. III. Mutational origins of wrinkly spreader diversity. *Genetics* **176(1)**: 441–453.

Barton, A.D., Irwin, A.J., Finkel, Z.V., and Stock, C.A. (2016). Anthropogenic climate change drives shift and shuffle in North Atlantic phytoplankton communities. *Proceedings of the National Academy of Sciences* **7**: 201519080.

Barton, A.D., Lozier, M.S., and Williams, R.G. (2015). Physical controls of variability in north Atlantic phytoplankton communities. *Limnology and Oceanography* **60(1)**: 181–197.

Behrenfeld, M.J., O'Malley, R.T., Siegel, D.a., McClain, C.R., Sarmiento, J.L., Feldman, G.C., Milligan, A.J., Falkowski, P.G., Letelier, R.M., and Boss, E.S. (2006). Climate-driven trends in contemporary ocean productivity. *Nature* **444(7120)**: 752–755.

Behrenfeld, M.J., Randerson, J.T., McClain, C.R., Feldman, G.C., Los, S.O., Tucker, C.J., Falkowski, P.G., Field, C.B., Frouin, R., Esaias, W.E., Kolber, D.D., and Pollack, N.H. (2001). Biospheric primary production during an ENSO transition. *Science* **291(5513)**: 2594–2597.

Bell, G. (1997*a*). *The basics of selection*. Chapman & Hall, New York.

Bell, G.A.C. (1997*b*). Experimental evolution in Chlamydomonas. I . Short-term selection in uniform and diverse environments. *Heredity* **78**: 490–497.

Benner, I., Diner, R.E., Lefebvre, S.C., Li, D., Komada, T., Carpenter, E.J., and Stillman, J.H. (2013). Emiliania huxleyi increases calcification but not expression of calcification-related genes in long-term exposure to elevated temperature and pCO2. *Philosophical Transactions of the Royal Society B: Biological Sciences* **368(1627)**.

Bennett, A.F. and Lenski, R.E. (1993). Evolutionary Adaptation to Temperature II. Thermal Niches of Experimental Lines of Escherichia coli. *Evolution* **47(1)**: 1–12.

Bentkowski, P., Oosterhout, C.V., and Mock, T. (2015). A model of genome size evolution for prokaryotes in stable and fluctuating environments. *Genome biology and evolution* **7(8)**: 2344–2351.

Berges, J.A., Varela, D.E., and Harrison, P.J. (2002). Effects of temperature on growth rate, cell composition and nitrogen metabolism in the marine diatom Thalassiosira pseudonana (Bacillariophyceae). *Marine Ecology Progress Series* **225**: 139–146.

Bermúdez, R., Feng, Y., Roleda, M.Y., Tatters, A.O., Hutchins, D.A., Larsen, T., Boyd, P.W., Hurd, C.L., Riebesell, U., and Winder, M. (2015). Long-term conditioning to elevated pCO2 and warming influences the fatty and amino acid composition of the diatom Cylindrotheca fusiformis. *PLoS ONE* **10(5)**: 1–15.

Berry, J.A. and Raison, J.K. (1981). *Responses of Macrophytes to Temperature*. pp. 277–338. Springer Berlin Heidelberg, Berlin, Heidelberg.

Biémont, C. and Vieira, C. (2006). Genetics: junk DNA as an evolutionary force. *Nature* **443(7111)**: 521–524.

Bird, A. (2007). Perceptions of epigenetics. *Nature* **447(7143)**: 396–398.

Blake, J.A., Christie, K.R., Dolan, M.E., Drabkin, H.J., Hill, D.P., Ni, L., Sitnikov, D., Burgess, S., Buza, T., Gresham, C., McCarthy, F., Pillai, L., Wang, H., Carbon, S., Dietze, H., Lewis, S.E., Mungall, C.J., Munoz-Torres, M.C., Feuermann, M., Gaudet, P., Basu, S., Chisholm, R.L., Dodson, R.J., Fey, P., Mi, H., Thomas, P.D., Muruganujan, A., Poudel, S., Hu, J.C., Aleksander, S.A., McIntosh, B.K., Renfro, D.P., Siegele, D.A., Attrill, H., Brown, N.H., Tweedie, S., Lomax, J., Osumi-Sutherland, D., Parkinson, H., Roncaglia, P., Lovering, R.C., Talmud, P.J., Humphries, S.E., Denny, P., Campbell, N.H., Foulger, R.E., Chibucos, M.C., Giglio, M.G., Chang, H.Y., Finn, R., Fraser, M., Mitchell, A., Nuka, G., Pesseat, S., Sangrador, A.,

Scheremetjew, M., Young, S.Y., Stephan, R., Harris, M.A., Oliver, S.G., Rutherford, K., Wood, V., Bahler, J., Lock, A., Kersey, P.J., McDowall, M.D., Staines, D.M., Dwinell, M., Shimoyama, M., Lauledeerkind, S., Hayman, G.T., Wang, S.J., Petri, V., D'Eustachio, P., Matthews, L., Balakrishnan, R., Binkley, G., Cherry, J.M., Costanzo, M.C., Demeter, J., Dwight, S.S., Engel, S.R., Hitz, B.C., Inglis, D.O., Lloyd, P., Miyasato, S.R., Paskov, K., Roe, G., Simison, M., Nash, R.S., Skrzypek, M.S., Weng, S., Wong, E.D., Berardini, T.Z., Li, D., Huala, E., Argasinska, J., Arighi, C., Auchincloss, A., Axelsen, K., Argoud-Puy, G., Bateman, A., Bely, B., Blatter, M.C., Bonilla, C., Bougueleret, L., Boutet, E., Breuza, L., Bridge, A., Britto, R., Casals, C., Cibrian-Uhalte, E., Coudert, E., Cusin, I., Duek-Roggli, P., Estreicher, A., Famiglietti, L., Gane, P., Garmiri, P., Gos, A., Gruaz-Gumowski, N., Hatton-Ellis, E., Hinz, U., Hulo, C., Huntley, R., Jungo, F., Keller, G., Laiho, K., Lemercier, P., Lieberherr, D., Macdougall, A., Magrane, M., Martin, M., Masson, P., Mutowo, P., O'Donovan, C., Pedruzzi, I., Pichler, K., Poggioli, D., Poux, S., Rivoire, C., Roechert, B., Sawford, T., Schneider, M., Shypitsyna, A., Stutz, A., Sundaram, S., Tognolli, M., Wu, C., Xenarios, I., Chan, J., Kishore, R., Sternberg, P.W., Van Auken, K., Muller, H.M., Done, J., Li, Y., Howe, D., and Westerfeld, M. (2015). Gene ontology consortium: Going forward. *Nucleic Acids Research* **43(D1)**: D1049–D1056.

Blödner, C., Goebel, C., Feussner, I., Gatz, C., and Polle, A. (2007). Warm and cold parental reproductive environments affect seed properties, fitness, and cold responsiveness in Arabidopsis thaliana progenies. *Plant, Cell & Environment* **30(2)**: 165–175.

Blount, Z.D., Barrick, J.E., Davidson, C.J., and Lenski, R.E. (2012). Genomic analysis of a key innovation in an experimental Escherichia coli population. *Nature* **489(7417)**: 513–8.

Blount, Z.D., Borland, C.Z., and Lenski, R.E. (2008). Historical contingency and the evolution of a key innovation in an experimental population of Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America* **105(23)**: 7899–906.

Bluhm, B., Gebruk, A., Gradinger, R., Hopcroft, R., Huettmann, F., Kosobokova, K., Sirenko, B., and Weslawski, M. (2011). Arctic Marine Biodiversity: An Update of Species Richness and Examples of Biodiversity Change. *Oceanography* **24**: 232–248.

Boetius, A., Albrecht, S., Bakker, K., Bienhold, C., Felden, J., Fernández-Méndez, M., Hendricks, S., Katlein, C., Lalande, C., Krumpen, T., Nicolaus, M., Peeken, I., Rabe, B., Rogacheva, A., Rybakova, E., Somavilla, R., and Wenzhöfer, F. (2013). Export of algal biomass from the melting Arctic sea ice. *Science (New York, N.Y.)* **339(6126)**: 1430–2.

Bond, D.M. and Baulcombe, D.C. (2015). Epigenetic transitions leading to heritable, RNA-mediated de novo silencing in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences* **2014(13)**: 201413053.

Booth, B.C. and Horner, R.A. (1997). Microalgae on the Arctic ocean section, 1994: Species abundance and biomass. *Deep-Sea Research Part II: Topical Studies in Oceanography* **44(8)**: 1607–1622.

Bossdorf, O., Arcuri, D., Richards, C.L., and Pigliucci, M. (2010). Experimental alteration of DNA methylation affects the phenotypic plasticity of ecologically relevant traits in Arabidopsis thaliana. *Evolutionary Ecology* **24(3)**: 541–553.

Bossdorf, O., Richards, C.L., and Pigliucci, M. (2008). Epigenetics for ecologists. *Ecology Letters* **11(2)**: 106–115.

Bowler, C., Allen, A.E., Badger, J.H., Grimwood, J., Jabbari, K., Kuo, A., Maheswari, U., Martens, C., Maumus, F., Otillar, R.P., Rayko, E., Salamov, A., Vandepoele, K., Beszteri, B., Gruber, A., Heijde, M., Katinka, M., Mock, T., Valentin, K., Verret, F., Berges, J.a., Brownlee, C., Cadoret, J.P., Chiovitti, A., Choi, C.J., Coesel, S., De Martino, A., Detter, J.C., Durkin, C., Falciatore, A., Fournet, J., Haruta, M., Huysman, M.J.J., Jenkins, B.D., Jiroutova, K., Jorgensen, R.E., Joubert, Y., Kaplan, A., Kröger, N., Kroth, P.G., La Roche, J., Lindquist, E., Lommer, M., Martin-Jézéquel, V., Lopez, P.J., Lucas, S., Mangogna, M., McGinnis, K., Medlin, L.K., Montsant, A., Oudot-Le Secq, M.P., Napoli, C., Obornik, M., Parker, M.S., Petit, J.L., Porcel, B.M., Poulsen, N., Robison, M., Rychlewski, L., Rynearson, T.a., Schmutz, J., Shapiro, H., Siaut, M., Stanley, M., Sussman, M.R., Taylor, A.R., Vardi, A., von Dassow, P., Vyverman, W., Willis, A., Wyrwicz, L.S., Rokhsar, D.S., Weissenbach, J., Armbrust, E.V., Green, B.R., Van de Peer, Y., and Grigoriev, I.V. (2008). The Phaeodactylum genome reveals the evolutionary history of diatom genomes. *Nature* **456(7219)**: 239–44.

Bowler, C., Vardi, A., and Allen, A.E. (2010). Oceanographic and biogeochemical insights from diatom genomes. *Annual Review of Marine Science* **2**: 333–365.

Boyd, P.W. (2011). Beyond ocean acidification. *Nature Geoscience* **4(5)**: 273–274.

