THE EFFECT OF THE TOXIC DINOFLAGELLATE *ALEXANDRIUM FUNDYENSE* ON THE CALANOID COPEPOD *CALANUS FINMARCHICUS*

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION
OF THE UNIVERSITY OF HAWAIʻI AT MĀNOA IN PARTIAL
FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY
IN

ZOOLOGY

AUGUST 2015

By
Vittoria Roncalli

Dissertation Committee:

Daniel K. Hartline, Chairperson
Petra H. Lenz, Thesis advisor
Andrew E. Christie
Michael G. Hadfield
Erica Goetze

Keywords: Harmful algal blooms, Detoxification, Transcriptomics, Stress response, Glutathione S-transferase

To my mom…

Acknowledgments

The completion of this dissertation could not have occurred without the incredible support of many individuals. First and foremost, I would like to thank my advisor, Petra for her endless patience and support throughout my graduate career. She has been an incredible mentor, always standing up for me when I needed and ready support me during the numerous steps in my doctoral journey when I was stuck. I'd like to thank the members of my committee (in alphabetical order) for their time, advice and support. Andy Christie was there for me since day one of my project and his kindness and generosity are unmatched. Mike Hadfield provided invaluable insights into my work and for challenging me to think about problems from a diverse scientific angle. Dan Hartline continues to inspire me with his dedication to research quality, and was invaluable for his critiques throughout the entire project, always ready with new challenges. And last but not least, Erica Goetze has provided me with invaluable insights into my work, particularly with regards to the role of copepod from an the oceanographic perspective and every chapter has been substantially improved by her input.

I offer thanks to the Bekesy lab members, for all that they have done to support me as a person and scientist throughout the years. A special thank to Lynn Hata, Mattew Ciecklak, Brad Jones, Ian Cooke, Ann Castelfranco, and everyone else who I have worked with over the past four years.

A special thank to Adrianna Ianora and Antonio Miralto that made my dream come true. They shared with me their infinite love for science since I was a BA student and since then, they never stopped to support my scientific career. Without them I would have earned this degree today and I wouldn't be the person I am.

Finally, I must thank my family for without it, there is no way that I would be where I am today. I want to thank my dad Sasa', my uncle Claudio, my brothers Dede', Fabrizio and Rigel for a lifetime of guidance and love and for being present over those four years although two oceans apart. My friends, here in Hawaiʻi that have been with me through the up and down sides that happen on the journey towards this doctoral degree. I especially thank Michelle Jungbluth for being my best friend and because of her I am a better scientist, a better surfer and a better person. But most of all, I thank my mum, Rosaria Colucci, for her inconditionate love and for her support during those 4 years of my journey.

Mahalo nui loa

This research is based upon work supported by the the National Science Foundation under grants OCE-1040597 and OCE-1235549 to Petra H. Lenz, grant IOS-1353023 to Andrew E. Christie, Cades Foundation of Honolulu, Hawaii, Mount Desert Island Biological Laboratory's David W. Towle Graduate Fellowship Summer 2012 and PBRC support to Vittoria Roncalli.

Contribution of Authors

Though none of this dissertation would have been possible without my advisor, Petra H. Lenz, she is listed as co-author for all the Chapters. Petra extensively contributed to the realization of this study by conceiving and designing the experiment, by helping in the data interpretation and writing the paper. Petra also provided institutional support that was vital to the completion of the work. Daniel K. Hartline and Andrew E. Christie have extensively advised me during the data interpretation and the writing of all the Chapters of this dissertation. Matthew Cieslak has contributed to the project to analyze the data by providing computer support. Chapter 2 is also co-authored by Jefferson T. Turner, Daniel D. Anderson and Dave Kulis. Jefferson T. Turner helped to perform the experiments and provided his expertise during the writing of the paper. Donald D. Anderson and Dave Kulis from Woods Hole Oceanographic Institution (WHOI) provided the algae used in the experiment and performed the toxin analysis. They also provided their comments on the paper. Chapter 4 is co- authored by Matthew Cieslak, Andrew E. Christie, Yale Passamanek and Petra H. Lenz. The work has been conceived and designed by myself, Petra H. Lenz and Andrew E. Christie. Analysis of the data involved also Matthew Cieslak, and Yale Passamanek . The paper has been written by myself, Petra H. Lenz and Andrew E. Christie.

Abstract

Harmful algal blooms have represented a serious problem for local economies and public health worldwide. The propagation of toxins through the marine web chain, and the negative effect on higher consumers, has raised questions on how primary consumers, such as herbivorous zooplankton respond to toxic algae. Inshore and offshore waters of the Gulf of Maine (USA) are seasonally dominated by the toxic dinoflagellate *Alexandrium fundyense*, which is responsible for paralytic shellfish poisoning (PSP) in humans. The calanoid copepod *Calanus finmarchicus* co-occurs with *A. fundyense* during spring and summer blooms. High survival has been reported for this copepod exposed to the dinoflagellate, however, little is known about other effects. The goal of my dissertation work was to study the effect of *A. fundyense* on the fitness and physiology of *C. finmarchicus* in controlled laboratory experiments. Fitness was measured as survival and reproductive success. Changes in global gene expression using RNA-Seq technology were used to measure the physiological response. The results suggest that blooms of *A. fundyense* are an environmental challenge for *C. finmarchicus* with a negative effect on copepod population growth. The negative effects included a decrease in reproduction, as well as metabolic adjustments, such as a decrease in lipid biosynthesis and protein synthesis/cell growth. My results demonstrate that transcriptomics is a tool that can be used to investigate the physiological ecology of non-model species such as copepods in order to identify biological processes affected by non-optimal environmental conditions.

Table of Contents

Acknowledgments..iii

Contribution of Authors...v

Abstract...vi

List of Tables…..viii

List of Figures...ix

Chapter 1. Introduction…...1

Chapter 2. The effect of the toxic dinoflagellate *Alexandrium fundyense* on the fitness of the calanoid copepod *Calanus finmarchicus*...12

Chapter 3. Global gene expression analysis of *C. finmarchicus* in response to *A. fundyense*...43

Chapter 4. Glutathione S-Transferase (GST) Gene Diversity in the Crustacean *Calanus finmarchicus* – Contributors to Cellular Detoxification..74

Chapter 5. Relative expression of Glutathione S-transferase genes in *Calanus finmarchicus* females feeding on the toxic dinoflagellate *Alexandrium fundyense* measured by RNA-Seq and quantitative reverse transcription PCR (RT-qPCR)...114

Chapter 6. Summary and Conclusions...131

References...137

Appendices..145

List of Tables

Table 2.1. Experimental treatments..35

Table 2.2. Summary of parameters monitored in *C. finmarchicus*..36

Table 2.3. Summary of statistic results..37

Table 3.1. RNA-sequencing and mapping results...66

Table 3.2. *C. finmarchicus* detoxification enzymes...67

Table 3.3. GO enrichment analysis of up- and down- regulated genes at 2 days.....................68

Table 3.4. GO enrichment analysis of up and down regulated genes at 5 days........................69

Table 4.1. Search results from *in silico* mining of GSTs in *Calanus finmarchicus*..................105

Table 4.2. Annotation of putative glutathione S-transferase-encoding....................................106

Table 4.3. Number of GST *Calanus finmarchicus*, *Tigriopus japonicus* and *Daphnia pulex* and the insect *Drosophila melanogaster*..107

Table 4.4. Comparison between putative *Calanus finmarchicus* glutathione S-transferases (GSTs) from the Gulf of Maine and the Norwegian Sea populations......................................108

Table 4.5. Putative *Calanus finmarchicus* glutathione S-transferases (GSTs) showing differences between the Gulf of Maine and the Norwegian Sea *de novo* assemblies.................................109

Table 5.1. *C. finmarchicus* candidate reference genes...127

List of Figures

Figure 2.1. Toxin profile of the dinoflagellate *Alexandrium fundyense*....................................38

Figure 2.2. Grazing activity for *C. finmarchicus* fed with LD and HD of *A. fundyense* and the control *Rhodomonas* spp. for 7 days..39

Figure 2.3. *C. finmarchicus* egg production rate after exposure to LD and HD of *A. fundyense* and the control *Rhodomonas* spp. for 7-days..40

Figure 2.4. Egg viability and effective recruitment (nauplii female^{-1}) of *C. finmarchicus* adult females fed with LD and HD of *A. fundyense* and the control *Rhodomonas* spp. for 7 days......41

Figure 2.5. *C. finmarchicus* survival rates after exposure to LD and HD of *A. fundyense* and the control *Rhodomonas spp.* for 7 days..42

Figure 3.1. Differentially expressed genes in *C. finmarchicus* adult females feeding on *A. fundyense* diets (LD, HD) for 2 days and 5 days...70

Figure 3.2. Fold change expression change in *C. finmarchicus* adult females feeding on *A. fundyense* (LD, HD) for 2 and 5 days..71

Figure 3.3. *C. finmarchicus* profile of DE genes involved in stress response.............................72

Figure 3.4. *C. finmarchicus* DE genes involved energy metabolism, reproduction and growth..73

Figure 4.1. Alignment of selected *Calanus finmarchicus* Delta GST (Calfi-Delta-IV) and Mu GST (Calfi-Mu-III) proteins with their top arthropod protein BLAST hits...............................110

Figure 4.2. Alignment of *Calanus finmarchicus* microsomal glutathione S-transferase subclass 1 proteins with mGST-1s from other crustaceans...111

Figure 4.3. Phylogenetic tree for cytosolic GSTs from the crustacean *Calanus finmarchicus* and other selected crustacean and insect species...112

Figure 4.4. Relative expression of selected *Calanus finmarchicus* cytosolic, mitochondrial and microsomal GST-encoding transcripts across six developmental stages...................................113

Figure 5.1. Relative expression measured with RNA-Seq and RT-qPCR.................................128

Figure 5.2. Glutathione S-transferase relative expression over time in *C. finmarchicus* feeding on *A. fundyense* diets for 5 days..129

Figure 5.3. Comparison of relative expression levels measured with RT-qPCR for June and July experiments...130

Chapter 1

1. INTRODUCTION

The goal of this dissertation was to investigate the effect of a toxic dinoflagellate, *Alexandrium fundyense* on the fitness and global gene expression of the calanoid copepod *Calanus finmarchicus*. During the summer, *C. finmarchicus* co-occurs with harmful algal blooms (HABs) of *A. fundyense,* which are common in the offshore waters of the Gulf of Maine. This dinoflagellate produces saxitoxins (STXs), which are blockers of the voltage-gated sodium channel. The saxitoxins are highly toxic to humans and other vertebrates, but not to many invertebrates. The dinoflagellate is ingested by invertebrate filter feeders, including the copepod *C. finmarchicus*. Although it has been shown that there is no increase in mortality in the copepod, it is not known whether the toxic compounds produced by this dinoflagellate might have sub-lethal effects on the copepod. The two primary objectives of my dissertation were to assess the effect of *A. fundyense* on: 1) reproduction, feeding and survival (Chapter 2); and 2) physiology, by analyzing changes in global gene expression (Chapter 3). Adult females were collected from the Gulf of Maine and incubated for a 7-day period on one of three diets: a control with a non-toxic alga as food, a low dose of *A. fundyense* with algal concentration corresponding to natural blooms, or a high dose of *A. fundyense*. Global gene expression using the RNA-Seq technology was used to investigate the biological, cellular and molecular processes that were affected by the introduction of the dinoflagellate into the diet after 2- and 5-days. A third objective was to characterize detoxification enzymes as likely candidates of differentially expressed genes in response to the toxic alga. I identified and characterized the glutathione S-transferase (GST) superfamily using a bioinformatic workflow, and investigated the expression pattern in females exposed to the dinoflagellate (Chapters 4 and 5). The final objective was to

validate expression data produced by the RNA-Seq technology using real-time quantitative polymerase chain reaction (RT-qPCR). Relative expression of selected glutathione S-transferase genes was obtained by RT-qPCR and compared with RNA-Seq results (Chapter 5).

Survival of *C. finmarchicus* was not affected by *A. fundyense* even while the copepod was feeding exclusively on the toxic dinoflagellate for 7 days. However, a reduction in fecundity and a decrease in egg viability were observed in both high and low dose treatments. The physiological (gene expression) response at 2 days involved 2,454 differentially expressed genes and covered a broad range of processes showing many similarities with the cellular (CSR) stress responses described in other organisms. At 5 days, there was an overall decrease in transcriptional activity suggesting a cellular homeostasis response (CHR), which corresponds to a physiological adjustment to a new steady state characterized by fewer differentially expressed genes (1,529). In spite of similar ingestion rates (measured as μg carbon/female/day) among all three treatments, I observed a putative effect on indicators of the energy balance (the allocation of caloric intake to different bodily processes and heat) in *C. finmarchicus* feeding on the toxic diets at both time points. This putative energy reallocation was reflected by strong down-regulation in the expression of genes involved in lipid biosynthesis, growth and reproduction and RNA metabolism in females on the dinoflagellate diets. Surprisingly detoxification was not the most prominent process of the copepod physiological response to the toxic dinoflagellate. Less than 1% of the differentially expressed genes (16 genes out of over 2,000) were regulated (up or down) in detoxification (phases I, II). These results suggest that the physiological response did not include broad up-regulation of cellular detoxification pathways. In addition, the results of the RNA-Seq were supported by RT-qPCR, a technique typically considered the "gold standard" for gene expression studies. In conclusion this experimental study suggests that blooms of *A.*

fundyense could be a stressor for *C. finmarchicus* in the Gulf of Maine, with negative effects on its fitness and energy balance.