Boyd, P.W., Dillingham, P.W., McGraw, C.M., Armstrong, E.A., Cornwall, C.E., Feng, Y.y., Hurd, C.L., Gault-Ringold, M., Roleda, M.Y., Timmins-Schiffman, E., and Nunn, B.L. (2015*a*). Physiological responses of a Southern Ocean diatom to complex future ocean conditions. *Nature Climate Change* **6**: 207–216.

Boyd, P.W., Lennartz, S.T., Glover, D.M., and Doney, S.C. (2015*b*). Biological ramifications of climate-change-mediated oceanic multi-stressors. *Nature Climate Change* **(5)**: 71–79.

Boyd, P.W., Rynearson, T.A., Armstrong, E.A., Fu, F., Hayashi, K., Hu, Z., Hutchins, D.A., Kudela, R.M., Litchman, E., Mulholland, M.R., Passow, U., Strzepek, R.F., Whittaker, K.A., Yu, E., and Thomas, M.K. (2013). Marine Phytoplankton Temperature versus Growth Responses from Polar to Tropical Waters Outcome of a Scientific Community-Wide Study. *PLoS ONE* **8(5)**: e63091.

Boyko, A. and Kovalchuk, I. (2011). Genome instability and epigenetic modification-heritable responses to environmental stress? *Current opinion in plant biology* **14(3)**: 260–6.

Bray, J. and Curtis, J. (1957). An Ordination of the upland forest community of southern Wisconsin.

Brody, S.R. and Lozier, M.S. (2015). Characterizing upper-ocean mixing and its effect on the spring phytoplankton bloom with in situ data. *ICES Journal of Marine Science: Journal du Conseil* **72(6)**: 1961–1970.

Brzezinski, M.A. (1985). The Si:C:N ratio of marine diatoms: Interspecific variability and the effect of some environmental variables. *Journal of Phycology* .

Burke, M.K., Dunham, J.P., Shahrestani, P., Thornton, K.R., Rose, M.R., and Long, A.D. (2010). Genome-wide analysis of a long-term evolution experiment with Drosophila. *Nature* **467(7315)**: 587–590.

Caldeira, K. and Wickett, M.E. (2003). Anthropogenic carbon and ocean pH. *Nature* **425**: 365.

Campbell, D., Zhou, G., Gustafsson, P., Oquist, G., and Clarke, a.K. (1995). Electron transport regulates exchange of two forms of photosystem II D1 protein in the cyanobacterium Synechococcus. *The EMBO journal* **14(22)**: 5457–5466.

Carpenter, E.J., Montoya, J.P., Burns, J., Mulholland, M.R., Subramaniam, A., and Capone, D.G. (1999). Extensive bloom of a N2-fixing diatom/cyanobacterial association in the tropical Atlantic Ocean. *Marine Ecology Progress Series* **185(1977)**: 273–283.

Casacuberta, E. and González, J. (2013). The impact of transposable elements in environmental adaptation. *Molecular Ecology* **22(6)**: 1503–1517.

Caspeta, L., Chen, Y., Ghiaci, P., Feizi, A., Buskov, S., Hallström, B.M., Petranovic, D., and Nielsen, J. (2014). Biofuels. Altered sterol composition renders yeast thermotolerant. *Science (New York, N.Y.)* **346(6205)**: 75–8.

Caspeta, L. and Nielsen, J. (2015). Thermotolerant Yeast Strains Adapted by Laboratory Evolution Show Trade-Off at Ancestral Temperatures and Preadaptation to Other. *mBio* **6(4)**: 1–9.

Cavrak, V.V., Lettner, N., Jamge, S., Kosarewicz, A., Bayer, L.M., and Mittelsten Scheid, O. (2014). How a Retrotransposon Exploits the Plant's Heat Stress Response for Its Activation. *PLoS Genetics* **10(1)**: e1004115.

Cermeño, P., de Vargas, C., Abrantes, F., and Falkowski, P.G. (2010). Phytoplankton biogeography and community stability in the ocean. *PLoS ONE* **5(4)**.

Cermeno, P., Dutkiewicz, S., Harris, R.P., Follows, M., Schofield, O., and Falkowski, P.G. (2008). The role of nutricline depth in regulating the ocean carbon cycle. *Proceedings of the National Academy of Sciences* **105(51)**: 20344–20349.

Chaal, B.K. and Green, B.R. (2005). Protein import pathways in 'complex' chloroplasts derived from secondary endosymbiosis involving a red algal ancestor. *Plant Molecular Biology* **57(3)**: 333–342.

Chan, C.X., Baglivi, F.L., Jenkins, C.E., and Bhattacharya, D. (2013). Foreign gene recruitment to the fatty acid biosynthesis pathway in diatoms. *Mobile genetic elements* **3(5)**: e27313.

Chang, R.L., Andrews, K., Kim, D., Li, Z., Godzik, A., and Palsson, B.O. (2013). Structural Systems Biology Evaluation of Metabolic Thermotolerance in Escherichia coli. *Science (New York, N.Y.)* **340(6137)**: 1220–1223.

Charmantier, A., McCleery, R.H., Cole, L.R., Perrins, C., Kruuk, L.E.B., and Sheldon, B.C. (2008). Adaptive phenotypic plasticity in response to climate change in a wild bird population. *Science* **320(5877)**: 800–803.

Chepurnov, V.A., Mann, D.G., Sabbe, K., and Vyverman, W. (2004). Experimental studies on sexual reproduction in diatoms. *International review of cytology* **237**: 91–154.

Chevin, L.M., Lande, R., and Mace, G.M. (2010). Adaptation, Plasticity, and Extinction in a Changing Environment: Towards a Predictive Theory. *PLoS Biol* **8(4)**: e1000357.

Chisholm, S.W. (1992). *Primary Productivity and Biogeochemical Cycles in the Sea*. chap. Phytoplank, pp. 213–237. Springer US, Boston, MA.

Chisholm, S.W. (2000). Stirring times in the Southern Ocean. *Nature* **407(6805)**: 685–687.

Chust, G., Irigoien, X., Chave, J., and Harris, R.P. (2013). Latitudinal phytoplankton distribution and the neutral theory of biodiversity. *Global Ecology and Biogeography* **22(5)**: 531–543.

Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118 ; iso-2; iso-3. *Landes Bioscience* **1(2)**: 152–161.

Coesel, S., Oborník, M., Varela, J., Falciatore, A., and Bowler, C. (2008). Evolutionary origins and functions of the carotenoid biosynthetic pathway in marine diatoms. *PLoS ONE* **3(8)**.

Collins, S. (2011). Competition limits adaptation and productivity in a photosynthetic alga at elevated CO2. *Proceedings of the Royal Society B: Biological Sciences* **278(1703)**: 247–255.

Collins, S. (2013). New model systems for experimental evolution. *Evolution* **67(7)**: 1847–1848.

Collins, S. and Bell, G. (2006). Evolution of natural algal populations at elevated CO2. *Ecology Letters* **9(2)**: 129–135.

Collins, S. and De Meaux, J. (2009). Adaptation to different rates of environmental change in Chlamydomonas. *Evolution* **63(11)**: 2952–2965.

Conrad, T.M., Lewis, N.E., and Palsson, B.Ø. (2011). Microbial laboratory evolution in the era of genome-scale science. *Molecular systems biology* **7(509)**: 509.

Cooper, T.F., Rozen, D.E., and Lenski, R.E. (2003). Parallel changes in gene expression after 20,000 generations of evolution in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America* **100(3)**: 1072–1077.

Daines, S.J., Clark, J.R., and Lenton, T.M. (2014). Multiple environmental controls on phytoplankton growth strategies determine adaptive responses of the N:P ratio. *Ecology Letters* **17(4)**: 414–425.

Dallinger, W.H. (1887). The President's Address. *Royal Microscopical Society* **7**: 184–199.

Davison, I.R. (1991). Environmental effects on algal photosynthesis: Temperature. *Journal of Phycology* **27**: 2–8.

De Mazancourt, C., Johnson, E., and Barraclough, T.G. (2008). Biodiversity inhibits species' evolutionary responses to changing environments. *Ecology Letters* **11(4)**: 380–388.

De Riso, V., Raniello, R., Maumus, F., Rogato, A., Bowler, C., and Falciatore, A. (2009). Gene silencing in the marine diatom Phaeodactylum tricornutum. *Nucleic Acids Research* **37(14)**: e96.

De Senerpont Domis, L.N., Mooij, W.M., and Huisman, J. (2007). Climate-induced shifts in an experimental phytoplankton community: a mechanistic approach. *Hydrobiologia* **584(1)**: 403–413.

De Senerpont Domis, L.N., Waal, D.B.V.D., Helmsing, N.R., van Donk, E., and Mooij, W.M. (2014). Community stoichiometry in a changing world : combined effects of warming and eutrophication on phytoplankton dynamics. *Ecology* **95(6)**: 1485–1495.

De Wit, P., Pespeni, M.H., Ladner, J.T., Barshis, D.J., Seneca, F., Jaris, H., Therkildsen, N.O., Morikawa, M., and Palumbi, S.R. (2012). The simple fool's guide to population genomics via RNA-Seq: An introduction to high-throughput sequencing data analysis. *Molecular Ecology Resources* **12(6)**: 1058–1067.

Debiasse, M.B. and Kelly, M.W. (2016). Plastic and Evolved Responses to Global Change : What Can We Learn from Comparative Transcriptomics ? *Journal of Heredity* **107(1)**: 71–81.

Delong, E.F. (2009). The Microbial Ocean from Genomes to Biomes. *Nature* **459(7244)**: 200–2006.

Deschamps, P. and Moreira, D. (2012). Reevaluating the green contribution to diatom genomes. *Genome biology and evolution* **4(7)**: 683–8.

Dhar, R., Sägesser, R., Weikert, C., Yuan, J., and Wagner, A. (2011). Adaptation of Saccharomyces cerevisiae to saline stress through laboratory evolution. *Journal of evolutionary biology* **24(5)**: 1135–53.

Doney, S.C., Abbott, M.R., Cullen, J.J., Karl, D.M., and Rothstein, L. (2004). From genes to ecosystems : the ocean ' s new frontier In a nutshell :. *Frontiers in Ecology and the Environment* **2(9)**: 457–466.

Doney, S.C., Fabry, V.J., Feely, R.A., and Kleypas, J.A. (2009). Ocean Acidification: The Other CO 2Problem. *Annual Review of Marine Science* **1(1)**: 169–192.

Doney, S.C., Ruckelshaus, M., Duffy, J.E., Barry, J.P., Chan, F., English, C.A., Galindo, H.M., Grebmeier, J.M., Hollowed, A.B., Knowlton, N., Polovina, J., Rabalais, N.N., Sydeman, W.J., and Talley, L.D. (2012). Climate change impacts on marine ecosystems. *Ann Rev Mar Sci* **4**: 11–37.