2. BACKGROUND

2.1. The increase of harmful algal blooms

Harmful algae blooms (HABs) are dominated by toxic phytoplankton species, which, via the production of natural toxins, have the potential to alter or disrupt marine ecosystems, affecting public health and coastal economies across the planet (Etheridge, 2010; Anderson et al., 2012, 2014). In the past three decades, the impacts of HABs on the public health and the economy (closure of shellfish beds, lost production in fisheries) have increased in frequency, intensity and geographic distribution (Rossini, 2014). An example is the increased global distribution of paralytic shellfish poisoning (PSP), which can be fatal to humans and is caused by the ingestion of toxins produced by the dinoflagellate *Alexandrium fundyense*. Before 1970, PSP cases were limited to temperate waters off Europe, North America and Japan, but by the 1990s PSP events were common throughout the Southern Hemisphere as well (Rossini, 2014). Toxins from harmful algae are transferred to humans through pelagic and benthic organisms, and accumulated at intermediate trophic levels of the marine food web. However, information on their effects is still limited. Thus, the issue of a global increase in harmful algal blooms is a major concern for the marine food web, for the economies of coastal communities that are mostly based on fisheries and also for public health (Anderson et al., 2008).

2.2. HAB of the dinoflagellate *Alexandrium* spp. and paralytic shellfish poisoning (PSP)

Dinoflagellates of the genus *Alexandrium* are among the most important HAB species in terms of severity, diversity, and distribution of blooms (Anderson et al., 2012); its members are widespread globally, with species present in coastal, shelf and slope waters of subarctic, temperate and tropical regions of the Northern and Southern Hemispheres (Anderson et al., 2012). *Alexandrium* spp. produce water-soluble sodium-channel blocker toxins collectively known as saxitoxins (STXs). Saxitoxins are a family of approximately 57 naturally occurring neurotoxic alkaloids, which can be divided into several sub-groups based on their structure and toxicity (Weise et al., 2010). Based on this classification, STXs are divided into gonyautoxin (gonyautoxin1, 2, 3, 4), carbamate (saxitoxin & neosaxitoxin), N-sulfocarbamoyl (N-sulfocarbamoyl- gonyautoxins 1&2, gonyautoxin 5) and decarbamoyl toxins (decarbamoylsaxitoxin, decarbamoylneosaxitoxin, decarbamoylgonyautoxin 2&3) (Weise et al., 2010).

By producing STXs, *Alexandrium* spp. is responsible for outbreaks of potentially lethal paralytic shellfish poisoning (PSP), which is caused by ingestion of seafood containing saxitoxins (Lehane, 2001; Llewellyn, 2006). Upon ingestion of contaminated food, saxitoxins are rapidly accumulated through the gastrointestinal mucosa and progressively inhibit nerve transmission, causing relaxation of smooth muscle, and in extreme cases, paralysis and respiratory failure in a variety of organisms, including humans, seabirds, fishes and marine mammals (Llewellyn, 2006; Weise et al., 2010). Saxitoxins become accumulated in and transported to the higher levels through benthic and pelagic marine food webs. Toxins enter pelagic food webs through ingestion by herbivorous zooplankton and fishes. These organisms serve as toxin vectors by accumulating saxitoxins in their tissue and then being ingested by higher tertiary consumers such as piscivorous fishes and squid, carnivorous gastropods and

crustaceans, marine mammals and birds (Turner and Tester, 1997).

2.3. *Alexandrium fundyense* and *Calanus finmarchicus* in the Gulf of Maine

In the Gulf of Maine, the toxic dinoflagellate *A. fundyense* dominates the annual algal blooms that occur in inshore and offshore waters (Martin and White, 1988; Anderson et al. 2014). Although blooms are highly variable (spatially and temporally), dense regional blooms (10^3-10^4 cell L^{-1}) occur every summer (Anderson, 1997; Deeds et al., 2014). The effect of *A. fundyense* toxins are substantial for the local economies; in the last decade, fishery closures due to the accumulation of STXs in suspension-feeding shellfish have occurred every year (Kleindinst et al., 2014).

During the summer when *A. fundyense* blooms are high, the zooplankton biomass of the Gulf of Maine includes a high abundance of the calanoid *Calanus finmarchicus* (Davis, 1987; Sherman et al., 1987; Meise and O'Reilly, 1996; Miller et al., 1998; Durbin et al., 2000). The *C. finmarchicus* population is dominated by the second generation (G1) represented mostly by mid- and late-stage copepodites (CIV and CV) that are either preparing for diapause or maturing directly into adults and producing a third generation (G2) (Miller et al., 1998; Durbin et al., 2000; Fiksen, 2000). *C. finmarchicus* is important for the local economy, mainly because it is consumed by many economically important fishes such as cod, mackerel and herring (Runge, 1988; Darbyson et al., 2003).

During bloom conditions *C. finmarchicus* ingest *A. fundyense* as measured in multiple field and laboratory studies (Campbell et al., 2005; Turner and Borkman, 2005; Turner et al., 2005; Turner, 2006; 2010), even when the dinoflagellate is not the only food source (Teegarden et al., 2001; 2008; Hassett, 2003). However, those aforementioned laboratory studies were

mostly limited to exposure times of approximately 24 hr (Teegarden et al., 2001; 2008; Hassett, 2003; Turner and Borkman, 2005; Turner, 2010). Surprisingly, although the interaction between the copepod and *A. fundyense* has been extensively studied, none of these studies investigated the possibility of sub-lethal effects on *C. finmarchicus* (Turner, 2014).

2.4. Effect of *A. fundyense* blooms on copepods

Effects of algal toxins on their primary grazers, zooplankton filter feeders, are variable, in some cases having clear deleterious effects on these animals, but in other cases, having little or no obvious effects (reviewed in Turner, 2014).

In general, copepods are highly resistant to algal toxins, with no effect on the survivorship in response to harmful algae diets (Teegarden et al., 2001; Colin and Dam, 2002; Kozlowsky-Suzuki et al., 2006; Turner and Borkman, 2005). Moreover, these animals may retain and accumulate toxins to become vectors for the transfer of toxins to higher trophic levels, including fishes, whales and humans (Turner and Tester, 1997; Campbell et al., 2005; Doucette et al., 2005; 2006; Petitpas et al., 2014). Laboratory studies have shown that the copepods *Acartia tonsa*, *Centropages hamatus* and *Eurytemora herdmani* are able to avoid *Alexandrium* spp. when the dinoflagellate represented only a small proportion of the diet (Turriff et al., 1995; Teegarden, 1999; Teegarden et al., 2001; Tang et al., 2001; Dutz et al., 1998). However, field studies suggest that copepods are ingesting *Alexandrium fundyense* with no selective avoidance even when the dinoflagellate represents a small fraction of the total assemblage of natural phytoplankton (reviewed in Turner, 2014).

Although survival rate is high, there is some evidence that *A. fundyense* affects the overall fitness of invertebrate filter feeders. Sub-lethal effects, such as reduced feeding, low egg

production and low egg hatching success, have been reported for several copepods (Sopanen et al., 2011; Dutz, 1998; Guisande et al., 2002; Colin and Dam, 2007; Teegarden et al., 2008; Turner, 2014). For example, low ingestion rates were reported for *A. tonsa* and *E. herdmani* feeding on *A. fundyense* (Teegarden and Cembella, 1996), in which *Acartia clausi* responded to exposure to *Alexandrium* spp. with decreased growth and lower fecundity. This suggest these copepods either reduced caloric intake (low feeding activity) or possibly reallocated energy into detoxification (Dutz, 1998; Frangópulos et al., 2000; Guisande et al., 2002). Biogeographic factors limit the interaction between the copepod *Acartia hudsonica* and *A. fundyense*; populations that had been historically exposed to the dinoflagellate (Gulf of Maine) had higher fitness (grazing, egg production, egg hatching, growth, survival) when tested in the lab than copepods from areas (e.g. New Jersey) that had little or no historical exposure (Dam, 2013). It was hypothesized that differences in STX resistance between the two populations were due to a mutation in the voltage-gated Na channel however the physiological mechanism that underlies this difference in resistance unknown (Chen et al., 2015).

2.5. Cellular stress and homeostasis responses

Organisms activate cellular stress-defense systems with the aim to prevent physiological damage caused by the stressor such as high/cold temperature, toxic chemicals or diseases (Van Straallen and Roelofs, 2006; Stillman and Hurt, 2015). Stress responses are well studied in eukaryotes that share common features among many organisms. The stress response is transient. After the introduction of an environmental challenge there is an immediate activation of a uniform typically non-specific rescue operation known as the “cellular stress response” (CSR) that involves large-scale transcriptional activity. This response activates proteins involved in

redox regulation, DNA damage and protein degradation control, lipid and energy metabolisms and molecular chaperoning (Kültz, 2003, 2005). The cellular stress response is followed by a second set of stressor-specific adaptations, commonly referred to as the "cellular homeostasis response" (CHR), which leads to a physiological adjustment toward a new "steady state" (Kültz, 2005). Both the CSR and CHR require energy, which is achieved via the up-regulation of genes involved in energy-generation functions (e.g. carbohydrate and lipid catabolism) and the down-regulation of growth-related functions, such as RNA metabolism, reproduction and cell-proliferation (Kültz, 2003, 2005). Stress responses are also known to be graded to the severity of the stressor (Gasch et al., 2000); in studies testing severe conditions, observed differences in expression may reach a 100-fold change compared to the control level. In contrast, when organisms are exposed to more realistic sub-lethal conditions, differential expression may be subtle and modest. For example, in a study investigating gene expression of a eurythermal fish, *Austrofundulus limnaeus*, in response to temperature changes that mimic natural environmental fluctuations, differences in gene expression typically ranged between 2- and 4-fold. This response indicates a tight regulation of steady state levels of mRNA transcripts during temperature changes that *A. limnaeus* was adapted to experience in its environment (Podrabsky and Somero, 2004). Similarly, in several ecotoxicology studies when the cladoceran *Daphnia* spp. was exposed to stress conditions that did not affect survival (e.g. 1/10 LC50), the majority of differentially expressed genes showed only a 1- to 4-fold change compared with the control levels (Poynton 2007; 2012). These studies suggest that even mild stressors can affect the overall physiology of an organism, and that these effects can be detected using gene expression techniques. Given that blooms of *A. fundyense* are common in the Gulf of Maine, I expected the magnitude of response of *C. finmarchicus* to the dinoflagellate, in particular to the low dose

treatment, to be equivalent to the these modest responses.

2.6. Use of transcriptomics in physiological ecology studies

In the last decade, high-throughput nucleotide sequencing technology has been used to study the cellular stress response in many organisms, including non-model species (e.g. Kültz, 2003, 2005; Somero, 2012). The advent of RNA-Seq technology and the development of software programs that assemble short sequences have made it possible to develop genomic and transcriptomic resources for non-model species (Riesgo et al., 2012; Lenz et al. 2014). Resources developed with this technology have been used in gene expression studies to investigate the physiological response of non-model species to a variety of environmental stressors (Wang et al., 2009). Advantages of this approach include high-throughput sequencing of short reads that allows for a global assessment of the physiological response, and detection of many biological processes that might be regulated simultaneously (Mortazavi et al., 2008).

RNA-Seq provides a snapshot of all the mRNA expressed in an organism/tissue at the time of RNA extraction. The underlying assumption of these gene expression studies is that differences in mRNA serve as a proxy for differences in protein levels that are then manifested in different phenotypes (Gry et al., 2009). However, while mRNA expression levels commonly function as a good proxy corresponding to changes in protein expression levels (Liao and Weng, 2015), the correlation between mRNA and protein expression can be variable (Gry et al., 2009). For instance, changes in gene transcription related to the experimental treatment may either represent an increase of protein level associated with the stressor, or may simply represent the attempt to maintain the current level of protein against other changes in protein translation or degradation (Podrabsky and Somero, 2004). However, changes in transcript abundance, whether

they reflect an increase in the protein level or not, are likely to reflect changes in the physiology but one must be cautious when interpreting these types of data when no direct measurements of protein expression or other physiological parameters have been made.

Although the application of RNA-Seq technology represents a powerful tool for many ecological studies, data interpretation still remains a challenge. RNA-Seq is a relatively new technology, and bioinformatics tools are still being developed to aid in data processing and interpretation. Although well-suited for investigating the physiological ecology of non-model species, there are challenges in generating reference transcriptomes with valid "annotation", that is determinations of the identities of the genes uncovered. Furthermore, even if differentially expressed genes are annotated, the physiological significance of a modest change in expression is not always clear. Nevertheless, the application of RNA-Seq technology has recently provided new insights in the mechanisms that regulate the physiological response of many organisms to a variety of environmental challenges (Stillman and Armstrong, 2015). The fluctuations in marine ecosystem caused by environmental changes (e.g. temperature, ocean acidification, pollutions and HABs) require a better understanding of the physiological ecology on non-model organisms (e.g. copepods) that play critical roles in the pelagic food web. Thus, gene expression studies using RNA-Seq represent powerful tool to investigate interactions between organisms and the environment (Stillman and Armstrong, 2015; Beaugrand, 2015).

2.7. Calanus finmarchicus de novo transcriptome

Relative gene expression has been used to investigate biological, cellular and molecular processes that are regulated in *C. finmarchicus* developmentally, environmentally, seasonally and/or under experimental conditions (Lenz et al., 2012, 2014; Tarrant et al., 2008, 2014; Unal et

al., 2010; Aruda et al. 2011). With the rapid advancement in high-throughput sequencing technologies and whole transcriptome profiling (RNA-Seq), two *C. finmarchicus de novo* assemblies have been generated in the past two years (Lenz et al., 2014; Tarrant et al., 2014). In this study, I used the *C. finmarchicus de novo* transcriptome generated from individuals from the Gulf of Maine population (Lenz et al., 2014). The *de novo* transcriptome was produced from mRNA collected from six developmental stages: embryo, early nauplius (NI-II), late nauplius (NV-VI), early copepodite (CI-II), late copepodite (CV) and adult (CVI) female. The reference transcriptome consisted of 96,090 transcripts, of which 40% were annotated using Gene Ontology (Lenz et al., 2014). Overall, the coverage of the Gulf of Maine transcriptome was estimated to be 65% or better (Lenz et al., 2014).

In summary, *A. fundyense* blooms represent non-optimal conditions for *C. finmarchicus*. Under experimental conditions, I observed a negative impact on fecundity and energy balance in adult females fed on either high or low doses of *A. fundyense*. The significance of the project consists of demonstrating that in spite of no difference in long-term (7 days) survival, the dinoflagellate diets had a major effect on the biology of the copepod. The results of the low dose treatment suggest that *A. fundyense* blooms may have a significant effect on *C. finmarchicus* in the Gulf of Maine. During summers of intense and large-scale blooms, there may be a significant negative impact on the population dynamics of *C. finmarchicus* as the second generation starts to reproduce. In addition, this is one of the few studies in which the effect of toxic algae on the copepod has been investigated by combining the traditional fitness measurements (egg production, grazing, survival) with global gene expression studies as a proxy of the physiological response.

Chapter 2

The effect of the toxic dinoflagellate *Alexandrium fundyense* on the fitness of the calanoid copepod *Calanus finmarchicus*

ABSTRACT

Inshore and offshore waters of the Gulf of Maine (USA) have spring/summer harmful algal blooms (HABs) of the toxic dinoflagellate *Alexandrium fundyense*, which is responsible for paralytic shellfish poisoning (PSP) in humans. The calanoid copepod *Calanus finmarchicus* co-occurs with *A. fundyense* during the seasonal blooms. At that time, *C. finmarchicus* population abundances are high, dominated by immature copepods preparing for diapause and by actively reproducing adults. High survival has been reported for copepods exposed to toxic *A. fundyense*, but little is known about possible sub-lethal effects of toxins on these animals. In this study, *C. finmarchicus* adult females were fed either a control diet of non-toxic *Rhodomonas* spp. or one of two experimental diets containing either low (LD) or high (HD) doses of toxic *A. fundyense* for a total of 7 days in two independent experiments. The effect of the dinoflagellate on the copepod fitness was measured using survival rate, grazing rate and reproductive success. LD and HD cell concentrations were 50 and 200 cells mL^{-1} respectively. As expected, ingestion of the dinoflagellate had no effect on survival and grazing activity. However, significant reduction of egg production and viability was measured for *C. finmarchicus* females fed either of the experimental treatment. After the 7-day experiment, total naupliar recruitment of the females fed with LD and HD was reduced by 35% to 75%, respectively, compared with the control females. These results suggest that blooms of *A. fundyense* can represent an environmental challenge for *C. finmarchicus* populations, with a potential negative effect on copepod recruitment.

1. INTRODUCTION

Harmful algal blooms (HABs) dominated by the dinoflagellate *Alexandrium fundyense* occur annually in offshore and inshore waters of the Gulf of Maine (Martin and White, 1988; Shumway et al., 1988; Anderson et al. 2005; 2014). Although blooms are highly variable both spatially and temporarily, dense regional blooms (10^3-10^4 cell L^{-1}) of *A. fundyense* occur every summer (Anderson, 1997; Deeds et al., 2014; McGillicuddy et al., 2014). By producing saxitoxins (STXs), potent neurotoxins, *A. fundyense* is responsible for outbreaks of paralytic shellfish poisoning (PSP), which is potentially fatal to humans (Llewellyn, 2006). Although highly toxic to most vertebrates, numerous invertebrate herbivores, such as copepods and shellfish, are able to ingest the toxic algae without affecting their survival (Shumway, 1990; Bricelj and Shumway, 1998; Teegarden et al., 2003; Petitpas et al., 2014). These invertebrates retain and accumulate toxins in their tissues, becoming vectors for transfer of the toxins to higher trophic levels, including fishes, whales and humans (Anderson and White, 1992; Turner and Tester, 1997; Campbell et al., 2005; Doucette et al., 2005; 2006; Petitpas et al., 2014). It is less clear whether and how *A. fundyense* affects the overall fitness of these invertebrate filter feeders (Turner, 2014).

Calanus finmarchicus is one of the more abundant calanoid copepods in the North Atlantic, extending from the mid-Atlantic Shelf off the US east coast to the Barents Sea north of Norway (Conover, 1988; Planque et al., 1997). In the Gulf of Maine, the *C. finmarchicus* population increases during the spring (Davis, 1987; Meise and O'Reilly, 1996) with the copepod serving as food for planktivorous fishes such as larval herring and mackerel (Darbyson et al., 2003). During the summer, when *A. fundyense* blooms are present, *C. finmarchicus* dominates the zooplankton biomass (Davis, 1987; Sherman et al., 1987), reaching abundances of

10,000 m^3 or higher (Meise and O'Reilly, 1996). During this period the *C. finmarchicus* population is dominated by mid- and late-stage copepodites (Miller et al., 1998; Durbin et al., 2000). Not surprisingly, during bloom conditions *C. finmarchicus* and other zooplankters ingest *A. fundyense* as measured in multiple field studies (Campbell et al., 2005; Turner and Borkman, 2005; Turner et al., 2005; Turner, 2006; 2010). Survival in these studies, as well as in laboratory grazing studies, was high even when *A. fundyense* was the only food source (Teegarden et al., 2001; 2008; Hassett, 2003). However, previous grazing studies were mostly limited to incubations of approximately 24 hr (Teegarden et al., 2001; 2008; Hassett, 2003; Turner and Borkman, 2005; Turner, 2010). In addition, none of these studies investigated the possibility of other effects on *C. finmarchicus* (Turner, 2014).

Studies on other copepods have shown that *Alexandrium* spp. can have adverse effects such as low feeding rates, low egg production and low egg hatching success (Sopanen et al., 2011; Dutz, 1998; Guisande et al., 2002; Colin and Dam, 2007; Teegarden et al., 2008; Turner, 2014). Reduced feeding rates have been measured for *Acartia tonsa* and *Eurytemora herdmani* fed on *A. fundyense* at very high concentrations (ca. 500-2000 cell mL^{-1}) (Teegarden and Cembella, 1996), while *Acartia clausi* responded to exposure to *Alexandrium* spp. with decreased growth and fecundity, suggesting either a reduced caloric intake (low feeding activity), reduced food assimilation, or a possible reallocation of energy into detoxification (Dutz, 1998; Frangópulos et al., 2000; Guisande et al., 2002). Overall, these sub-lethal effects, induced by *Alexandrium* spp. in different copepod species, suggest that the presence of the toxic dinoflagellate could be a significant environmental stressor. Thus, HABs could disrupt existing pelagic communities in ways that are difficult to predict based on studies focused solely on survivorship.

The goal of the present study was to better understand if *A. fundyense* is an environmental challenge for *C. finmarchicus* with negative effects on copepod fitness. Adult female survival, feeding rates and reproductive success were monitored during 7-day-long experiments in which the animals were maintained on three different diets: a control with no *A. fundyense*, a low dose (LD) with (25:75 by cell volume *A. fundyense* and *Rhodomonas* sp.) and a high dose (HD) with 100% *A. fundyense*. The experimental algal concentration for the LD was comparable to bloom conditions in the Gulf of Maine (Anderson, 1997; McGillicuddy et al., 2014; Petitpas et al., 2014).

2. MATERIALS AND METHODS

2.1. Field collection and maintenance of *Calanus finmarchicus*

Calanus finmarchicus were collected in June and July 2012 in the Gulf of Maine near Mount Desert Rock (Lat: 44° 2'N; Long: 68°3'W) by slowly towing a 75 cm diameter (560 µm mesh) net vertically from 75 m depth to the surface. Plankton collections were immediately diluted into buckets containing 10 L of subsurface seawater and placed on ice in coolers for transportation to the Mount Desert Island Biological Laboratory (Salisbury Cove, ME) for the experimental procedures (generally within 3 hours of collection). Healthy adult females and adult males were sorted from the diluted plankton samples and transferred into 3.5 L jars of filtered seawater (FSW) with 15-20 individuals per jar (1:3 male and females) with *Rhodomonas* sp. added *ad libitum,* and placed overnight into an incubator (Percival Model I-36VL, Percival Scientific, Inc., Perry, IA, USA) maintained at 10°C on a 14:10 h light:dark cycle.

2.2. Experimental design

Calanus finmarchicus adult females were fed with the toxic *A. fundyense* over a 7-day period; one experiment was performed in June and replicated in July (Table 2.2). The experimental design included three treatments: control, low dose and high dose of *A. fundyense* (Table 2.1). The control group consisted of a monoalgal diet of the non-toxic flagellate *Rhodomonas* sp. (8000 cells mL^{-1}), which does not appear to impair copepod egg production and hatching success and is routinely used for maintenance of copepods in culture settings (Helland et al., 2003). The low dose treatment (LD) consisted of a mixed diet of *A. fundyense* (50 cells mL^{-1}) and *Rhodomonas* sp. (6000 cells mL^{-1}), which corresponded approximately to a 1:3 proportion by algal volume. The "high dose" treatment (HD) corresponded to 100% of *A. fundyense* at a concentration of 200 cells mL^{-1}. Although nutritional profiles can be variable, comparable lipid and amino acid profiles have been reported for *Rhodomonas* spp. and *A. fundyense* (Ianora et al., 2004; Seixas et al., 2009). Equivalent carbon concentrations were computed for each treatment for *Rhodomonas* sp. and *A. fundyense* based on measurements of carbon to volume relationships from Menden-Deuer and Lessard (2000) (Table 2.1). Carbon concentrations for the experimental food levels were similar, albeit slightly higher in the HD treatment (Table 2.1). Furthermore, carbon content per cell for *A. fundyense* can be highly variable depending on growth and culture conditions (Anderson et al., 1990). Thus, the carbon values provided in Table 2.1 are estimates. In addition, dinoflagellate toxin content (= toxin cell quota) can vary by a factor of 2 or more within a period of 24 hours (Anderson et al. 1990). Therefore, during the duration of the June and July experiments *A. fundyense* samples were collected daily for toxin content analysis (see section 2.3, Table 2.2).

2.3. Phytoplankton and *A. fundyense* toxin profile

The flagellate *Rhodomonas* sp. was used for the control diet in both experiments, however, the clones differed between June and July. The *Rhodomonas* sp. clone (LOT#120406) used in June was obtained from ALGAGEN LLC with a starting volume of 4 L that had an initial cell density of 10^6 cells mL^{-1}. Cell densities during the experimental week ranged from 0.8×10^6 to 1.6×10^6 cells mL^{-1} in the stock culture. The clone (CCMP739) was used in July. It was obtained from NOAA/NMFS (CT, USA, isolated in 1951 by R. Lasker) with initial cell density of 10^6 cells mL^{-1} in cultures of 4L. During the experiment, cell densities in the stock culture ranged from 1×10^6 to 2×10^6 cells mL^{-1}. For both *Rhodomonas* sp. clones, cultures were maintained at 15-16 ºC and ambient outside light and diluted by 50% with f/2 medium every three days (Guillard, 1973). The toxic dinoflagellate *A. fundyense* (clone GTCA28) was isolated from the western Gulf of Maine in 1985 and maintained at 15 ºC on a 14:10 h light:dark cycle. One liter volumes of culture were initially grown in modified f/2-Si medium (Anderson et al., 1994) at 15 ºC and then transferred to 10 ºC during mid-exponential growth for temperature equilibration prior to shipment. Upon receiving the *Alexandrium* culture, it was grown at 10 ºC on a 14:10 h light: dark and diluted by 50% every two days with f/2-Si medium (Guillard, 1973). During the experiments, cell densities in the stock cultures ranged from 17×10^3 to 32×10^3 cells mL^{-1} in June and 14×10^3 to 19×10^3 cells mL^{-1} in July. Stock cultures of *Rhodomonas* spp. and *A. fundyense* were checked 3 times per week between experiments and daily during experiments to assure that cells looked healthy and were swimming actively.