Draghi, J.a. and Whitlock, M.C. (2012). Phenotypic plasticity facilitates mutational variance, genetic variance, and evolvability along the major axis of environmental variation. *Evolution* **66(9)**: 2891–902.

Dubacq, J.P. and Tremolieres, A. (1983). Occurrence and function of phosphatidylglycerol containing delta 3-trans-hexadecenoic acid in photosynthetic lamellae. *Physiologie vegetale* .

Duncan, E.J., Gluckman, P.D., and Dearden, P.K. (2014). Epigenetics, plasticity, and evolution: How do we link epigenetic change to phenotype? *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* **322(4)**: 208–220.

Durselen, C.D. and Rick, H.J. (1999). Spatial and temporal distribution of two new phytoplankton diatom species in the German Bight in the period 1988 and 1996. *Sarsia* **84(5-6)**: 367–377.

Dyhrman, S.T., Jenkins, B.D., Rynearson, T.A., Saito, M.A., Mercier, M.L., Alexander, H., Whitney, L.P., Drzewianowski, A., Bulygin, V.V., Bertrand, E.M., Wu, Z., Benitez-Nelson, C., and Heithoff, A. (2012). The Transcriptome and Proteome of the Diatom Thalassiosira pseudonana Reveal a Diverse Phosphorus Stress Response. *PLoS ONE* **7(3)**: e33768.

Elser, J.J., Dobberfuhl, D.R., Mackay, N.a., and Schampel, J.H. (1996). Size , and Life Stoichiomet Toward a unified view of cellular and ecosystem processes. *BioScience* **46(9)**: 674–684.

Elser, J.J., Sterner, R.W., Gorokhova, E., Fagan, W.F., Markow, T.A., Cotner, J.B., Harrison, J.F., Hobbie, S.E., Odell, G.M., and Weider, L.J. (2000). Biological Ecosystems From Genes To Ecosystems. *Ecology Letters* **3**: 540–550.

Ensminger, I., Busch, F., and Huner, N.P.A. (2006). Photostasis and cold acclimation: Sensing low temperature through photosynthesis. *Physiologia Plantarum* **126(1)**: 28–44.

Eppley, R.W. (1972). Temperature and phytoplankton growth in the sea. *Fishery Bulletin* **70(4)**: 1063–1085.

Fablet, M. and Vieira, C. (2011). Evolvability, epigenetics and transposable elements.

Falcone, D.L., Ogas, J.P., and Somerville, C.R. (2004). Regulation of membrane fatty acid composition by temperature in mutants of Arabidopsis with alterations in membrane lipid composition. *BMC plant biology* **4**: 17.

Falkowski, P.G. (2004). The Evolution of Modern Eukaryotic Phytoplankton. *Science* **305(5682)**: 354–360.

Falkowski, P.G., Fenchel, T., and Delong, E.F. (2008). The microbial engines that drive Earth's biogeochemical cycles. *Science* **320(5879)**: 1034–9.

Feely, R.a., Sabine, C.L., Lee, K., Berelson, W., Kleypas, J., Fabry, V.J., Millero, F.J., and Anonymous (2004). Impact of anthropogenic CO2 on the CaCO3 system in the oceans. *Science* **305(5682)**: 362–366.

Feller, U., Crafts-Brandner, S.J., and Salvucci, M.E. (1998). Moderately High Temperatures Inhibit Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase (Rubisco) Activase-Mediated Activation of Rubisco. *Plant Physiol* **116(2)**: 539–546.

Feng, Y., Hare, C.E., Leblanc, K., Rose, J.M., Zhang, Y., DiTullio, G.R., Lee, P., Wilhelm, S., Rowe, J.M., and Sun, J. (2009). The Effects of Increased pCO2 and Temperature on the North Atlantic Spring Bloom: I. The Phytoplankton Community and Biogeochemical Response. *Marine Ecology Progress Series* **388**: 13–25.

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

Finkel, Z.V., Beardall, J., Flynn, K.J., Quigg, A., Rees, T.A.V., and Raven, J.A. (2010). Phytoplankton in a changing world: cell size and elemental stoichiometry. *Journal of Plankton Research* **32(1)**: 119–137.

Finn, R.D., Attwood, T.K., Babbitt, P.C., Bateman, A., Bork, P., Bridge, A.J., Chang, H.Y., Dosztányi, Z., El-Gebali, S., Fraser, M., Gough, J., Haft, D., Holliday, G.L., Huang, H., Huang, X., Letunic, I., Lopez, R., Lu, S., Marchler-Bauer, A., Mi, H., Mistry, J., Natale, D.A., Necci, M., Nuka, G., Orengo, C.A., Park, Y., Pesseat, S., Piovesan, D., Potter, S.C., Rawlings, N.D., Redaschi, N., Richardson, L., Rivoire, C., Sangrador-Vegas, A., Sigrist, C., Sillitoe, I., Smithers, B., Squizzato, S., Sutton, G., Thanki, N., Thomas, P.D., Tosatto, S.C.E., Wu, C.H., Xenarios, I., Yeh, L.S., Young, S.Y., and Mitchell, A.L. (2016). InterPro in 2017-beyond protein family and domain annotations. *Nucleic acids research* **45(November 2016)**: gkw1107.

Fisher, R.A. (1930). *The Genetical Theory of Natural Selection*. Oxford University Press 1999. oxford uni ed.

Flores, K.B., Wolschin, F., and Amdam, G.V. (2013). The role of methylation of DNA in environmental adaptation. *Integrative and Comparative Biology* **53(2)**: 359–372.

Flynn, K.J., Raven, J.A., Rees, T.A.V., Finkel, Z., Quigg, A., and Beardall, J. (2010). Is the growth hypothesis applicable to microalgae? *Journal of Phycology* **46(1)**: 1–12.

Follows, M.J., Dutkiewicz, S., Grant, S., and Chisholm, S.W. (2007). Emergent Biogeography of Microbial Communities in a Model Ocean. *Science (New York, N.Y.)* **315**: 1843–1846.

Fong, S.S., Joyce, A.R., and Palsson, B.Ø. (2005). Parallel adaptive evolution cultures of Escherichia coli lead to convergent growth phenotypes with different gene expression states. *Genome Research* **15(10)**: 1365–1372.

Fragoso, G.M., Poulton, A.J., Yashayaev, I.M., Head, E.J., Stinchcombe, M., and Purdie, D.A. (2016). Biogeographical patterns and environmental controls of phytoplankton communities from contrasting hydrographical zones of the Labrador Sea. *Progress in Oceanography* **141(January)**: 212–226.

Franssen, S.U., Gu, J., Winters, G., Huylmans, A.K., Wienpahl, I., Sparwel, M., Coyer, J.A., Olsen, J.L., Reusch, T.B., and Bornberg-Bauer, E. (2014). Genome-wide transcriptomic responses of the seagrasses Zostera marina and Nanozostera noltii under a simulated heatwave confirm functional types. *Marine Genomics* **15**: 65–73.

French, F.W. and Hargraves, P.E. (1985). Spore formation in the life cycles of the diatoms Chaetoceros diadema and Leptocylindrus danicus. *Journal of Phycology* **21(3)**: 477–483.

G. Falkowski, P. (1977). The adenylate energy charge in marine phytoplantkon: The effect of temperature on the physiological state of Skeletonema costatum (Grev.) Cleve. *Journal of Experimental Marine Biology and Ecology* **27(1)**: 37–45.

Galderisi, U., Jori, F.P., and Giordano, A. (2003). Cell cycle regulation and neural differentiation. *Oncogene* **22(33)**: 5208–5219.

Gallagher, J. (1983). Cell Enlargement in Skeletonema Costatum (Bacillariophyceae).

Geider, R. and La Roche, J. (2002). Redfield revisited: variability of C:N:P in marine microalgae and its biochemical basis. *European Journal of Phycology* **37(1)**: 1–17.

Geider, R.J. (1987). Light and temperature dependence of the carbon to chlorophyll a ratio in microalgae and cyanobacteria: Implications for physiology and growth of phytoplankton. *New Phytologist* **106**: 1–34.

Geider, R.J., MacIntyre, H.L., and Kana, T.M. (1998). A dynamic regulatory model of phytoplanktonic acclimation to light, nutrients, and temperature. *Limnology and Oceanography* **43(4)**: 679–694.

Gillooly, J.F., Allen, A.P., Brown, J.H., Elser, J.J., Martinez del Rio, C., Savage, V.M., West, G.B., Woodruff, W.H., and Woods, H.A. (2005). The metabolic basis of whole-organism RNA and phosphorus content. *Proceedings of the National Academy of Sciences USA* **102(33)**: 11923–7.

Gillooly, J.F., Brown, J.H., and West, G.B. (2001). Effects of Size and Temperature on Metabolic Rate. *Science* **293(September)**: 2248–2252.

Gleason, L.U. and Burton, R.S. (2015). RNA-seq reveals regional differences in transcriptome response to heat stress in the marine snail Chlorostoma funebralis. *Molecular Ecology* **24(3)**: 610–627.

Gobet, A., Böer, S.I., Huse, S.M., van Beusekom, J.E.E., Quince, C., Sogin, M.L., Boetius, A., and Ramette, A. (2012). Diversity and dynamics of rare and of resident bacterial populations in coastal sands. *The ISME Journal* **6(3)**: 542–553.

Godhe, A., Asplund, M.E., Härnström, K., Saravanan, V., Tyagi, A., and Karunasagar, I. (2008). Quantification of diatom and dinoflagellate biomasses in coastal marine seawater samples by real-time PCR. *Applied and Environmental Microbiology* **74(23)**: 7174–7182.

Gosselin, M., Levasseur, M., Wheeler, P.A., Horner, R.A., and Booth, B.C. (1997). New measurements of phytoplankton and ice algal production in the Arctic Ocean. *Deep-Sea Research Part II: Topical Studies in Oceanography* **44(8)**: 1623–1644.

Gould, S.B., Waller, R.F., and Mcfadden, G.I. (2008). Plastid evolution. *Annual Review of Plant Biology* **59**: 491–517.

Gradinger, R.R. and Baumann, M.E.M. (1991). Distribution of phytoplankton communities in relation to the large-scale hydrographical regime in the Fram Strait. *Marine Biology* **111(2)**: 311–321.

Graversen, R.G., Mauritsen, T., Tjernström, M., Källén, E., and Svensson, G. (2008). Vertical structure of recent Arctic warming. *Nature* **451(January)**: 53–56.

Gregory, T.R. (2001). Coincidence, coevolution, or causation? DNA content, cell size, and the C-value enigma. *Biological Reviews of the Cambridge Philosophical Society* **76(1)**: 65–101.

Gruber, A., Rocap, G., Kroth, P.G., Armbrust, E.V., and Mock, T. (2015). Plastid proteome prediction for diatoms and other algae with secondary plastids of the red lineage. *Plant Journal* **81(3)**: 519–528.

Guillard, R.R. and Ryther, J.H. (1962). Studies of marine planktonic diatoms. I. Cyclotella nana Hustedt, and Detonula confervacea (cleve) Gran. *Can J Microbiol.* **8**: 229–239.

Hardge, K., Kilias, E., Bauerfeind, E., Cherkasheva, A., Ga, S., Kraft, A., Kidane, Y.M., Lalande, C., Piontek, J., Thomisch, K., and Wurst, M. (2015). Summertime plankton ecology in Fram Strait a compilation of long- and short-term observations. *Polar Research* **1**: 1–18.

Hare, M.P., Nunney, L., Schwartz, M.K., Ruzzante, D.E., Burford, M., Waples, R.S., Ruegg, K., and Palstra, F. (2011). Understanding and Estimating Effective Population Size for Practical Application in Marine Species Management. *Conservation Biology* **25(3)**: 438–449.

Harris, R.S., Feng, G., Ross, K.J., Sidhu, R., Thulin, C., Longerich, S., Szigety, S.K., Winkler, M.E., and Rosenberg, S.M. (1997). Mismatch repair protein MutL becomes limiting during stationary-phase mutation. *Genes and Development* **11(18)**: 2426–2437.

Havaux, M. (1993). Rapid photosynthetic adaptation to heat stress triggered in potato leaves by moderately elevated temperatures. *Plant, Cell & Environment* **16(4)**: 461–467.

Hazel, J.R. (1995). Thermal Adaptation in Biological-Membranes - Is Homeoviscous Adaptation the Explanation. *Annual Review of Physiology* **57(94)**: 19–42.

Hegseth, E.N. and Sundfjord, A. (2008). Intrusion and blooming of Atlantic phytoplankton species in the high Arctic. *Journal of Marine Systems* **74(1-2)**: 108–119.

Hein, M., Pedersen, M.F., and Sand-jensen, K. (1995). Size-dependent nitrogen uptake in micro- and macroalgae. *Marine Ecology Progress Series* **118**: 247–253.

Heinle, M. (2013). The effects of light , temperature and nutrients on coccolithophores and implications for biogeochemical models **(September)**.

Helliwell, K.E., Collins, S., Kazamia, E., Purton, S., Wheeler, G.L., and Smith, A.G. (2015). Fundamental shift in vitamin B12 eco-physiology of a model alga demonstrated by experimental evolution. *The ISME Journal* **9(6)**: 1446–1455.

Hermann, D., Egue, F., Tastard, E., Nguyen, D.H., Casse, N., Caruso, A., Hiard, S., Marchand, J., Chénais, B., Morant-Manceau, A., and Rouault, J.D. (2014). An introduction to the vast world of transposable elements what about the diatoms? *Diatom Research* **29(1)**: 91–104.

Herrera, C.M., Pozo, M.I., and Bazaga, P. (2012). Jack of all nectars, master of most: DNA methylation and the epigenetic basis of niche width in a flower-living yeast. *Molecular Ecology* **21(11)**: 2602–2616.

Heywood, J.L., Zubkov, M.V., Tarran, G.A., Fuchs, B.M., and Holligan, P.M. (2006). Prokaryoplankton standing stocks in oligotrophic gyre and equatorial provinces of the Atlantic Ocean: Evaluation of inter-annual variability. *Deep-Sea Research Part II: Topical Studies in Oceanography* **53(14-16)**: 1530–1547.

Hinder, S.L., Hays, G.C., Edwards, M., Roberts, E.C., Walne, A.W., and Gravenor, M.B. (2012). Changes in marine dinoflagellate and diatom abundance under climate change. *Nature Clim. Change* **2(4)**: 271–275.

Ho, T.Y., Quigg, A., Finkel, Z.V., Milligan, A.J., Wyman, K., Falkowski, P.G., and Morel, F.M.M. (2003). The elemental composition of some marine phytoplankton. *Journal of Phycology* **39(6)**: 1145–1159.

Hofmann, M. and Schellnhuber, H.j. (2009). Oceanic acidification affects marine carbon pump. *Proceedings of the National Academy of Sciences* **106(9)**: 3017–3022.

Hopes, A. and Mock, T. (2015). Evolution of Microalgae and Their Adaptations in Different Marine Ecosystems. *eLS* pp. 1–9.

Horton, P., Ruban, A.V., and Young, A.J. (1999). *Regulation of the Structure and Function of the Light Harvesting Complexes of Photosystem II by the Xanthophyll Cycle*. pp. 271–291. Springer Netherlands, Dordrecht.

Horvath, G., Melis, A., Hideg, E., Droppa, M., and Vigh, L. (1987). Role of lipids in the organization and function of photosystem II studied by homogeneous catalytic hydrogenation of thylakoid membrane in situ. *Biochimica et Biophysica Acta* **891(1)**: 68–74.

Horváth, I., Glatz, A., Varvasovszki, V., Török, Z., Páli, T., Balogh, G., Kovács, E., Nádasdi, L., Benkö, S., Joó, F., and Vígh, L. (1998). Membrane physical state controls the signaling mechanism of the heat shock response in Synechocystis PCC 6803: Identification of hsp17 as a "fluidity gene". *Proceedings of the National Academy of Sciences of the United States of America* **95(7)**: 3513–3518.

Huey, R.B. and Kingsolver, J.G. (1989). Evolution of thermal sensitivity of ectotherm performance. *Trends in Ecology and Evolution* **4(5)**: 131–135.

Huner, N.P.A., Öquist, G., Hurry, V.M., Krol, M., Falk, S., and Griffith, M. (1993). Photosynthesis, photoinhibition and low temperature acclimation in cold tolerant plants. *Photosynthesis Research* **37(1)**: 19–39.

Hunter, S., Apweiler, R., Attwood, T.K., Bairoch, A., Bateman, A., Binns, D., Bork, P., Das, U., Daugherty, L., Duquenne, L., Finn, R.D., Gough, J., Haft, D., Hulo, N., Kahn, D., Kelly, E., Laugraud, A., Letunic, I., Lonsdale, D., Lopez, R., Madera, M., Maslen, J., Mcanulla, C., McDowall, J., Mistry, J., Mitchell, A., Mulder, N., Natale, D., Orengo, C., Quinn, A.F., Selengut, J.D., Sigrist, C.J.A., Thimma, M., Thomas, P.D., Valentin, F., Wilson, D., Wu, C.H., and Yeats, C. (2009). InterPro: The integrative protein signature database. *Nucleic Acids Research* **37(SUPPL. 1)**: 211–215.

Hutchins, D.a., Walworth, N.G., Webb, E.a., Saito, M.a., Moran, D., McIlvin, M.R., Gale, J., and Fu, F.X. (2015). Irreversibly increased nitrogen fixation in Trichodesmium experimentally adapted to elevated carbon dioxide. *Nature Communications* **6(SEPTEMBER)**: 8155.

IPCC (2014). *Summary for Policymakers*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Iversen, M.H., Nowald, N., Ploug, H., Jackson, G.A., and Fischer, G. (2010). High resolution profiles of vertical particulate organic matter export off Cape Blanc, Mauritania: Degradation processes and ballasting effects. *Deep-Sea Research Part I: Oceanographic Research Papers* **57(6)**: 771–784.

James, C.M., Al-Hinty, S., and Salman, A.E. (1989). Growth and w3 fatty acid and amino acid composition of microalgae under different temperature regimes. *Aquaculture* **77(4)**: 337–351.

Jensen, O.N. (2004). Modification-specific proteomics: Characterization of post-translational modifications by mass spectrometry. *Current Opinion in Chemical Biology* **8(1)**: 33–41.

Johns, G.C. and Somero, G.N. (2004). Evolutionary Convergence in Adaptation of Proteins to Temperature: A 4-Lactate Dehydrogenases of Pacific Damselfishes (Chromis spp.). *Molecular Biology and Evolution* **21(2)**: 314–320.

Johnson, Z.I., Zinser, E.R., Coe, A., McNulty, N.P., Woodward, E.M.S., and Chisholm, S.W. (2006). Niche partitioning among Prochlorococcus ecotypes along ocean-scale environmental gradients. *Science (New York, N.Y.)* **311(5768)**: 1737–1740.

Jones, J., Otu, H., Spentzos, D., Kolia, S., Inan, M., Beecken, W.D., Fellbaum, C., Gu, X., Joseph, M., Pantuck, A.J., Jonas, D., and Libermann, T.A. (2005). Gene signatures of progression and metastasis in renal cell cancer. *Clinical Cancer Research* **11(16)**: 5730–5739.

Jørgensen, E.G. (1968). The Adaptation of Plankton Algae. *Physiologia Plantarum* **21(2)**: 423–427.

Kaltz, O. and Bell, G. (2002). The ecology and genetics of fitness in Chlamydomonas. XII. Repeated sexual episodes increase rates of adaptation to novel environments. *Evolution; international journal of organic evolution* **56(9)**: 1743–1753.

Kawecki, T.J., Lenski, R.E., Ebert, D., Hollis, B., Olivieri, I., and Whitlock, M.C. (2012). Experimental evolution. *Trends in ecology & evolution* **27(10)**: 547–60.

Keeling, P.J. (2010). The endosymbiotic origin, diversification and fate of plastids. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **365(1541)**: 729–48.

Kilias, E., Kattner, G., Wolf, C., Frickenhaus, S., and Metfies, K. (2014). A molecular survey of protist diversity through the central Arctic Ocean. *Polar Biology* **37(9)**: 1271–1287.

Kilias, E., Wolf, C., Nöthig, E.M., Peeken, I., and Metfies, K. (2013). Protist distribution in the Western Fram Strait in summer 2010 based on 454-pyrosequencing of 18S rDNA. *Journal of Phycology* **49(5)**: 996–1010.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* **14(4)**: R36.

Kishimoto, T., Iijima, L., Tatsumi, M., Ono, N., Oyake, A., Matsuo, M., Okubo, M., Suzuki, S., Mori, K., Kashiwagi, A., Ying, B.w., and Yomo, T. (2010). Transition from Positive to Neutral in Mutation Fixation along with Continuing Rising Fitness in Thermal Adaptive Evolution. *PLoS genetics* **6(10)**.

Kitsios, G. and Doonan, J.H. (2011). Cyclin dependent protein kinases and stress responses in plants. *Plant signaling & behavior* **6(2)**: 204–9.

Klaas, C. and Archer, D.E. (2002). Association of sinking organic matter with various types of mineral ballast in the deep sea: Implications for the rain ratio. *Global Biogeochemical Cycles* **16(4)**: 1–14.

Klausmeier, C.A., Litchman, E., Daufresne, T., and Levin, S.A. (2004). Optimal nitrogen-to-phosphorus stoichiometry of phytoplankton. *Nature* **429**: 171–174.

Klenke, M. and Schenke, H.W. (2002). A new bathymetric model for the central Fram Strait. *Marine Geophysical Researches* **23(4)**: 367–378.

Klironomos, F.D., Berg, J., and Collins, S. (2013). How epigenetic mutations can affect genetic evolution: Model and mechanism. *BioEssays* **35(6)**: 571–578.