Food suspensions for the experimental treatments were prepared daily by diluting algal cells from the stock cultures into filtered seawater (FSW) at the target concentrations shown in Table 2.1. For the toxin analysis, three replicate samples of *A. fundyense* cells were obtained from the stock culture (1.5 mL per sample), transferred into Eppendorf tubes, centrifuged for 8

minutes at 3000 rpm and the supernatant removed. Subsequently, 0.5 mL of 0.05 M acetic acid was added to the pellet and each sample was homogenized using a pipette tip. The sample was shaken twice, stored immediately at 4ºC and transported on ice to Woods Hole Oceanographic Institution where the samples were sonicated in an ice water bath using a Branson Sonifier 250 D fitted with a micro-tip probe at a constant 40-watt output for 1 minute. This extract was then stored at – 20 ºC. Prior to analysis, the samples were thawed, centrifuged for 10 minutes at 3000 x g, and 200 µL of the supernatant was added to a limited volume autosampler vial. Toxin analyses were carried out using a modification of the Oshima (1989) post-column derivatization HPLC method (Anderson et al., 1994). Certified reference standard solutions purchased from the National Research Council (Halifax, Nova Scotia Canada), containing toxins C1, C2, GTX1–5, dcGTX2, 3, NEO, dcSTX and STX were run at the beginning of the sample queue and following every 4th sample. The resulting toxin content values were expressed in micromolar concentrations and these were used to calculate cellular toxin content on duplicate or triplicate measurements for each experimental time point. Abbreviations used through this text are: STX= saxitoxin; NEO= neosaxitoxin; GTX1, 4= gonyautoxins 1 and 4; GTX2,3= gonyautoxins 2 and 3; dcGTX2&3= decarbamoyl gonyautoxin 2 and 3; GTX5= gonyautoxin 5 (or B1, Hall 1982); C1&2 = toxins C1 and C2; dcSTX= decarbamoyl saxitoxin). The concentrations of toxins GTX1&4, GTX2&3 and C1& 2 were combined to account for possible epimerization of the toxin pairs. Toxicities (in STX equivalent cell^{-1}) were calculated from molar composition data using individual potencies provided by the NRC and the daily toxicity level was calculated following (Anderson et al., 1990). The calculation was based on the concentration of individual toxins and their specific toxicity in µg STX eq. µmol^{-1} as follows: C1, 2.61; C2, 41.6; GTX1, 429.4; GTX2, 155.2; GTX3, 275.6; GTX4, 313.7; GTX5, 27.8; dcGTX2, 66.5; dcGTX3, 162.7;

NEO 399.3; dcSTX, 221.7; STX, 432.0.

2.4. Egg production, fecal pellet production, egg viability and naupliar production

Thirty mature female *Calanus finmarchicus* were randomly selected from the stock jars (see above) after overnight acclimation and incubated individually in 100 mL tissue flasks in one of the three treatment food suspensions, and maintained with the same treatment for the next seven days. The experiment was performed in June and then replicated in July (Table 2.2). Every day, each female was transferred with a pipette into a new container, and supplied with a fresh food suspension. The jars from the previous 24 hours with the eggs and the fecal pellets produced by each female were incubated for another 24 hours in the 10ºC incubator to allow eggs to hatch, and then were preserved in 70% ethanol. The contents of the jars were then checked under a dissecting microscope and nauplii, unhatched eggs and fecal pellets were counted. The mass of each fecal pellet was estimated to be the same among the different treatments based on size. Nauplii were checked for possible malformations (none were found). Eggs were also checked for evidence of cannibalism (crumpled egg membrane, none were found).

2.5. Grazing activity

Daily grazing experiments were performed during the July experiment. Nine sets of two mature and healthy *C. finmarchicus* adult females were randomly selected from the stock jars (see section 2.1) and transferred into crystallizing dishes with 200 mL of one of the three treatment diets. Each treatment consisted of three replicates with food suspensions prepared daily by performing cell counts in stock cultures and adjusting the

concentration by dilution with FSW. Each day females were transferred into new containers with new experimental food suspensions. For each treatment two additional containers without copepods: a grazing control and an initial were prepared; the initial, representative of daily food suspensions was preserved with Lugol's solution (2%) at the beginning of each daily experiment while the grazing control was treated like the experimental jars and incubated in a Percival Model I-36VL Incubator System (Percival Scientific, Inc., Perry, IA, USA) at 10°C on a 14:10 h light:dark cycle at 50 μm $m^{-2}s^{-1}$. After a 22.5-24 hour incubation period, and the transfer of the females to new food suspensions, the contents of the containers (1 grazing control and 3 experimental per treatment) were preserved in 2% Lugol's solution. Phytoplankton cell numbers in preserved aliquots were counted in Sedgwick-Rafter cells, with a minimum of 400 cells counted in all cases, assuring ca. 10% precision (Guillard, 1973). Computed ingestion and clearance rates of *C. finmarchicus* were determined from differences in phytoplankton cell concentrations in initial, grazing control and experimental suspensions using the formulae described by Frost (1972). Briefly, daily copepod feeding activity was obtained by counting the number of algal cells left in the container with the two mature and healthy females, compared to the number of algae cells in the grazing control containers (no copepod) and corrected for algal growth (initial container counts).

Dead copepods were removed (1 dead female in C and LD and 2 in the HD) and replaced with new healthy females that had been kept in reserve containers (2 females per container) exposed to the same experimental conditions, including the daily transferring in new containers with fresh food suspensions. Grazing rates for dead copepods, were calculated assuming that they had lived for 1/2 day.

2.6. Survival

Female survival rate was monitored daily in all experimental incubations. In the "reproductive success" experiments (section 2.4) in June and July 10 females per treatment were checked under the microscope to assure they were undamaged and healthy (actively swimming) before their daily transfer into new containers. Daily survival rate was recorded. In addition, in both June and July three replicates of two females per treatment were kept in 200 mL of food suspensions, and checked daily for swimming activity before being transferred to fresh seawater with new food. This second set of females was the one monitored for grazing activity in July (see section 2.5). Dead females, if any, were replaced with new healthy females (see section 2.5); thus survival rate in the 200 mL containers was calculated each day as the number of females still alive before the daily transfer to new food suspensions.

2.7. Statistical analysis

Statistical analyses were performed using the software Prism Graph Pad (v 6.0). Two-way analysis of variance (ANOVA) followed by *Tukey* test for post-hoc multiple comparisons ($P<0.05$) was performed to test for differences among treatments and over time for each measured fitness parameter using 10 replicates for egg production (see section 2. 6) and 6 replicates grazing (see section 2.5) experiments, respectively.

3. RESULTS

3.1. *Alexandrium fundyense* toxicity levels during the experiments

Toxin measurements for *A. fundyense* confirmed that the dinoflagellate in the cultures was neurotoxic during both the June and July experiments. No significant differences in toxin composition (molar % of total toxin) and total toxin content were found between the two experimental periods. Of the more than 20 naturally-occurring STX derivatives (Llewellyn, 2006), the *A. fundyense* culture contained nine derivatives represented by N-sulfocarbamoyl toxins (C1&2), gonyautoxins (GTX1&4, GTX2&3), decarbamoyl toxins (dcGTX3) neosaxitoxins (NEO) and saxitoxin (STX). The major toxins present (in order of relative abundance as molar % of total toxin) were N-sulfocarbamoyl toxins, followed by gonyautoxins, neosaxitoxins and saxitoxin (Figure 2.1A). The average toxin content (all derivatives) during the 7-day experiment was not significantly different between June and July (Student's t-test; P = 0.36) (Figure 2.1B). During the June experiment, toxin content ranged between 0.01 and 0.03 ng STX equivalents cell^{-1}, and there were no significant differences among experimental days. In July, toxin content was constant between Days 1 and 5 (mean: 0.014 ng STX equivalents cell^{-1}) (Figure 2.1C). However, significantly higher toxin content was measured for Day 6 and Day 7 (Figure 2.1C). The higher toxicities corresponded to changes of NEO and STX derivatives, which showed a 2- and 10-fold increase, respectively, compared with previous days. The changes in the amount of NEO and STX also affected the relative proportion of the different toxins (% mol), as the proportion of NEO increased from 30% to 40% and STX from 1% to 10%, while N-sulfocarbamoyl and gonyautoxins toxins decreased to 27% and 23% respectively.

C. finmarchicus females were exposed to mean toxicities respectively of 1.0 and 0.7 ng STX equivalents mL^{-1} for June (Days 1-7) and July (Days 1-5) in the LD treatment; mean toxicities for the HD treatment were 4-fold higher (4 and 2.8 ng STX equivalents mL^{-1} June and July respectively). In the July experiment, Day 6 and Day 7 had higher toxicities (see above)

respectively of 1.4 and 1 ng STX equivalents mL^{-1} for the LD and 5.5 and 4 ng STX equivalents mL^{-1} for the HD treatment. Thus, the dinoflagellate's toxicity in terms of toxin profile and toxin equivalency was similar throughout the two experiments, with the exception of the last 2 experimental days in July.

3.2. Grazing activity

3.2.1. Ingestion rates

Calanus finmarchicus ingested *A. fundyense* in both LD and HD treatments over the 7-day experimental period with no significant differences compared with the control diet (Figure 2.2A). Raw data of phytoplankton cells obtained each day from the counts of the initial, grazing control and experimental containers are listed in Supplemental Table 2.1. Computed ingestion rates converted to μg C $female^{-1}$ h^{-1} were similar among the three treatments with no significant changes over time (Figure 2.2A). Ingestion rates averaged 1 μg C $female^{-1}$ h^{-1} (±0.1SD) for control, LD and HD treatments (Figure 2.2A). Those values are highly comparable to the mean ingestion rates reported for *Calanus* spp. adult females feeding on algae of similar cell size (Frost, 1972). Daily ingestion of toxins was computed as STX equivalents per day from the toxicity measurements (see section 3.1). STX ingestion for *C. finmarchicus* females feeding on the LD and HD did not change over time for the first 5 days of experiment (Figure 2.2B). *C. finmarchicus* ingested on average 0.3 ng STX equivalents day^{-1} (±0.1SD) in the LD and 2 ng STX equivalents day^{-1} (±1SD) in the HD treatments between Day 1 and 5 (Figure 2.2B). On Days 6 and 7, cell toxicity levels were higher (Figure 2.2B) with corresponding higher STX ingestion rates in both LD and HD treatments (Figure 2.2B). Average daily carbon ingestion rates during the 7-day experiment were comparable among the treatments (Figure 2.2C); in the

LD treatment that contained a mixture of *Rhodomonas* sp. and *A. fundyense*, *C. finmarchicus* females consumed the toxic dinoflagellate in proportion to its concentration (Figure 2.2C), showing no evidence of selective feeding.

In conclusion, *C. finmarchicus* carbon ingestion rates in all three treatments (control, LD and HD) were comparable, excluding the possibility that the copepods, when incubated with the dinoflagellate as prey were non-feeding.

3.2.2. Fecal pellet production

Fecal pellet production during the 7-day experiment in June was not significantly different between the control and LD treatments while fecal pellet production was significantly lower in the HD treatment from Day 1 to Day 3 (Figure 2.2D). An overall decrease of the number of fecal pellets was observed over time in all treatments (Figure 2.2D); however, this decline was significant only for the control and LD treatment (Table 2.3). Cumulative fecal pellet production confirmed the significantly difference for females on the HD treatment compared with the LD and the control diets. Total fecal pellets production for females feeding on the HD treatment became significant lower on day 2 and reminded so for the experiment.

In the July experiment, fecal pellet production during the 7-day experiment was not significantly different between the control and the two toxic treatments (Figure 2.2E). Fecal pellet production rates were variable and declined by Day 3 in all three treatments, although this decline over time was only significant in the control (Figure 2.2E, Table 2.3). Cumulative fecal pellet production during the 7-day July experiment was not significantly different for females on the LD and HD diets compared with the fecal pellet production of females on the control diet.