Knies, J.L., Izem, R., Supler, K.L., Kingsolver, J.G., and Burch, C.L. (2006). The genetic basis of thermal reaction norm evolution in lab and natural phage populations. *PLoS Biology* **4(7)**: 1257–1264.

Koboldt, D.C., Chen, K., Wylie, T., Larson, D.E., McLellan, M.D., Mardis, E.R., Weinstock, G.M., Wilson, R.K., and Ding, L. (2009). VarScan: Variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics* **25(17)**: 2283–2285.

Kooistra, W.H. and Medlin, L.K. (1996). Evolution of the diatoms (Bacillariophyta). IV. A reconstruction of their age from small subunit rRNA coding regions and the fossil record. *Molecular phylogenetics and evolution* **6**: 391–407.

Koonin, E.V. and Wolf, Y.I. (2008). Genomics of bacteria and archaea: The emerging dynamic view of the prokaryotic world. *Nucleic Acids Research* **36(21)**: 6688–6719.

Kramer, K., van der Werf, B., and Schelhaas, M.J. (2014). Bring in the genes: genetic-ecophysiological modeling of the adaptive response of trees to environmental change. With application to the annual cycle. *Frontiers in plant science* **5(January)**: 742.

Krol, M., Maxwell, D.P., and Huner, N.P.a. (1997). Exposure of Dunaliella salina to low temperature mimics the high light-induced accumulation of carotenoids and the carotenoid binding protein (Cbr). *Plant and Cell Physiology* **38(2)**: 213–216.

Kuebler, J.E., Davison, I.R., and Yarish, C. (1991). Photosynthetic adaptation to temperature in the red algae Lomentaria baileyana and Lomentaria orcadensis. *British Phycological Journal* **26(1)**: 9–19.

Kumar, S.V. and Wigge, P.a. (2010). H2A.Z-Containing Nucleosomes Mediate the Thermosensory Response in Arabidopsis. *Cell* **140(1)**: 136–147.

Langmead, B. and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Meth* **9(4)**: 357–359.

Larsen, P.E., Scott, N., Post, A.F., Field, D., Knight, R., Hamada, Y., and Gilbert, J.A. (2015). Satellite remote sensing data can be used to model marine microbial metabolite turnover. *Isme Journal* **9(1)**: 166–179.

Lennon, E.F.E. and Houlton, B.Z. (2016). Coupled molecular and isotopic evidence for denitrifier controls over terrestrial nitrogen availability. *The ISME Journal* **11(3)**: 1–14.

Lenski, R.E. and Bennett, A.F. (1993). Evolutionary Response of Escherichia coli to Thermal Stress. *The American Naturalist* **142**: 47–64.

Lenski, R.E., Rose, M.R., Simpson, S.C., and Tadler, S.C. (1991). Long-Term Experimental Evolution in Escherichia coli. I. Adaptation and Divergence During 2,000 Generations. *The American Naturalist* **138(6)**: 1315–1341.

Leonardos, N. and Geider, R.J. (2004). Responses of elemental and biochemical composition of Chaetoceros muelleri to growth under varying light and nitrate:phosphate supply ratios and their influence on critical N:P. *Limnology and Oceanography* **49(6)**: 2105–2114.

Lewis, J.G., Learmonth, R.P., and Watson, K. (1995). Induction of heat, freezing and salt tolerance by heat and salt shock in Saccharomyces cerevisiae. *Microbiology* **141(3)**: 687–694.

Li, P., Filiault, D., Box, M.S., Kerdaffrec, E., Oosterhout, C.V., Wilczek, A.M., Schmitt, J., Mcmullan, M., Bergelson, J., Nordborg, M., and Dean, C. (2014). Multiple FLC haplotypes defined by independent cis-regulatory variation underpin life history diversity in Arabidopsis thaliana. *Genes Dev* **4**: 1635–1640.

Li, W. (2002). Macroecological patterns of phytoplankton in th enorthwestern North Atlantic Ocean. *Nature* **419(September)**: 154–157.

Litchman, E., Edwards, K.F., and Klausmeier, C.A. (2015). Microbial resource utilization traits and trade-offs : implications for community structure , functioning and biogeochemical impacts at present and in the future. *Frontiers in Microbiology* **6(254)**.

Litchman, E., Klausmeier, C.A., and Yoshiyama, K. (2009). Contrasting size evolution in marine and freshwater diatoms. *Proceedings of the National Academy of Sciences of the United States of America* **106(8)**: 2665–2670.

Logan, C.a. and Buckley, B.a. (2015). Transcriptomic responses to environmental temperature in eurythermal and stenothermal fishes. *Journal of Experimental Biology* **218(12)**: 1915–1924.

Lohbeck, K.T., Riebesell, U., Collins, S., and Reusch, T.B.H. (2013). Functional genetic divergence in high CO2 adapted Emiliania huxleyi populations. *Evolution* **67(7)**: 1892–1900.

Lohbeck, K.T., Riebesell, U., and Reusch, T.B.H. (2012). Adaptive evolution of a key phytoplankton species to ocean acidification. *Nature Geosci* **5(5)**: 346–351.

Loladze, I. and Elser, J.J. (2011). The origins of the Redfield nitrogen-to-phosphorus ratio are in a homoeostatic protein-to-rRNA ratio. *Ecology Letters* **14**: 244–250.

Longhurst, A. (1998). *Ecological geography of the sea*. Academic Press. 1st ed.

Los, D., Horvath, I., Vigh, L., and Murata, N. (1993). The temperature-dependent expression of the desaturase gene desA in Synechocystis PCC6803. *FEBS Letters* **318(1)**: 57–60.

Lovejoy, C., Massana, R., and Pedro, C. (2006). Diversity and Distribution of Marine Microbial Eukaryotes in the Arctic Ocean and Adjacent Seas Diversity and Distribution of Marine Microbial Eukaryotes in the Arctic Ocean and Adjacent Seas . *Applied and Environmental Microbiology* **72(5)**: 3085–3095.

Luddington, I.A., Kaczmarska, I., and Lovejoy, C. (2012). Distance and character-based evaluation of the V4 region of the 18S rRNA gene for the identification of diatoms (Bacillariophyceae). *PloS one* **7(9)**: e45664.

Mahadevan, a., D'Asaro, E., Lee, C., and Perry, M.J. (2012). Eddy-Driven Stratification Initiates North Atlantic Spring Phytoplankton Blooms. *Science* **337(6090)**: 54–58.

Malviya, S., Scalco, E., Audic, S., Vincent, F., Veluchamy, A., Bittner, L., Poulain, J., Wincker, P., Iudicone, D., de Vargas, C., Zingone, A., and Bowler, C. (2015). Insights into global diatom distribution and diversity in the world's ocean. *Proceedings of the National Academy of Sciences* **348(6237)**: in review.

Marañón, E., Holligan, P.M., Varela, M., Mouriño, B., and Bale, A.J. (2000). Basin-scale variability of phytoplankton biomass, production and growth in the Atlantic Ocean. *Deep Sea Research Part I: Oceanographic Research Papers* **47(5)**: 825–857.

Marinov, I., Doney, S.C., Lima, I.D., Lindsay, K., Moore, J.K., and Mahowald, N. (2013). North-South asymmetry in the modeled phytoplankton community response to climate change over the 21st century. *Global Biogeochemical Cycles* **27(4)**: 1274–1290.

Marshall, H.G., Burchardt, L., and Lacouture, R. (2005). A review of phytoplankton composition within Chesapeake Bay and its tidal estuaries. *Journal of Plankton Research* **27(11)**: 1083–1102.

Matsen, F.a., Kodner, R.B., and Armbrust, E.V. (2010). pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics* **11(1)**: 538.

Maumus, F. (2009). *Transcriptional and Epigenetic regulation in the marine diatom Phaeodactylum tricornutum*. Ph.D. thesis.

Maumus, F., Allen, A.E., Mhiri, C., Hu, H., Jabbari, K., Vardi, A., Grandbastien, M.A., and Bowler, C. (2009). Potential impact of stress activated retrotransposons on genome evolution in a marine diatom. *BMC genomics* **10**: 624.

Maumus, F., Rabinowicz, P., Bowler, C., and Rivarola, M. (2011). Stemming epigenetics in marine stramenopiles. *Current genomics* **12(5)**: 357–370.

Maxwell, D.P., Falk, S., Trick, C.G., and Huner, N. (1994). Growth at Low Temperature Mimics High-Light Acclimation in Chlorella vulgaris. *Plant physiology* **105(2)**: 535–543.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* **20(9)**: 1297–1303.

Melis, A. (1999). Photosystem-II damage and repair cycle in chloroplasts : what modulates the rate of photodamage in vivo ? *Trends in Plant Science* **4(4)**: 130–135.

Merico, A., Bruggeman, J., and Wirtz, K. (2009). A trait-based approach for downscaling complexity in plankton ecosystem models. *Ecological Modelling* **220(21)**: 3001–3010.

Mirouze, M. and Paszkowski, J. (2011). Epigenetic contribution to stress adaptation in plants. *Current opinion in plant biology* **14(3)**: 267–74.

Mock, T., Daines, S.J., Geider, R., Collins, S., Metodiev, M., Millar, A.J., Moulton, V., and Lenton, T.M. (2015). Bridging the gap between omics and earth system science to better understand how environmental change impacts marine microbes. *Global Change Biology* .

Mock, T. and Hoch, N. (2005). Long-Term Temperature Acclimation of Photosynthesis in Steady-State Cultures of the Polar Diatom Fragilariopsis cylindrus. *Photosynthesis Research* **85(3)**: 307–317.

Mojica, K.D.A., van de Poll, W.H., Kehoe, M., Huisman, J., Timmermans, K.R., Buma, A.G.J., van der Woerd, H.J., Hahn-Woernle, L., Dijkstra, H.A., and Brussaard, C.P.D. (2015). Phytoplankton community structure in relation to vertical stratification along a north-south gradient in the Northeast Atlantic Ocean. *Limnology and Oceanography* **60(5)**: 1498–1521.

Momigliano, P., Harcourt, R., and Stow, A. (2015). Conserving coral reef organisms that lack larval dispersal: are networks of Marine Protected Areas good enough? *Frontiers in Marine Science* **2**: 1–5.

Montagnes, D.J.S. and Franklin, D.J. (2001). Effect of Temperature on Diatom Volume, Growth Rate, and Carbon and Nitrogen Content: Reconsidering Some Paradigms. *Limnology and Oceanography* **46(8)**: 2008–2018.

Morgan-Kiss, R., Ivanov, A.G., Williams, J., Khan, M., and Hüner, N.P.A. (2002). Differential thermal effects on the energy distribution between photosystem II and photosystem I in thylakoid membranes of a psychrophillic and a mesophilic alga. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1561(2)**: 251–265.

Morin, L.G., Smucker, R.A., and Herth, W. (1986). Effects of two chitin synthesis inhibitors on Thalassiosira fluviatilis and Cyclotella cryptica. *FEMS Microbiology Letters* **37(3)**: 263 LP – 268.

Mousing, E.A., Richardson, K., Bendtsen, J., Cetini??, I., Perry, M.J., and Cornell, W. (2016). Evidence of small-scale spatial structuring of phytoplankton alpha- and beta-diversity in the open ocean. *Journal of Ecology* **104(6)**: 1682–1695.

Moustafa, A., Beszteri, B.B., Maier, U.G., Bowler, C., Valentin, K., and Bhattacharya, D. (2009). Genomic Footprints of a Cryptic Plastid Endosymbiosis in Diatoms. *Science* **324(5935)**: 1724–1726.

Nelson, D.M., Tréguer, P., Brzezinski, M.A., Leynaert, A., and Quéguiner, B. (1995). Production and dissolution of biogenic silica in the ocean: Revised global estimates, comparison with regional data and relationship to biogenic sedimentation. *Global Biogeochemical Cycles* **9(3)**: 359–372.

Ness, R.W., Morgan, A.D., Vasanthakrishnan, R.B., and Keightley, P.D. (2015). Extensive de novo mutation rate variation between individuals and across the genome of Chlamydomonas reinhardtii. *bioRxiv* **44(0)**: 1–30.

Nishida, I. and Murata, N. (1996). Chilling sensitivity in plants and cyanobacteria: the crucial contribution of membrane lipids. *Annual Review of Plant Physiology and Plant Molecular Biology* **47(1)**: 541–568.

Nishio, Y., Nakamura, Y., Kawarabayasi, Y., Usuda, Y., Kimura, E., Sugimoto, S., Matsui, K., Yamagishi, A., Kikuchi, H., Ikeo, K., and Gojobori, T. (2003). Comparative Complete Genome Sequence Analysis of the Amino Acid Replacements Responsible for the Thermostability of Corynebacterium efficiens. *Genome Research* **13**: 1572–1579.

Niyogi, K.K., Li, X.P., Rosenberg, V., and Jung, H.S. (2005). Is PsbS the site of non-photochemical quenching in photosynthesis? *Journal of Experimental Botany* **56(411)**: 375–382.

Norberg, J. (2004). Biodiversity and ecosystem functioning: A complex adaptive systems approach. *Limnology and Oceanography* **49(4)**: 1269–1277.

OBIS (????). Continuous Plankton Recorder (CPR) data (phytoplankton) from the Sir Alister Hardy Foundation for Ocean Science (SAHFOS).

O'Connor, M.I., Piehler, M.F., Leech, D.M., Anton, A., and Bruno, J.F. (2009). Warming and resource availability shift food web structure and metabolism. *PLoS Biology* **7(8)**: 3–8.

Ohlrogge, J. and Browse, J. (1995). Lipid biosynthesis. *The Plant Cell* **7**: 957–970.

Oksanen, A.J., Blanchet, F.G., Kindt, R., Minchin, P.R., Hara, R.B.O., Simpson, G.L., Soly, P., Stevens, M.H.H., and Wagner, H. (2016). Community Ecology Package .

Paasche, E. (1980). Silicon content of five marine plankton diatom species measured with a rapid filter method. *Limnology and Oceanography* **25(3)**: 474–480.

Padfield, D., Yvon-Durocher, G., Buckling, A., Jennings, S., and Yvon-Durocher, G. (2016). Rapid evolution of metabolic traits explains thermal adaptation in phytoplankton. *Ecology Letters* **19(2)**: 133–142.

Passow, U. (2002). Production of transparent exopolymer particles (TEP) by phyto- and bacterioplankton. *Marine Ecology Progress Series* **236(Passow 2000)**: 1–12.

Passow, U. and Carlson, C.A. (2012). The biological pump in a high CO2 world. *Marine Ecology Progress Series* **470(2)**: 249–271.

Perrineau, M.M., Gross, J., Zelzion, E., Price, D.C., Levitan, O., Boyd, J., and Bhattacharya, D. (2014). Using Natural Selection to Explore the Adaptive Potential of Chlamydomonas reinhardtii. *PLoS ONE* **9(3)**: e92533.

Perrineau, M.M., Zelzion, E., Gross, J., Price, D.C., Boyd, J., and Bhattacharya, D. (2013). Evolution of salt tolerance in a laboratory reared population of Chlamydomonas reinhardtii. *Environmental microbiology* .

Peterson, M.E., Daniel, R.M., Danson, M.J., and Eisenthal, R. (2007). The dependence of enzyme activity on temperature: determination and validation of parameters. *The Biochemical journal* **402(2)**: 331–337.

Philippart, C.J.M., Anad??n, R., Danovaro, R., Dippner, J.W., Drinkwater, K.F., Hawkins, S.J., Oguz, T., O'Sullivan, G., and Reid, P.C. (2011). Impacts of climate change on European marine ecosystems: Observations, expectations and indicators. *Journal of Experimental Marine Biology and Ecology* **400(1-2)**: 52–69.

Ploug, H., Iversen, M.H., Koski, M., and Buitenhuis, E.T. (2008). Production, oxygen respiration rates, and sinking velocity of copepod fecal pellets: Direct measurements of ballasting by opal and calcite. *Limnology and Oceanography* **53(2)**: 469–476.

Plucain, J., Hindré, T., Le Gac, M., Tenaillon, O., Cruveiller, S., Médigue, C., Leiby, N., Harcombe, W.R., Marx, C.J., Lenski, R.E., and Schneider, D. (2014). Epistasis and allele specificity in the emergence of a stable polymorphism in Escherichia coli. *Science (New York, N.Y.)* **343(6177)**: 1366–9.

Poloczanska, E.S., Brown, C.J., Sydeman, W.J., Kiessling, W., Schoeman, D.S., Moore, P.J., Brander, K., Bruno, J.F., Buckley, L.B., Burrows, M.T., Duarte, C.M., Halpern, B.S., Holding, J., Kappel, C.V., O/'Connor, M.I., Pandolfi, J.M., Parmesan, C., Schwing, F., Thompson, S.A., and Richardson, A.J. (2013). Global imprint of climate change on marine life. *Nature Clim. Change* **advance on**.

Polovina, J.J., Howell, E.A., and Abecassis, M. (2008). Ocean's least productive waters are expanding. *Geophysical Research Letters* **35(3)**: 2–6.

Poulin, M., Daugbjerg, N., Gradinger, R., Ilyash, L., Ratkova, T., and von Quillfeldt, C. (2011). The pan-Arctic biodiversity of marine pelagic and sea-ice unicellular eukaryotes: A first-attempt assessment. *Marine Biodiversity* **41(1)**: 13–28.

Poulin, M., Underwood, G.J., and Michel, C. (2014). Sub-ice colonial Melosira arctica in Arctic first-year ice. *Diatom Research* **29(2)**: 213–221.

Poulsen, N., Chesley, P., and Kröger, N. (2006). Molecular Genetic Manipulation of the Diatom Thalassiosira pseudonana (Bacillariophyceae). *Journal of Phycology* **1065(42)**: 1059–1065.

Prihoda, J., Tanaka, A., de Paula, W.B.M., Allen, J.F., Tirichine, L., and Bowler, C. (2012). Chloroplast-mitochondria cross-talk in diatoms. *Journal of experimental botany* **63(4)**: 1543–57.

Prokopowich, C.D., Gregory, T.R., and Crease, T.J. (2003). The correlation between rDNA copy number and genome size in eukaryotes. *Genome / National Research Council Canada = Genome / Conseil national de recherches Canada* **46(1)**: 48–50.

Puentes-téllez, P.E., Kovács, Á.T., Kuipers, O.P., and Elsas, J.D.V. (2014). Comparative genomics and transcriptomics analysis of experimentally evolved Escherichia coli MC1000 in complex environments. *Environmental microbiology* **16(3)**: 856–870.

Rae, P.M. (1976). Hydroxymethyluracil in eukaryote DNA: a natural feature of the pyrrophyta (dinoflagellates). *Science* **194(4269)**: 1062–1064.

Ragueneau, O., Schultes, S., Bidle, K., Claquin, P., and Moriceau, B. (2006). Si and C interactions in the world ocean: Importance of ecological processes and implications for the role of diatoms in the biological pump. *Global Biogeochemical Cycles* **20(4)**.

Rahavi, M. (2011). Transgenerational adaptation to heavy metal salts in Arabidopsis. *Frontiers in Plant Science* **2(December)**: 1–10.

Raison, J.K., Roberts, J.K., and Berry, J.A. (1982). Correlations between the thermal stability of chloroplast (thylakoid) membranes and the composition and fluidity of their polar lipids upon acclimation of the higher plant, Nerium oleander, to growth temperature. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **688(1)**: 218–228.

Ratcliff, W.C., Denison, R.F., Borrello, M., and Travisano, M. (2012). Experimental evolution of multicellularity. *Proceedings of the National Academy of Sciences* **109(5)**: 1595–1600.

Ratcliff, W.C., Herron, M.D., Howell, K., Pentz, J.T., Rosenzweig, F., and Travisano, M. (2013). Experimental evolution of an alternating uni- and multicellular life cycle in Chlamydomonas reinhardtii. *Nature communications* **4(May)**: 2742.

Raven, J.A. and Falkowski, P.G. (1999). Oceanic sinks for atmospheric CO2. *Plant, Cell and Environment* **22(6)**: 741–755.

Raven, J.A. and Geider, R.J. (1988). Temperature and algal growth. *New Phytologist* **110(4)**: 441–461.

Reay, D.S., Nedwell, D.B., Priddle, J., and Icrobiol, A.P.P.L.E.N.M. (1999). Temperature Dependence of Inorganic Nitrogen Uptake : Reduced Affinity for Nitrate at Suboptimal Temperatures in Both Algae and Bacteria **65(6)**: 2577–2584.

Reay, D.S., Priddle, J., Nedwell, D.B., Whitehouse, M.J., Ellis-Evans, J.C., Deubert, C., and Connelly, D.P. (2001). Regulation by low temperature of phytoplankton growth and nutrient uptake in the Southern Ocean. *Marine Ecology Progress Series* **219(1990)**: 51–64.

Redfield, A.C. (1958). The biological control of chemical factors in the environment. *American Scientist* **46**: 205–221.

Reid, P.C., Johns, D.G., Edwards, M., Starr, M., Poulin, M., and Snoeijs, P. (2007). A biological consequence of reducing Arctic ice cover: arrival of the Pacific diatom Neodenticula seminae in the North Atlantic for the first time in 800000 years. *Global Change Biology* **13(9)**: 1910–1921.

Richards, E.J. (2006). Inherited epigenetic variation - Revisiting soft inheritance. *Nature reviews. Genetics* **7(5)**: 395–402.

Richardson, A.J. and Schoeman, D.S. (2004). Climate impact on plankton ecosystems in the Northeast Atlantic. *Science (New York, N.Y.)* **305(5690)**: 1609–12.