3.3. Reproductive success

3.3.1. Egg production in the control

Daily egg production rates were highly variable over time with a decline observed in the controls in both experiments (Figure 2.3A,C). This decline was significant in the control and the experimental treatments (Figure 2.3A,C Table 2.3). The decline of *C. finmarchicus* egg production over time has been reported for other laboratory studies including those with similar experimental food levels (Hirche and Kwasniewski, 1997; Båmstedt et al., 1999; Niehoff et al., 2000; Jansen et al., 2006; Madsen et al., 2008). Thus, the decrease in egg production observed here for the control females feeding was expected. Moreover, the initial egg production rates (Day 1) for both June and July experiments are similar to the ones measured for *C. finmarchicus* collected in the Western Atlantic during a similar period (June <10 eggs female^{-1} d^{-1} - Plourde and Runge, 1993; July: 30-50 eggs female^{-1} d^{-1}; Runge et al., 2006; Melle et al., 2014). Initial egg production in June averaged 13 eggs female^{-1} d^{-1} and 50 in July eggs female^{-1} d^{-1} Furthermore, during the July experiment, the total number of eggs produced by the control females during the 7-day experiment (136 eggs female^{-1}) (Figure 2.3D) is comparable to egg production rates reported in other studies for similar incubation periods (100-150 eggs female^{-1} week^{-1}– Hirche and Kwasniewski, 1997; Båmstedt et al., 1999).

3.3.2. Effect of A. fundyense *on egg production*

Alexandrium fundyense negatively affected *C. finmarchicus* egg production rates. Egg production rates were not significant different among the three treatments on any specific day, but the cumulative number of eggs produced was lower for females in the HD treatment for both June and July (Figure 2.3).

In the June experiment, despite the low production of eggs in all treatments, the effect of the dinoflagellate was significant in the HD treatment (Figure 2.3A). Daily egg production rates were highly variable and there were no statistically significant differences among treatments on any given day (Figure 2.3A). A decline in egg production was observed starting on Day 2 in all treatments, although this decline was only significant in the control (Figure 2.3A, Table 2.3). Significant differences were observed in cumulative egg production between the HD and control, and also between HD and LD, but not between the LD and control treatments (Figure 2.3B). Total egg production per female on the HD diet was significantly lower starting on Day 4 and thereafter (Figure 2.3B). By Day 7, the cumulative number of eggs produced by females in the HD treatment averaged 10 (±2SD), and 29 and 27 (±4SD), in the control and LD treatments (Figure 2.3B). Thus during the 7-day exposure to *A. fundyense*, adult females in the HD treatment produced only 34% as many eggs as the control females.

In July, cumulative egg production during the 7-day (as distinguished from daily numbers) was significantly lower for females on the LD and HD diets compared with egg production of females on the control diet (Figure 2.3D). This difference was significant starting on Day 4 for females in both the LD and HD treatments, and remained significantly different for the remainder of the experiment (Figure 2.3D). For the duration of the experiment, females fed on *A. fundyense* diets produced only 52% (LD) and 53% (HD) of the total number of eggs produced by females on the control diet.

3.3.3. Egg viability and naupliar production

The negative effect of the dinoflagellate was even more pronounced in the hatching success. The number of viable eggs produced by females on both LD and HD diets of *A. fundyense* was reduced compared with eggs produced by females on the control diet in both experiments (Figure 2.4).

In the June experiment, hatching success was more variable and significant differences in hatching success between the control and LD and HD were observed on Days 1 through 4 but not thereafter (Figure 2.4A). Specifically, on Days 1 and 2 both treatments had significantly lower hatching success, while on Day 3 only the HD treatment was significantly lower, with a reversed pattern on Day 4 (Figure 2.4A). Hatching success for control females on Day 1 was high but declined thereafter (Figure 2.4A). Between Days 5 and 7 hatching success was very low in all treatments including the control and averaged 17 and 18 % viable eggs for control and the two toxic treatments respectively, with no significant difference between treatments (Table 2.3). The decline in egg viability during the experiment was significant in all three treatments (Table 2.3).

During the July experiment, egg viability (% hatching) for *C. finmarchicus* females fed *A. fundyense* (low and high doses) was significantly lower by Day 2 compared with the control diet, and remained significantly different for the remainder of the experiment (Figure 2.4B). Egg hatching success in the control was near 100% during the first two days, and then gradually declined to 42% by Day 7, and this decline was significant (Figure 2.4B, Table 2.3). Hatching success in the LD and HD treatments were not significantly different from each other, and for both LD and HD declined from 68-77% to 5-4% between Days 1 and 5, and 0 on Days 6 and 7 (Figure 2.4B). The decline in egg viability over time was significant for all three treatments (Table 2.3).

Effective recruitment was computed as the number of healthy nauplii produced during the 7-day experimental period for each female that survived the entire 7-day experiment. In both experiments a reduction was observed in the number of healthy nauplii produced by females fed with both the LD and HD diets compared with the control (Figure 2.4C). In June, this difference was only significant for the HD treatment while in July, the average number of nauplii per female was significantly lower in both LD and HD treatments (Figure 2.4C). In June on average, females in the LD and HD produced respectively 65% and 25% of the number of healthy nauplii produced by control females (Figure 2.4C). In July, females feeding on LD and HD treatments both produced on average 40% of the number produced by the control females (Figure 2.4C).

3.4. Survival

Calanus finmarchicus showed good survival rates during the 7-day experiments in both LD and HD treatments of *A. fundyense* with no significant differences compared with the control diet (Figure 2.5). Similar survival rates were measured in two sets of conditions (100 mL and 200 mL containers) during the June and July experiments (Figure 2.5). In June, in the 200 mL containers (see section 2.6), survival rates on Day 7 were 100% for control and LD and 83% for the HD (1 dead female) (Figure 2.5A). In July 100% of females fed the three treatments were still alive on Day 7 (Figure 2.5C). In the 100 mL flasks (see section 2.6), survival among treatments was similar but lower than in the 200 mL containers. In both June and July, survival rates on Day 7 ranged between 60% to 80% (Figure 2.5B,D). Overall the copepod showed itself to be resistant to the

toxic dinoflagellate with survival rates ranging between 60-100% over a long period (7 days) in all three treatments.

4. DISCUSSION

Diets of *A. fundyense* had little effect on copepod survival and grazing activity, however, there were effects on copepod reproductive success. No difference in survival was found between the control and experimental treatments. Grazing activity was comparable between the control and the experimental treatments, and estimated daily carbon ingested was similar in all three treatments as well. However, overall egg production and hatching success were lower in the two experimental treatments, suggesting that feeding on the dinoflagellate reduced reproductive success, even at the lower dose.

The calanoid copepod *C. finmarchicus* co-occurs with *A. fundyense* blooms that are common in the waters of the Gulf of Maine (Petitpas et al., 2014). *Alexandrium fundyense* starts to appear in very low numbers during the annual spring bloom (Anderson, 1997), which is usually dominated by diatoms of the genera *Thalassiosira, Chaetoceros* and *Skeletonema* (10^3-10^5 cells mL^{-1})(Bigelow et al., 1926; Starr et al., 1999). During the summer, *A. fundyense* cell densities range between 1 to 100 cells mL^{-1} (Anderson, 1997; Anderson et al. 2005); comparable to the LD treatment used in this study (50 cells mL^{-1}). In contrast, the HD treatment, is comparable to severe blooms > 100 cells mL^{-1} which have been reported occasionally in the Bay of Fundy (Anderson et al., 2014; McGillicuddy et al., 2014; Petitpas et al., 2014). During bloom conditions, *C. finmarchicus* ingests *A. fundyense* even at modest bloom densities of 1 cell mL^{-1} (Teegarden et al. 2001); using clarance rates measured for *C. finmarchcius* late copepodites (CV)

(average 6 mL CV^{-1} h^{-1}) during *A. fundyesne* blooms (average 1 cell $mL^{-}1$ and 1.8 μg C $L-^{1}$) from Teegarden et al. (2001), I estimated that ingestion rates during those bloom conditions would be 0.3 μg C $copepodite^{-1}$ d^{-1}.Thus, assuming similar clearance rates, ingestions rates during blooms of 100 cells mL^{-1} would be similar to those in the LD treatment (30-40 μg C fem^{-1} d^{-1}, Figure 2A). During the June experiment and the first five days of the experiment in July estimated toxin ingestion rates averaged 1 and 0.7 ng STX equivalents $copepod^{-1}$ d^{-1} in the LD and and 4 and 2.1 ng STX equivalents $copepod^{-1}$ d^{-1} in the HD treatment, suggesting that toxin ingestion by females in the HD treatment can be expected to occur under field conditions.

Resistance to toxic *Alexandrium* spp. in the form of high survival rates is widespread among copepods, including *C. finmarchicus* (Teegarden and Cembella, 1996; Teegarden, 1999; Colin and Dam, 2002; Liu and Wang, 2002; Hassett, 2003), although most of these studies were limited to 24 hours. The current experiment demonstrates that survival is not affected even when *C. finmarchicus* feeds exclusively on *A. fundyense* for a period of seven days. Similar results have been reported only for the copepod *Acartia clausi* fed on a monoalgal diet of the toxic *Alexandrium lusitanicum* for six consecutive days (Dutz, 1998). This suggests that copepods exposed to blooms of long duration (weeks to months) may show no increase in mortality due to the toxins, and would have an extended time period to accumulate toxins.

However, the 7-day exposure to, *A. fundyense* diets affected the reproductive success of *Calanus finmarchicus.* During the experimental period, egg production (total number of eggs) was significantly lower for adult females feeding on both experimental diets (July) and in the HD treatment (June) compared with the control. Similar to *C. finmarchicus*, Dutz (1998) reported a reduction in the total number of eggs for *A. clausi* fed for six days on a monoalgal diet of the toxic *A. lusitanicum* that resulted in a reduction of 30% of the total number of eggs

produced by females on the control *Rhodomonas baltica*. This reduction in egg production is greater than what was observed in this study for the HD diet.

The effect of the dinoflagellate diet was even more evident when the number of hatched and healthy nauplii is considered as a measure of reproductive success. In July, on average, LD and HD females produced less than half of the number of viable eggs as in the control. In June, the average number of nauplii produced by the LD females was about 65% of that produced in the control, while the effect of the HD diet was even greater than in July, with only 25% viable eggs compared with the control diet. The effect of diet on egg viability in copepods has been studied both in the field and in laboratory experiments (Ianora and Miralto, 2010). Low egg viability has been reported for many different copepods, including *Calanus* spp. feeding on diatoms (Uye and Takamatsu, 1990; Miralto et al., 1995, 1999; Poulet et al., 1995; Starr et al., 1999; Ceballos and Ianora, 2003; Ianora et al., 2003, 2004). In some copepods, no viable eggs are produced after three days on diets of the oxylipin-producing *Skeletonema marinoi* (Ceballos and Ianora, 2003). Egg viability in *C. finmarchicus* appears to be less affected by diatoms (Starr et al., 1999; Gerecht et al., 2013). Reduction of egg viability in *C. finmarchicus* feeding either exclusively on *Thalassiorira nordenskioldii* (10^4 cells mL^{-1}) or on a mixture of diatoms (*Thalassiorira nordenskioldii, Chaetoceros debilis,* and *Navicula* sp.) was observed only after a prolonged exposure with a decline on the order of 20% (Starr et al., 1999). This is in contrast to the current study where hatching success declined to 50% and 30% after only two days even on the LD diet in June and July respectively. As expected, in the HD treatment the decline was even stronger with hatching success of 25% (June) and 22%(July) compared with the control diet. These results suggest that *A. fundyense* blooms could have a significant impact on *C. finmarchicus* fitness even at *A. fundyense* densities that are common in the Gulf of Maine

(Anderson, 1997; Anderson et al. 2005).

Whether the saxitoxins were responsible for the observed effects on *C. finmarchicus* reproductive success remains unclear. The experiments reported here used *Rhodomonas* sp. as the control diet. Although it was confirmed that *A. fundyense* used in this study (strain GTC28) produced saxitoxins during the two 7-day experiments, there is the possibility that the negative effects were not caused by those toxins, but rather by other metabolites produced by *A. fundyens*e. A dramatic reduction in egg production and hatching success in the copepod *Temora stylifera* fed on a non-toxic *Alexandrium tamarense* was attributed to other compounds that might be interfering with fertilization (Ianora et al., 2004). In addition, compounds that inhibit the growth of microalgae and heterotrophic protists have been identified in toxic and non-toxic strains of *Alexandrium* spp. (Tillmann and John, 2002; Tillmann et al., 2008). Thus, in the current experimental design, it is difficult to determine whether the effects on *C. finmarchicus* reproductive success were caused by the saxitoxins and/or by other compounds produced by the dinoflagellate. Alternatively, it has been suggested that toxic dinoflagellates are nutritionally inadequate (e.g. *Karenia brevis*-Prince et al., 2006; Waggett et al., 2012); however, even if the fatty acid profile of the dinoflagellate was not measured in this study, it is known that the fatty acid composition of *Alexandrium* spp. is similar to many other dinoflagellate species that are considered a quality food for copepods (Hammann et al., 2013; Ianora et al., 2004).