Riddle, N.C. and Richards, E.J. (2002). The control of natural variation in cytosine methylation in Arabidopsis. *Genetics* **162(1)**: 355–363.

Riebesell, U., Körtzinger, A., and Oschlies, A. (2009). Sensitivities of marine carbon fluxes to ocean change. *Proceedings of the National Academy of Sciences of the United States of America* **106(49)**: 20602–20609.

Riebesell, U., Zondervan, I., Rost, B., Tortell, P.D., Zeebe, R.E., and Morel, F.M. (2000). Reduced calcification of marine plankton in response to increased atmospheric CO2. *Nature* **407(September)**: 364–367.

Riehle, M.M., Bennett, A.F., Lenski, R.E., and Long, A.D. (2003). Evolutionary changes in heat-inducible gene expression in lines of Escherichia coli adapted to high temperature. *Physiological Genomics* **14(1)**: 47–58.

Robinson, C., Poulton, A.J., Holligan, P.M., Baker, A.R., Forster, G., Gist, N., Jickells, T.D., Malin, G., Upstill-Goddard, R., Williams, R.G., Woodward, E.M.S., and Zubkov, M.V. (2006). The Atlantic Meridional Transect (AMT) Programme: A contextual view 1995-2005. *Deep-Sea Research Part II: Topical Studies in Oceanography* **53(14-16)**: 1485–1515.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2009). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26(1)**: 139–140.

Robinson, M.D. and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome biology* **11(3)**: R25.

Rockman, M.V. (2008). Reverse engineering the genotype-phenotype map with natural genetic variation. *Nature* **456(7223)**: 738–744.

Roemmich, D. and McGowan, J. (1995). Climatic warming and the decline of zooplankton in the california current. *Science (New York, N.Y.)* **267(5202)**: 1324–1326.

Rossoll, D., Bermundez, R., Hauss, H., Schulz, K.G., Riebesell, U., Sommer, U., and Winder, M. (2012). Ocean acidification-induced food quality deterioration constrains trophic transfer. *PLoS ONE* **7(4)**: 2–7.

Round, F.E., Crawford, R.M., and Mann, D.G. (1990). *The diatoms: Biology and morphology of the genera*. Cambridge University Press.

Rousseaux, C.S. and Gregg, W.W. (2015). Recent decadal trends in global phytoplankton composition. *Global Biogeochemical Cycles* **29**: 1674–1688.

Roy, S. and Morse, D. (2012). A full suite of histone and histone modifying genes are transcribed in the dinoflagellate Lingulodinium. *PloS one* **7(4)**: e34340.

Rudolph, B., Gebendorfer, K.M., Buchner, J., and Winter, J. (2010). Evolution of Escherichia coli for growth at high temperatures. *Journal of Biological Chemistry* **285(25)**: 19029–19034.

Rusch, D.B., Halpern, A.L., Sutton, G., Heidelberg, K.B., Williamson, S., Yooseph, S., Wu, D., Eisen, J.A., Hoffman, J.M., Remington, K., Beeson, K., Tran, B., Smith, H., Baden-Tillson, H., Stewart, C., Thorpe, J., Freeman, J., Andrews-Pfannkoch, C., Venter, J.E., Li, K., Kravitz, S., Heidelberg, J.F., Utterback, T., Rogers, Y.H., Falc??n, L.I., Souza, V., Bonilla-Rosso, G., Eguiarte, L.E., Karl, D.M., Sathyendranath, S., Platt, T., Bermingham, E., Gallardo, V., Tamayo-Castillo, G., Ferrari, M.R., Strausberg, R.L., Nealson, K., Friedman, R., Frazier, M., and Venter, J.C. (2007). The Sorcerer II Global Ocean Sampling expedition: Northwest Atlantic through eastern tropical Pacific. *PLoS Biology* **5(3)**: 0398–0431.

Saint-Ruf, C. and Matic, I. (2006). Environmental tuning of mutation rates. *Environmental Microbiology* **8(2)**: 193–199.

Salvucci, M.E. and Crafts-Brandner, S.J. (2004). Inhibition of photosynthesis by heat stress: The activation state of Rubisco as a limiting factor in photosynthesis. *Physiologia Plantarum* **120(2)**: 179–186.

Sarmiento, J.L., Dunne, J., Gnanadesikan, A., Key, R.M., Matsumoto, K., and Slater, R. (2002). A new estimate of the CaCO3 to organic carbon export ratio. *Global Biogeochemical Cycles* **16(4)**: 1107.

Schaum, E. and Collins, S. (2014). Plasticity predicts evolution in a marine algae. *Proceedings of the Royal Society B: Biological Sciences* **281**: 20141486.

Schaum, E., Rost, B., Millar, A.J., and Collins, S. (2013). Variation in plastic responses of a globally distributed picoplankton species to ocean acidification. *Nature Climate Change* **3(3)**: 298–302.

Scheinin, M., Riebesell, U., Rynearson, T.A., Lohbeck, K.T., and Collins, S. (2015). Experimental evolution gone wild. *J. R. Soc. Interface* **12**: 1–5.

Schlüter, L., Lohbeck, K.T., Gutowska, M.a., Gröger, J.P., Riebesell, U., and Reusch, T.B.H. (2014). Adaptation of a globally important coccolithophore to ocean warming andacidification. *Nature Climate Change* pp. 1–14.

Schmieder, R. and Edwards, R. (2011). Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **27(6)**: 863–864.

Schmitz, R.J., Schultz, M.D., Lewsey, M.G., O'Malley, R.C., Urich, M.A., Libiger, O., Schork, N.J., and Ecker, J.R. (2011). Transgenerational Epigenetic Instability Is a Source of Novel Methylation Variants. *Science* **334(6054)**: 369–373.

Schoemann, V., Becquevort, S., Stefels, J., Rousseau, V., and Lancelot, C. (2005). \textit{Phaeocystis} blooms in the global ocean and their controlling mechanisms: A review. *Journal of Sea Research* **53(1-2 SPEC. ISS.)**: 43–66.

Selman, M., Pardo, A., Barrera, L., Estrada, A., Watson, S.R., Wilson, K., Aziz, N., Kaminski, N., and Zlotnik, A. (2006). Gene expression profiles distinguish idiopathic pulmonary fibrosis from hypersensitivity pneumonitis. *American Journal of Respiratory and Critical Care Medicine* **173(2)**: 188–198.

Shuter, B.J., Thomas, J.E., Taylor, W.D., and Zimmerman, A.M. (1983). Phenotypic Correlates of Genomic DNA Content in Unicellular Eukaryotes and Other Cells. *The American Naturalist* **122(1)**: 26–44.

Siegel, D.A., Doney, S.C., and Yoder, J.A. (2002). The North Atlantic Spring Phytoplankton Bloom and Sverdrup's Critical Depth Hypothesis. *Science* **296(5568)**: 730–733.

Sims, P.a., Mann, D.G., and Medlin, L.K. (2006). Evolution of the diatoms: insights from fossil, biological and molecular data. *Phycologia* **45(4)**: 361–402.

Smetacek, V. (1999). Diatoms and the ocean carbon cycle. *Protist* **150(1)**: 25–32.

Smetacek, V. (2000). Oceanography: The giant diatom dump. *Nature* **406(6796)**: 574–575.

Sobrino, C. and Neale, P.J. (2007). Short-term and long-term effects of temperature on photosynthesis in the diatom Thalassiosira pseudonana under UV exposures. *Journal of Phycology* **43(3)**: 426–436.

Somero, G.N. (2004). Adaptation of enzymes to temperature: Searching for basic "strategies". *Comparative Biochemistry and Physiology - B Biochemistry and Molecular Biology* **139**: 321–333.

Sørensen, J.G., Nielsen, M.M., and Loeschcke, V. (2007). Gene expression profile analysis of Drosophila melanogaster selected for resistance to environmental stressors. *Journal of Evolutionary Biology* **20(4)**: 1624–1636.

Spilling, K., Ylöstalo, P., Simis, S., and Seppälä, J. (2015). Interaction Effects of Light, Temperature and Nutrient Limitations (N, P and Si) on Growth, Stoichiometry and Photosynthetic Parameters of the Cold-Water Diatom Chaetoceros wighamii. *Plos One* **10(5)**: e0126308.

Stapleford, L.S. and Smith, R.E.H. (1996). The interactive effects of temperature and silicon limitation on the psychrophilic ice diatom Pseudonitszchia seriata. *Polar Biology* **16**: 589–594.

Stecher, A., Neuhaus, S., Lange, B., Frickenhaus, S., Beszteri, B., Kroth, P.G., and Valentin, K. (2015). rRNA and rDNA based assessment of sea ice protist biodiversity from the central Arctic Ocean. *European Journal of Phycology* **0262(November)**: 1–16.

Steinacher, M., Joos, F., Foelicher, T.L., Bopp, L., Cadule, P., Cocco, V., Doney, S.C., Gehlen, M., Lindsay, K., Moore, J.K., Schneider, B., and Segschneider, J. (2010). Projected 21st century decrease in marine productivity: a multi-model analysis. *Biogeosciences* **7**: 979–1005.

Sterner, R.W. and Elser, J.J. (2002). *Ecological stoichiometry: the biology of elements from molecules to the biosphere*. Princeton University Press.

Stockwell, D.A., Whitledge, T.E., Zeeman, S.I., Coyle, K.O., Napp, J.M., Brodeur, R.D., Pinchuk, A.I., and Hunt, G.L. (2001). Anomalous conditions in the south-eastern Bering Sea, 1997: Nutrients, phytoplankton and zooplankton. *Fisheries Oceanography* **10(1)**: 99–116.

Stoebel, D.M., Hokamp, K., Last, M.S., and Dorman, C.J. (2009). Compensatory evolution of gene regulation in response to stress by Escherichia coli lacking RpoS. *PLoS Genetics* **5(10)**: 1–9.

Stramski, D., Sciandra, A., and Claustre, H.H. (2002). Effects of temperature, nitrogen, and light limitation on the optical properties of the marine diatom Thalassiosira pseudonana. *Limnology and Oceanography* **47(2)**: 392–403.

Strauss, J. (2012). *A genomic analysis using RNA - Seq to investigate the adaptation of the psychrophilic diatom Fragilariopsis cylindrus to the polar environment*. Ph.D. thesis.

Supek, F., Bošnjak, M., Škunca, N., and Šmuc, T. (2011). Revigo summarizes and visualizes long lists of gene ontology terms. *PLoS ONE* **6(7)**.

Suzumura, M. (2008). Persulfate chemical wet oxidation method for the determination of particulate phosphorus in comparison with a high-temperature dry combustion method. *Limnology and Oceanography: Methods* **6**: 619–629.

Takeda, S., Sugimoto, K., Kakutani, T., and Hirochika, H. (2001). Linear DNA intermediates of the Tto1 retrotransposon in Gag particles accumulated in stressed tobacco and Arabidopsis thaliana. *Plant Journal* **28(3)**: 307–317.