In conclusion, while survival rates were high, lower reproductive success was observed at both high (HD) and moderate (LD) levels of *A. fundyense* in the diet of *C. finmarchicus* females. This effect was significant after several days on the experimental diets, and could be missed in experiments of short duration (24-48 hrs). The results of this study suggest that blooms of *A. fundyense* in the Gulf of Maine may not affect *C. finmarchicus* survival, but they are likely to

lower the reproductive success of females maturing in June and July. In the summer, when high abundances of *A. fundyense* are common, particularly in the eastern Gulf of Maine, *C. finmarchicus* populations are dominated by the second generation (G1) (Miller et al., 1998). Starting in June, the population is represented by mid- and late-stage copepodites (CIV and CV) that are either preparing for diapause or maturing directly into adults and producing a third generation (G2) (Miller et al., 1998; Durbin et al., 2000; Fiksen, 2000). It is not clear whether *A. fundyense* might interfere with preparation for diapause in copepodites (stages CIV and CV). However, lower reproductive success by females encountering *A. fundyense* blooms could potentially reduce recruitment to the third generation (G2). This reduction in reproductive success might affect recruitment in the following winter/spring if the G2 contributes significantly to the winter adult population (e.g., Miller et al., 1998; Saumweber and Durbin, 2006).

ACKNOWLEDGMENTS

Co-authors
Jefferson T. Turner, David Kulis, Donald M. Anderson and Petra H. Lenz.

Others
I wish to extend my appreciation to the many colleagues who generously contributed to this study from the initial planning stages to its completion. We would like to thank the Elizabeth T. crew (Lobster boat cruises, MA), Adrianna Ianora, Adriana Zingone and Diana Sarno from Stazione Zoologica A. Dohrn (Naples), Daniel K. Hartline from the University of Hawaii at Manoa as well as Rajdeep Roy (ISRO-NRSC, Balanagar, Hyderabad, India) and Juliette Smith (Virginia Institute of Marine Science) for the toxin analyses. This research was supported by the National Science Foundation Grant OCE-1040597 and OCE-1235549 to Petra Lenz, Cades Fundation of Honolulu grant and Mount Desert Island Biological Laboratory's David W. Towle

Fellowship 2012 to Vittoria Roncalli.Support for D. M. Kulis, D. M. Anderson, and J.T. Turner was provided through the Woods Hole Center for Oceans and Human Health, National Science Foundation Grant OCE-1314642 and National Institute of Environmental Health Sciences Grant 1-P01-ES021923-01.

Table 2.1. Experimental treatments: daily food added to *C. finmarchicus* adult females kept at 1 female per 100 mL during two 7-day experiments

	Treatment		
	Control	Low dose	High dose
	100% *Rhodomonas* spp.	25Alex: 75 Rho	100% *A. fundyense*
Cell mL^{-1}	8000	Rho=6000	200
		Alex=50	
		Total=6050	
μg C L^{-1}			
	304*	Rho=228*	358**
		Alex**=89	
		Total=**317**	

Each treatment had the same total volume. Carbon volume relationship has been calculated using the formula from Mended-Deuer&Lessard (2000).

Carbon content for for *Rhodomonas baltica* (38 pgC $cell^{-1}$) (Fields et al., 2015)

Carbon content estimated for *A. fundysense* (1790 pgC $cell^{-1}$). Cell volume has been calculated using formula shape-2 from (Sun and Liu, 2003).

Table 2.2. Summary of parameters monitored in *C. finmarchicus* during 7-day experiments in June and July 2012							
Conditions	WHOI	200 mL containers		100 mL containers			
# females	--	2 per container		1 per container			
Measurements	*Algal toxicity*	*Survival*	*Grazing rate**	*Survival*	*Fecal pellet production*	*Egg production*	*Egg viability*
June	X	X	-	X	X	X	X
July	X	X	X	X	X	X	X

* Grazing rate was measured as clearance and ingestion rate
WHOI= Woods Hole Oceanographic institution

Table 2.3. Statistic results for *C. finmarchicus* fitness parameters significantly affected by treatment or time of exposure. For each parameter, a two way-ANOVA was performed using P<0.05 as significant cutoff.

	Factor	June				July			
		df	*MS*	*F*	*p*	*df*	*MS*	*F*	*p*
Fecal pellets	Treatment	2	1807	25.16	**< 0.0001**	2	247.6	2.53	0.0818
	Time	6	1940	27.01	**< 0.0001**	6	2136	21.89	**<0.0001**
	TxT	12	187.9	2.61	**0.0031**	12	262.9	2.69	**0.0022**
Egg production daily	Treatment	2	70.48	1.98	0.1409	2	168.7	1.70	0.1847
	Time	6	211.4	5.94	**< 0.0001**	6	3868	39.07	**< 0.0001**
	TxT	12	36.18	1.01	0.4347	12	163.7	1.65	0.0801
Egg viability	Treatment	2	9635	8.92	**0.0003**	2	45141	58.69	**< 0.0001**
	Time	6	4247	3.93	**0.0015**	6	13820	17.97	**< 0.0001**
	TxT	12	2142	1.98	**0.0336**	12	602.8	0.78	0.6664

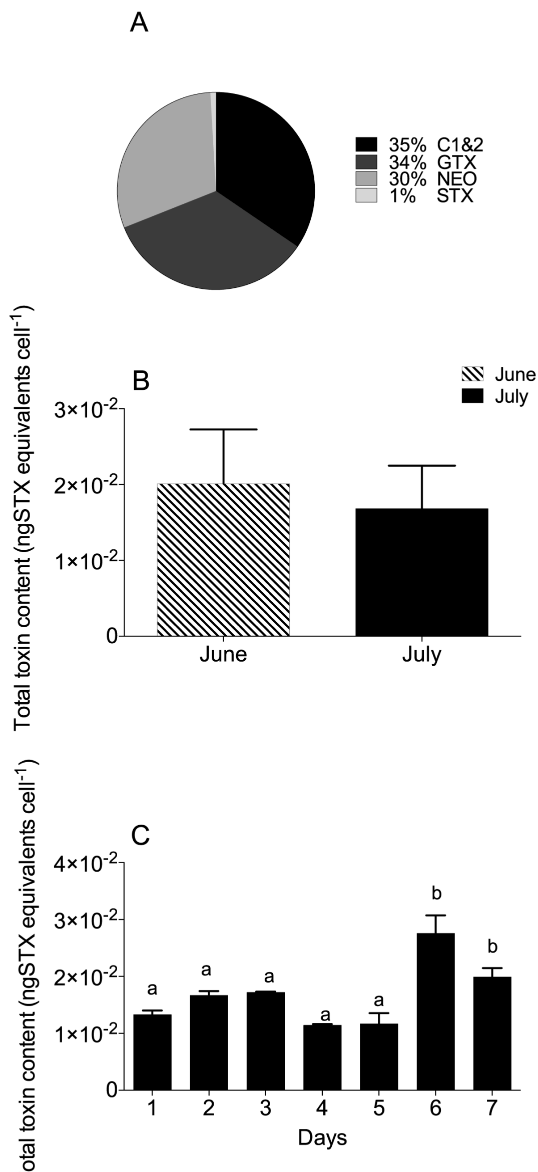


Figure 2.1. Toxin profile of the dinoflagellate *Alexandrium fundyense*.
A) Percent molar PSP toxin composition for the July experiment.
B) Average of total toxin content during 7-day experiments in June and July.
C) Daily total toxin content measured during the July experiment; results of One-way ANOVA are reported on each bar: bars with the same letters are not significantly different.

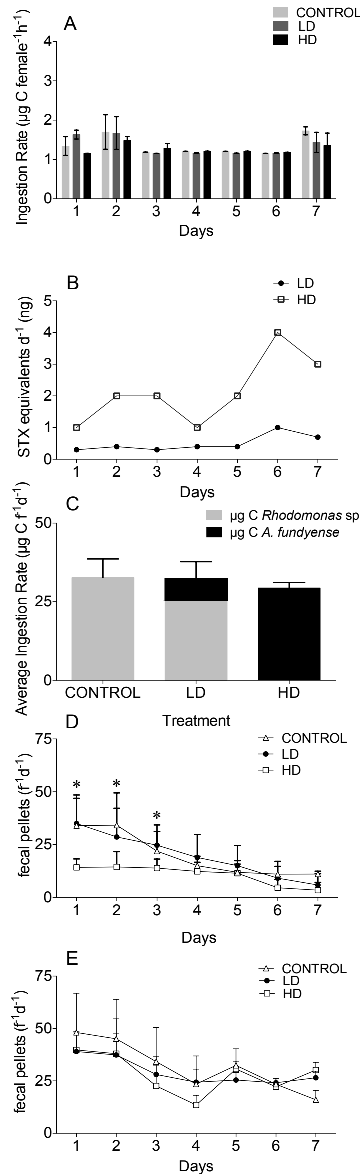


Figure 2.2. Grazing activity for *C. finmarchicus* fed with LD and HD of *A. fundyense* and the control *Rhodomonas* spp. for 7 days. Results of grazing for July experiment with ingestion rates measured as Carbon content (A), ingestion rates measured as STX equivalents (B) and average of daily ingestion rate (C). In D and E, fecal pellet production rate measured for females individually incubated in 100 mL containers (see text for details) in June (D) and July (E) experiments.

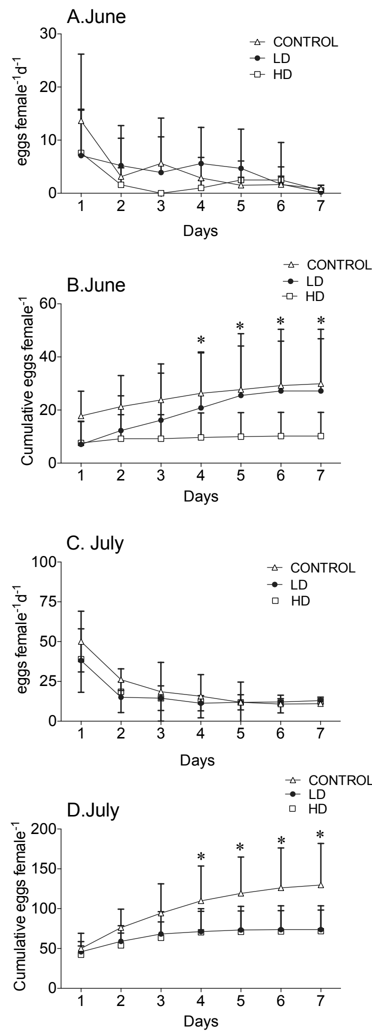


Figure 2.3. ***C. finmarchicus*** **egg production rate after exposure to LD and HD of *A. fundyense* and the control *Rhodomonas* spp. for 7-days.** Daily (A) and cumulative (B) egg production rates over the 7-day experiment in June. In C) daily and cumulative (D) egg production rate over the 7-day experiment in July.

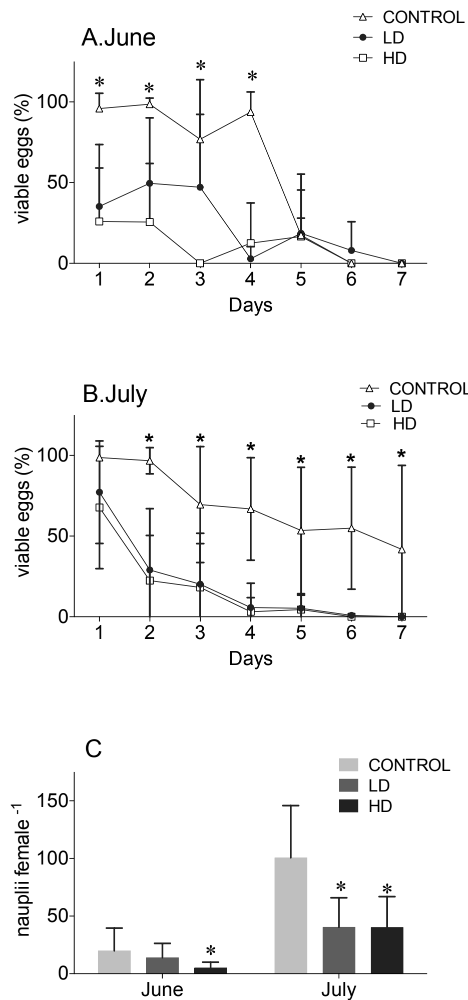


Figure 2.4. Egg viability and effective recruitment (nauplii female^{-1}) of *C. finmarchicus* adult females fed with LD and HD of *A. fundyense* and the control *Rhodomonas* spp. *for 7 days.* Egg viability (%) for June (A) and July (B) experiments and effective recruitment (C) for both experiments. In C, the number of nauplii female^{-1} after 7 days exposure has been calculated for females that were still alive by Day 7. This corresponds to n=8 for C and LD and n=6 for HD in the June experiment and n=7 for C and LD and n=6 for HD in the July experiment.
* indicates significant differences between the treatments observed after statistical test (2-way ANOVA P<0.05).

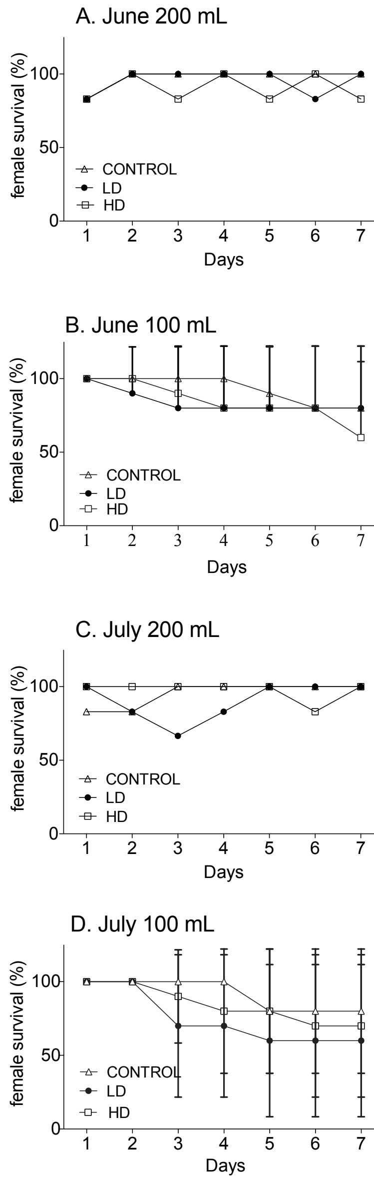


Figure 2.5. *C. finmarchicus* survival rates after exposure to LD and HD of *A. fundyense* and the control *Rhodomonas spp.* for 7 days. In A and B survival rates measured in June for females incubated in 100 mL (A) and 200 mL (B) containers. In C and D, survival rates measured in July for females incubated in 100 mL (C) and 200 mL (D) containers (see text for details). In the 200 mL containers dead females were replaced which is why the graph reset to 100% each day. For the 100 mL containers no replacements were made so mortality is cumulative.

Chapter 3

Global gene expression of *Calanus finmarchicus* in response to the toxic *Alexandrium fundyense*

ABSTRACT

The calanoid *Calanus finmarchicus* co-occurs with harmful blooms of the dinoflagellate *Alexandrium fundyense* in the Gulf of Maine. Although the copepod's survival is not affected by the dinoflagellate even at high doses, sub-lethal effects in the form of reduced reproductive success have been observed at both high and low doses. Here, global gene expression was measured in adult females fed on either control or *A. fundyense* diets using RNA-Seq technology as a proxy for the physiological response to *A. fundyense*. Global changes in gene expression were investigated after a short (2 days) and a long (5 days) period on the experimental diet and compared with females feeding a control diet consisting of the flagellate *Rhodomonas* sp. . A significant time and dose response was observed in *C. finmarchicus* with over 1,000 genes identified as differentially expressed (DE) in response to *A. fundyense*. At 2 days, the response was comparable to the cellular stress response with genes involved in energy metabolism (lipid and carbohydrate), chaperone and redox reactions and protein turnover. At 5 days, the transcriptional response involved fewer genes, presumably reflecting a physiological adjustment towards new environmental conditions. Although the two experimental treatments shared a large number of DE genes, many fewer genes were shared between time points in either treatment. Down-regulation of lipid biosynthesis, reproduction, growth and protein synthesis suggested a difference in energy balance between the control and experimental females. Surprisingly, detoxification was not a major component of the copepod's response to the dinoflagellate while lipid and carbohydrate regulation occurred at both time points. Overall, the results suggest that *A.*

fundyense affects the copepod energy budget, with fewer resources available for growth, which might explain lower egg production and viability.

1. INTRODUCTION

In their natural environments, organisms are affected by multiple abiotic and biotic factors that affect their life history and physiology (Slater, 1997). As environmental conditions move organisms away from their physiological optimum, the "sub-optimal conditions" will impact their development, growth and reproduction even though there may be no increase in mortality (Slater, 1997; Roff, 2002). Organisms exposed to sub-optimal conditions require higher expenditures of metabolic energy for maintenance, and this energy is deflected from processes such as reproduction and growth (Koehn and Bayne, 1989; Slater, 1997), which often leads to a reduction in fitness.

Understanding the mechanisms that regulate trade-offs between life-history traits is key to understanding the ecology of an organism, or even a community. The study of trade-offs in life-history theory has been extensively approached from a "phenotypic" perspective (Stearns, 1989; Roff, 2002, 2007) that might predict the optimal equilibrium but cannot predict how that optimum is reached (Roff, 2007). On the other hand, transcriptomics may help to address how a trade-off that is expressed at the phenotypic level is manifested or modulated at the level of genetic regulation (Van Straallen and Roelofs, 2012). Transcriptomics is the collection of RNA transcripts that are expressed at any given time in an organism or a specific tissue. The quantitative assessment of tens of thousands of genes at a given time, global gene expression pattern, can be determined using high-throughput methods such as RNA-Seq. The application of global gene expression studies for addressing questions about life-history patterns is one recent

technological advance that is increasing our understanding of the molecular basis of physiological adaptation to environmental conditions (Van Straallen and Roelofs, 2012). With the rapid advance in high-throughput sequencing technologies and whole transcriptome profiling, new transcriptomic resources have been developed for non-model species that are ecologically important including marine zooplankton such as the copepod *Calanus finmarchicus* (Lenz et al., 2014; Tarrant et al., 2014). This provides new opportunities to study gene expression responses to different environmental conditions, leading to a better understanding of the physiological ecology of these non-model organisms (Wang et al., 2009; Stillman and Armstrong, 2015; Beaugrand, 2015).

When experiencing environmental stress, organisms activate multiple cellular stress defense mechanisms with the aim to counteract the perturbed state (Van Straallen and Roelofs, 2006). Stress responses vary widely depending on the stressors and the duration of exposure. The magnitude of response (number of genes and relative expression) highly depends on the severity of the stressor (Gasch et al., 2000). In organisms exposed to severe, potentially lethal conditions, differences in expression may reach 100-fold change compared with the control. When exposed to more realistic sub-lethal conditions, differential expression may be subtle and modest (e.g., 2-4 fold difference in expression compared with the control) (Podrabsky and Somero, 2004; Poynton, 2007; 2012). It is known that in eukaryotes a stereotypical defense response is activated after the introduction of a stressor, which is known as the "cellular stress response" (CSR) and consists of a large transcriptional activity with activation of a common set of genes involved in general aspects of cellular protection such as DNA repair and protein stabilization (Kültz, 2003, 2005). Genes involved in chaperone activity (heat shock proteins), redox and detoxification regulation (e.g. thioredoxin, glutathione peroxidase, glutathione reductase), energy metabolism

(e.g. lipid) and intracellular signaling (e.g. serine threonine kinase) are highly up-regulated during this initial response (Kültz, 2003; 2005). In addition, cells that undergo division (DNA replication and cell cycle) are more susceptible to the stress, and activation of cell cycle checkpoints is another aspect of the CSR that includes maintaining cellular integrity and completing cell division under stress. The cellular stress response is followed by the "cellular homeostasis response" (CHR) that consists of regulation of fewer genes, and is typically stressor-specific (Kültz, 2005). This response is sustained while sub-optimal environmental conditions are maintained and its aim is to restore cellular homeostasis (Kültz, 2003, 2005). CHR is not as well understood as the initial CSR.

In pelagic marine ecosystems, extreme conditions are the exception, while subtle changes in the physical, chemical and biological environment that could represent sub-optimal conditions for an organism, are the rule (Beaugrand, 2015). Pelagic ecosystems are nutrionally dilute, and the spectrum of potential food within this environment is characterized by frequent changes in abundance and quality. During their life cycle, copepods are likely to encounter sub-optimal nutritional resources that will require physiological adjustment. Examples include low food abundance, low food quality (e.g. detritus), and toxic food particles. Evidence of sub-optimal food conditions encountered by copepods in the natural environment include small size, low RNA/DNA ratios, reduced egg production and low hatching success (Turner and Tester, 1997; Miralto et al., 1999; Campbell et al., 2001; Ianora et al., 2004; Turner, 2014).

In the Gulf of Maine, *C. finmarchicus* co-occurs with the toxic dinoflagellate *Alexandrium fundyense*. Although ingestion of this alga does not affect survival, the question remains whether the dinoflagellate represents an environmental stressor for the copepod. Thus, the goals here were: 1) to use global gene expression to establish the presence/absence of a

cellular stress response that is characterized by an immediate and transient changes in transcription; and 2) to characterize the cellular homeostatis response in adult females ingesting the toxic dinoflagellate using gene expression as a proxy for physiology. Adult female *C. finmarchicus* were fed on a control diet or one of two experimental diets composed of either 25% or 100% *A. fundyense*. Relative gene expression was determined in whole individuals after 2 and 5 days following the introduction of the experimental diets using high-throughput sequencing. The differential expression of genes between the control and the two experimental treatments on day 2 suggest that a cellular stress response was induced in both low and high doses of *A. fundyense*. Many fewer differentially expressed genes were observed at 5 days, and the response was characterized by an adjustment to the energy balance in the experimental treatments compared with the control. Relatively few genes involved in detoxification were up-regulated at either 2 or 5-days after the introduction of the experimental diet.

2. MATERIALS AND METHODS

2.1. Field collection and maintenance of *Calanus finmarchicus*

Calanus finmarchicus were collected using a vertical net tow (75 cm diameter, 560 µm mesh) in July of 2012 in the Gulf of Maine near Mount Desert Rock (Lat: 44° 2'N; Long: 68°3'W) as described in Chapter 2. Adult females and adult males were transferred into 3.5 L jars at 5-10 individuals per liter with *Rhodomonas* sp. added *ad libitum,* and allowed to acclimate overnight at 10°C on a 14:10 h light:dark cycle in an incubator (Percival Model I-36VL, Percival Scientific, Inc., Perry, IA, USA).

2.2. Experimental design

Experimental incubations for the RNA-Seq experiment were conducted in parallel with the survival, grazing and reproduction experiments described in Chapter 2. For the RNA-Seq experiment, 18 containers were set-up with adult females to be harvested either at 2 days or 5 days. Three biological replicates were set up containing 15 females in 1.5 L containers for each time point and treatment, control, low dose and high dose. In the control group, copepods were fed the non-toxic flagellate *Rhodomonas* sp. (8000 cells $mL^{-1}d^{-1}$). In the "low dose" group (LD) copepods were fed 50 cells $mL^{-1}d^{-1}$ *A. fundyense* and 6000 cells $mL^{-1}d^{-1}$ *Rhodomonas* sp., which corresponded to a 25:75 proportion by algal volume. The "high dose" (HD) group was fed 200 cells $mL^{-1}d^{-1}$. New food suspensions were added daily and copepod survival rates were monitored. The three experimental food suspensions corresponded in terms of carbon (see Table 2.1, Chapter 2). All experimental animals were kept in a Percival Model I-36VL Incubator System (Percival Scientific, Inc., Perry, IA, USA) at 10°C on a 14:10 h light-dark cycle. At 2 and 5 days, a total of 10 adult females were harvested for each biological replicate and treatment, and immediately transferred in 0.5 mL RNAlater (Ambion) and stored at -80°C until RNA extraction.

2.3. RNA-Seq overview

In this study I followed a workflow used in many RNA-Seq studies to determine differential gene expression between *C. finmarchicus* females in the experimental vs control diets at 2 and 5 days. Briefly, as reviewed by Oshlack et al. (2010), the workflow consisted of four steps: 1) generation of millions of short reads of the transcripts (section 2.4); 2) quantitative mapping of reads against a reference genome or transcriptome to quantify relative gene expression (section 2.5); 3) identification of differentially expressed genes (DE) (section 2.6); and 4) biological insights through functional interpretation of DE genes (section 2.7).

2.4. RNA extraction, gene library preparation and sequencing

Total RNA was extracted from whole *C. finmarchicus* adult females using QIAGEN RNeasy Mini Kit (QIAGEN Inc., Valencia, CA, USA), in conjunction with a Qiashredder column (QIAGEN Inc.), following the instructions of the manufacturer, and with a final elution volume of 30 µl. RNA concentration and quality were checked using an Agilent Model 2100 Bioanalyzer (Agilent Technologies, Inc., Santa Clara, CA, USA). A total of 18 samples (3 biological replicates x 3 treatments x 2 time points) were shipped on dry ice to the University of Georgia Genomics Facility for library preparation and Illumina sequencing. There, the TruSeq RNA sample preparation kit (Illumina) was used to prepare double-stranded multiplexed cDNA libraries starting from the total RNA following manufacturer's instructions. Briefly, RNA samples were first purified with two oligo-dT selection (poly (A) enrichment using oligo-dT beds) to select for only mRNAs, and then fragmented and reverse transcribed into double-stranded complementary cDNA. cDNA libraries were prepared with a 350 bp insert and primed using random hexamers. Each sample was tagged with an indexed adapter prior to shipping to University of Missouri DNA Core Facility (*http://biotech.missouri.edu/dnacore*) for sequencing. At the Missouri facility the samples were loaded into a single lane to be run on Illumina HiSeq 2000 instrument using paired-end sequencing (100 bp).

2.5. Mapping of short reads

The 18 RNA-Seq libraries obtained from the sequencing step (section 2.4) were assessed for quality using FASTQC software (v 0.10.0). Removal of low quality sequences was done using the FASTX Toolkit (version 0.013; http://hannonlab.cshl.edu/fastx_toolkit/); this involved

trimming of the first nine and the last 29 bases followed by the elimination of low quality reads (an average "Phred" score lower then 20). This resulted in the removal of an average of 34% of reads leaving from 10 to 24 million reads for relative gene expression analysis in each sample.