Tang, E.P.Y. and Vincent, W.F. (1999). Strategies of thermal adaptation by high-latitude cyanobacteria. *New Phytologist* **142(2)**: 315–323.

Tarran, G.A., Heywood, J.L., and Zubkov, M.V. (2006). Latitudinal changes in the standing stocks of nano- and picoeukaryotic phytoplankton in the Atlantic Ocean. *Deep-Sea Research Part II: Topical Studies in Oceanography* **53(14-16)**: 1516–1529.

Tatters, A.O., Roleda, M.Y., Schnetzer, A., Fu, F., Hurd, C.L., Boyd, P.W., Caron, D.A., Lie, A.A.Y., Hoffmann, L.J., and Hutchins, D.A. (2013*a*). Short- and long-term conditioning of a temperate marine diatom community to acidification and warming. *Philosophical Transactions of the Royal Society B: Biological Sciences* **368(1627)**: 20120437.

Tatters, A.O., Schnetzer, A., Fu, F., Lie, A.Y.a., Caron, D.a., and Hutchins, D.a. (2013*b*). Short- versus long-term responses to changing CO2 in a coastal dinoflagellate bloom: implications for interspecific competitive interactions and community structure. *Evolution* **67(7)**: 1879–91.

Thompson, J.D., Desmond, H.G., and Gibson, T.J. (1994). Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22(22)**: 4673–4680.

Thompson, P. (1999). The response of growth and biochemical composition to variations in daylenghth, temperature and irradiance in the marine diatom Thalassiosira pseudonana. *Journal of Phycology* **35**: 1215–1223.

Thompson, P., Guo, M., and Harrison, P. (1992*a*). Effect of variation in temperature. I. on the biochemical composition of eight species of marine phytoplankton.

Thompson, P.A., Guo, M.X., Harrison, P.J., and Whyte, J.N.C. (1992*b*). Effects of Variation in Temperature .II. on the Fatty-Acid Composition of eight Species of Marine Phytoplankton.

Toseland, A., Daines, S.J., Clark, J.R., Kirkham, A., Strauss, J., Uhlig, C., Lenton, T.M., Valentin, K., Pearson, G.A., Moulton, V., and Mock, T. (2013). The impact of temperature on marine phytoplankton resource allocation and metabolism. *Nature Climate Change* **3(11)**: 979–984.

Trainor, F.R. (1993). Cyclomorphosis in Scenedesmus subspicatus (Chlorococcales, Chlorophyta): stimulation of colony development at low temperature. *Phycologia* **32(6)**: 429–433.

Tréguer, P., Nelson, D.M., Van Bennekom, a.J., Demaster, D.J., Leynaert, A., and Quéguiner, B. (1995). The silica balance in the world ocean: A reestimate. *Science* **268(5209)**: 375–379.

Urabe, J., Togari, J., and Elser, J.J. (2003). Stoichiometric impacts of increased carbon dioxide on a planktonic herbivore. *Global Change Biology* **9(6)**: 818–825.

Van Doorslaer, W., Stoks, R., Jeppesen, E., and De Meester, L. (2007). Adaptive microevolutionary responses to simulated global warming in Simocephalus vetulus: A mesocosm study. *Global Change Biology* **13(4)**: 878–886.

van Wijk, S.J., Taylor, M.I., Creer, S., Dreyer, C., Rodrigues, F.M., Ramnarine, I.W., Van Oosterhout, C., and Carvalho, G.R. (2013). Experimental harvesting of fish populations drives genetically based shifts in body size and maturation. *Frontiers in Ecology and the Environment* **11(4)**: 181–187.

Vanormelingen, P., Verleyen, E., and Vyverman, W. (2008). The diversity and distribution of diatoms: From cosmopolitanism to narrow endemism. *Biodiversity and Conservation* **17(2)**: 393–405.

Vargas, C.D., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., Lara, E., Berney, C., Bescot, N.L., Probert, I., Carmichael, M., Poulain, J., and Romac, S. (2015). Eukaryotic plankton diversity in the sunlit ocean. *Science* **348(6237)**: 1–11.

Várkonyi, Z., Masamoto, K., Debreczeny, M., Zsiros, O., Ughy, B., Gombos, Z., Domonkos, I., Farkas, T., Wada, H., and Szalontai, B. (2002). Low-temperature-induced accumulation of xanthophylls and its structural consequences in the photosynthetic membranes of the cyanobacterium Cylindrospermopsis raciborskii: an FTIR spectroscopic study. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 2410–2415.

Vélez-Bermúdez, I.C. and Schmidt, W. (2014). The conundrum of discordant protein and mRNA expression. Are plants special? *Frontiers in plant science* **5(November)**: 619.

Veluchamy, A., Lin, X., Maumus, F., Rivarola, M., Bhavsar, J., Creasy, T., O'Brien, K., Sengamalay, N.A., Tallon, L.J., Smith, A.D., Rayko, E., Ahmed, I., Le Crom, S., Farrant, G.K., Sgro, J.Y., Olson, S.A., Bondurant, S.S., Allen, A.E., Allen, A.E., Rabinowicz, P.D., Sussman, M.R., Bowler, C., Tirichine, L., O'Brien, K., Sengamalay, N.A., Tallon, L.J., Smith, A.D., Rayko, E., Ahmed, I., Crom, S.L., Farrant, G.K., Sgro, J.Y., Olson, S.A., Bondurant, S.S., Allen, A.E., Rabinowicz, P.D., Sussman, M.R., Bowler, C., and Tirichine, L. (2013). Insights into the role of DNA methylation in diatoms by genome-wide profiling in Phaeodactylum tricornutum. *Nature communications* **4**: 2091.

Via, S. and Lande, R. (1985). Genotype-Environment Interaction and the Evolution of Phenotypic Plasticity. *Evolution* **39(3)**: 505–522.

Vigh, L., Los, D.a., Horváth, I., and Murata, N. (1993). The primary signal in the biological perception of temperature: Pd-catalyzed hydrogenation of membrane lipids stimulated the expression of the desA gene in Synechocystis PCC6803. *Proceedings of the National Academy of Sciences of the United States of America* **90(19)**: 9090–9094.

Vogt, G., Woell, S., and Argos, P. (1997). Protein thermal stability, hydrogen bonds, and ion pairs. *Journal of Molecular Biology* **269(4)**: 631–643.

Volk, T. and Hoffert, M.I. (1985). Ocean Carbon Pumps: Analysis of Relative Strengths and Efficiencies in Ocean-Driven Atmospheric CO2 Changes. in *The Carbon Cycle and Atmospheric CO2: Natural Variations Archean to Present*. pp. 99–110. American Geophysical Union.

Von Dassow, P., Petersen, T.W., Chepurnov, V.A., and Virginia Armbrust, E. (2008). Inter- and intraspecific relationships between nuclear DNA content and cell size in selected members of the centric diatom genus Thalassiosira (Bacillariphycea). *Journal of Phycology* **44(2)**: 335–349.

Waal, D.B.V.D., Verschoor, A.M., Verspagen, J.M.H., van Donk, E., and Huisman, J. (2010). Climate-driven changes in the ecological stoichiometry of aquatic ecosystems. *Frontiers in Ecology and the Environment* **8(3)**: 145–152.

Walther, G.r., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T.J.C., Fromentin, J.m., I, O.H.g., and Bairlein, F. (2002). Ecological responses to recent climate change. *Nature* **416**: 389–395.

Wei, K.y. and Kennett, J.P. (1988). Phyletic Gradualism and Punctuated Equilibrium in the Late Neogene Planktonic Foraminiferal Clade Globoconella. *Paleobiology* **14(4)**: 345–363.

West-Eberhard, M.J. (2005). Developmental plasticity and the origin of species differences. *Proceedings of the National Academy of Sciences of the United States of America* **102 Suppl(2)**: 6543–6549.

White, F.N. and Somero, G. (1982). Acid-base regulation and phospholipid adaptations to temperature: time courses and physiological significance of modifying the milieu for protein function. *Physiological Reviews* **62(1)**: 40–90.

Wilhelm, C., Büchel, C., Fisahn, J., Goss, R., Jakob, T., LaRoche, J., Lavaud, J., Lohr, M., Riebesell, U., Stehfest, K., Valentin, K., and Kroth, P.G. (2006). The Regulation of Carbon and Nutrient Assimilation in Diatoms is Significantly Different from Green Algae. *Protist* **157(2)**: 91–124.

Winder, M. and Schindler, D.E. (2004). Climate change uncouples trophic interactions in an aquatic ecosystem. *Ecology* **85(8)**: 2100–2106.

Wodniok, S., Brinkmann, H., Glöckner, G., Heidel, A.J., Philippe, H., Melkonian, M., and Becker, B. (2011). Origin of land plants: do conjugating green algae hold the key? *BMC evolutionary biology* **11(1)**: 104.

Xu, K., Fu, F.X., and Hutchins, D.a. (2014). Comparative responses of two dominant Antarctic phytoplankton taxa to interactions between ocean acidification, warming, irradiance, and iron availability. *Limnology and Oceanography* **59(6)**: 1919–1931.

Yallop, M.L. (2001). Distribution patterns and biomass estimates of diatoms and autotrophic dinoflagellates in the NE Atlantic during June and July 1996. *Deep Sea Research Part II: Topical Studies in Oceanography* **48(July 1996)**: 825–844.

Yan, W. and Hunt, L.A. (1999). An equation for modelling the temperature response of plants using only the cardinal temperatures. *Annals of Botany* **(84)**: 607–614.

Ying, B.W., Matsumoto, Y., Kitahara, K., Suzuki, S., Ono, N., Furusawa, C., Kishimoto, T., and Yomo, T. (2015). Bacterial transcriptome reorganization in thermal adaptive evolution. *BMC Genomics* **16(1)**: 802.

Young, M.D., Wakefield, M.J., Smyth, G.K., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: Accounting for selection bias. *Genome biology* **11(2)**: R14.

Yvon-durocher, G., Dossena, M., Trimmer, M., Woodward, G., and Allen, A.P. (2015). Temperature and the biogeography of algal stoichiometry. *Global Ecology and Biogeography* pp. 1–9.

Zhou, X. and Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. *Nature genetics* **44(7)**: 821–4.

Zhu, F., Massana, R., Not, F., Marie, D., and Vaulot, D. (2005). Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiology Ecology* **52(1)**: 79–92.

Zimmermann, J., Jahn, R., and Gemeinholzer, B. (2011). Barcoding diatoms: evaluation of the V4 subregion on the 18S rRNA gene, including new primers and protocols. *Organisms Diversity & Evolution* **11(3)**: 173–192.

Zubkov, M.V., Sleigh, M.A., Tarran, G.A., Burkill, P.H., and Leakey, R.J.G. (1998). Picoplanktonic community structure on an Atlantic transect from 50\degree N to 50\degree S. *Deep-Sea Research Part I: Oceanographic Research Papers* **45(8)**: 1339–1355.

Zuur, A.F., Ieno, E.N., and Elphick, C.S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution* **1(1)**: 3–14.