Each quality filtered RNA-Seq library was then mapped to a *C. finmarchicus* reference transcriptome (96,090 contigs; Lenz et al., 2014) using Bowtie software (version, 2.0.6) (Langmead et al., 2009) with a 2-nucleotide mismatch tolerance. Information about the generation, quality and coverage of the *de novo* assembled *C. finmarchicus* transcriptome used as a reference can be found in Lenz et al. (2014).

2.6. Identification of differentially expressed genes

Differential gene expression analysis and calculation of fold change difference in expression were performed with the BioConductor package edgeR (Robinson et al., 2010), which uses a negative binominal distribution model and shrinkage estimation for dispersions (Robinson et al., 2010). As implemented by edgeR, the first step is the removal of genes with low counts. As recommended, in this filtering step all genes with fewer than 1 count per million reads (1cpm) were removed since these are unlikely to provide statistical power for detection of differentially expressed genes (Robinson et al., 2010). Because the number of reads per sample differed, the next step involved a between-libraries normalization (TMM-trimmed-mean of M values; Robinson et al., 2010). In addition, variation among biological replicates were calculated, and coefficients of variation were found to be below 30%. This value is in the range of what is expected in gene expression studies (Webster and Santos, 2015) suggesting that the variance in our experimental conditions did not represent an issue for the identification of differentially expressed genes.

Transcriptional expression profiles were analyzed for two factors: duration of exposure (2 and 5 days) and treatment (LD and HD). At each time point, the number of significant differentially expressed genes was established by pairwise comparison of libraries using three comparisons: CONTROL vs LD, CONTROL vs HD and LD vs HD. For the statistical analysis relative expression was quantified as a ratio of Log_2 (experimental/control). Transcripts were recognized as differentially expressed (DE) using the "exact test" ($p<0.05$) and a multiple comparison correction using the Benjamini-Hochberg method (false discovery rate <5%) implemented by edgeR (Robinson et al., 2010). In the results, the magnitude of differential expression is presented as the fold-change difference between the experimental and the control.

2.7. Functional analysis

The annotation process for the differentially expressed genes consist of: 1) identification of genes based on the similarity with known proteins deposited in the National Center for Biotechnology Information (NCBI); 2) assignment of genes into functional categories using Gene Ontology, a major bioinformatics initiative to standardize the vocabulary describing gene products (http://amigo.geneontology.org/amigo); and 3) determining which of those GO categories (biological process, cellular component and molecular function) are significantly represented among the DE dataset compared with the reference transcriptome (enrichment analysis). The first two steps were part of the annotation of the reference transcriptome (Lenz et al., 2014). Genes were identified (step 1) by using *blastx* algorithm to search against the NCBI non-redundant (nr) protein database (maximum E-value 10^{-3}), which was downloaded (February 2013) onto a local Beowulf Linux computer cluster. Gene ontology (GO terms) annotations (step

2) for biological process, molecular function and cellular component were assigned using Blast2GO with a maximum E-value of 10^{-6}.

Annotation of the *C. finmarchicus* reference transcriptome resulted in the identification of 40% of the genes with similarity to known proteins in the NCBI database (Lenz et al., 2014). Using this subset of annotated genes, I was able to identify 46% of the differentially expressed genes. DE genes were categorized into broad functions (stress response, lipid and carbohydrate metabolism, protein turnover, detoxification, cell cycle checkpoint, reproduction and growth) based on assignments found in the literature (e.g., Zinke et al., 2002; Podrabsky and Somero, 2004; Kültz, 2005), GO terms (Gene Ontology) and pathway analysis (Kyoto Encyclopedia of Gene and Genomes [KEGG]). Heat maps, based on magnitude of expression, were generated for each of these categories using *heatmap.2* function implemented by R software.

A second functional analysis involved the identification of GO terms (biological process, molecular function and cellular component) that were overrepresented (=enriched) among the DE genes. Enrichment analysis was performed for the DE genes with GO terms (293) against the 10,344 transcripts with assigned GO terms in the reference transcriptome (Lenz et al., 2014). The analysis was implemented using the software BLAST2GO (version 2.6.4) performing the Fisher's Exact Test followed by Multiple Testing correction of False Discovery rate (FDR <5%) (Conesa et al., 2005). It is important to note that in many cases multiple functions are assigned to individual genes.

Sequence data have been submitted to the National Center of Biotechnology Information (NCBI; www.ncbi.nlm.nih.gov) under the Bioproject PRNJNA236528 containing also the *C. finmarchicus de novo* transcriptome. The list of differentially expressed genes has been submitted to the Biological and Chemical Oceanographic Data Management Office Center of

Biotechnology Information (BCO-DMO; www.bco-dmo.org) under the Project CFINTRANSCRIPT (www.bco-dmo.org/dataset/528312).

3. RESULTS

3.1. Sequencing, mapping and differential gene expression

Illumina sequencing of the 18 libraries (3 replicates x 3 treatments x 2 time points) representing *C. finmarchicus* feeding on the control *Rhodomonas* sp. and *A. fundyense* (LD and HD) treatments for 2 days and 5 days yielded over 484 million paired-end 100 bp reads (Table 3.1). Sequencing yields per library ranged from 15 to 36 million reads with an average of 26 million across all samples (Table 3.1). The number of mapped reads ranged between 10 and 24 millions among the 18 libraries. This number of mapped reads is estimated to be sufficient for quantitative gene expression of 90% of the transcripts (Nagalakshmi et al., 2008; Wang et al., 2009). The Illumina-generated reads were mapped against the *C. finmarchicus* reference transcriptome (96,090 contigs; Lenz et. al., 2014), which yielded an overall alignment rate of 71 to 72% per sample with fewer than 1% of the reads mapped more than once (Table 3.1). A total of 28,756 for the 2-day and 27,943 transcripts for the 5-day RNA-Seq datasets had large enough numbers of mapped reads for statistical testing of differentially expressed genes (DE). The remaining genes in the reference transcriptome were either not expressed in adult females (ca. 1/3 of transcripts; Lenz et al., 2014) or expression was below 1 cpm.

A larger number of differentially expressed genes was identified at 2 days with 1066 in the LD and 1388 in the HD treatments representing 4 to 5% of the transcripts with mapped reads (Figure 3.1). At 5 days 567 DE genes in the LD and 962 in the HD treatments were identified, representing 2 to 3% of the transcripts with mapped reads (Figure 3.1). At 2 days, a substantial

overlap was observed between the two treatments with 65% and 50% of the DE genes shared between the LD and HD datasets, respectively (Figure 3.1A). At 5 days, this overlap was less pronounced with 44% and 31% of the DE shared between the LD and HD datasets, respectively (Figure 3.1B).

In contrast, there was little overlap in DE genes between time points and within the same treatment. Many genes were either differentially expressed at 2 or 5 days with only 8-15% of the genes shared between time points (Figure 3.1C and D). In addition, the number of DE genes was higher in the HD than in the LD treatment and this was particularly apparent at 5 days (Figure 3.1D). These results suggest that the response to the toxic dinoflagellate (LD, HD) is more depended on the time of exposure then the toxic dosage provided in this experiment, although a pronounced graded response was observed.

3.2. Magnitude of response

The magnitude of the response measured as fold-change difference in expression between the experimental and control diets is shown in Figure 3.2 for both up-regulated (Figure 3.2A and B) and down-regulated (Figure 3.2C and D) genes. The largest number of DE genes that were either up- or down-regulated showed only a 2-4 fold change in expression at 2 and 5 days and in both experimental treatments (LD, HD). The number of up-regulated DE genes decreased between 2 days and 5 days (Figure 3.2A and B). In contrast, the distribution of down-regulated genes, whose overall number was lower then the up-regulated ones, was not affected by time (Figure 3.2C and D). At both 2 and 5 days, the number of down-regulated genes appeared to be dose dependent (Figure 3.2C and D).

3.3. Functional annotation

In order to investigate the biological processes that are regulated in *C. finmarchicus* females in response to the toxic dinoflagellate on short (2-days) or long (5-days) term time scales, the DE genes with significant annotations were organized according to their biological functions. DE genes that retrieved annotation (1,162) covered a broad range of conserved eukaryotic processes such as response to stress, intracellular signaling, protein turnover, cell cycle checkpoints, energy metabolism (carbohydrate, lipid), reproduction and growth. The pattern of DE genes support the interpretation that the females responded initially with a "cellular stress response" (CSR). At 5 days the differential gene expression pattern was characterized by metabolic differences between the experimental treatments and the control suggesting a "cellular homeostatis response" (CHR). The list of DE genes used to generate the heat maps in Figures 3.3 and 3.4 and their relative expression are found in Supplemental Table 3.1.

3.3.1. Stress response

Figure 3.3 is a heat map of DE genes encoding stress proteins for LD and HD treatments at 2 and 5 days. It shows a large number of DE genes (118) at 2 days, and many fewer at 5 days. The initial response included mostly up-regulated genes and these included heat shock proteins (HSP40, 70), caspases (apoptosis), cyclins (cell cycle checkpoint), serine kinases (intracellular signaling) and peptidases (protein turnover) (Figure 3.3, Supplemental Table 3.1). Genes involved in detoxification (Rushmore and Kong, 2002) included by 15 DE genes, which represent fewer than 7% of the detoxification genes in the reference transcriptome. These DE genes included four members of the cytochrome P450 family, four glutathione S-transferases

(GST), glutathione peroxidase, sulfotransferase and thioredoxin (phase II) and multi resistance proteins (MXR) (phase III) (Table 3.2).

At 5 days, the response included many fewer DE genes (57) involved in stress and many of these were down-regulated genes (Figure 3.3). The exception to this pattern was the presence of up-regulated genes involved in intracellular signaling, specifically several kinases (Figure 3.3). In addition, down-regulated enzymes included detoxification enzymes (3 cytochrome P450, and GST Omega; Table 3.2), chromosome segregation proteins (cell cycle check point), tyrosine kinases (intracellular signaling) and trypsins (protein digestion) (Figure 3.3).

3.3.2. Energy metabolism

The stress response is energetically costly (Van Straallen and Roelofs, 2006). Thus, one of the characteristics of the cellular stress response is the regulation of genes involved in energy metabolism (KEGG pathway analysis; Gasch et al., 2000; Zinke et al., 2002; Lenz et al., 2014). Here, in *C. finmarchicus* feeding on LD and HD diets, differential expression was observed for genes involved in lipid and carbohydrate metabolism (Figure 3.4A). A total of 33 DE genes were identified in lipid metabolism with the majority involved in either lipid degradation or lipid biosynthesis, with an additional nine DE genes involved in transport (Figure 3.4A). The transcriptional response changed between the two time points. The number of up-regulated genes was most pronounced at 2 days (13 DEs compared with 3 DEs at 5 days), and consistent with increased energy requirements during the cellular stress response: genes involved in lipid degradation were up-regulated, while genes involved in lipid biosynthesis were down-regulated in females on either the LD or HD *A. fundyense* diet (Figure 3.4A). At 2 days a total of 24 DE genes were involved in lipid degradation and these included lipases, phospholipases, oxygenases,

acyltransferases and phosphodiesterases (Figure 3.4A). The number of DE genes in HD females was higher than in LD (10 in LD vs 12 in HD), but the magnitude in terms of fold-change in expression was similar (Figure 3.4A). Genes involved in lipid transport such as lipoprotein and fatty acid binding proteins were also up-regulated (Figure 3.4A). In contrast, genes involved in lipid biosynthesis were not only down-regulated but showed a dose-dependent response with a single DE gene down-regulated in the LD and five in the HD treatment (Figure 3.4A).

After 5 days on the *A. fundyense* diets, *C. finmarchicus* gene expression pattern changed to fewer up-regulated genes, and an increase in the down-regulated ones (Figure 3.4A). The number of up-regulated genes involved in lipid degradation decreased to just 2 and 3 in the LD and HD treatments, respectively (Figure 3.4A). Several genes switched from up- to down-regulated, and there was an additional set of down-regulated DE genes (Figure 3.4A). Down-regulation of lipid biosynthesis was even more pronounced at 5 days than at 2 days with additional DE genes, higher fold-change and a pronounced dose effect in three down-regulated transcripts (ELOV 4, ELOV 7, steroid dehydrogenase; Figure 3.4A).

Forty-three genes involved in carbohydrate metabolism were differentially expressed in *C. finmarchicus* feeding on the dinoflagellate compared with the control diet (Figure 3.4A). The majority of DEs were involved in carbohydrate degradation (26 DEs up-regulated and 8 down-regulated) followed by transport (6 DEs) and biosynthesis (3 DEs) (Figure 3.4A). Up-regulation was pronounced at 2 days, (28 DEs) in particular for genes involved in degradation. Few transcripts (24) were found to be up-regulated in both experimental diets and time points. Interestingly, 15 of these DE genes were involved in carbohydrate degradation (endo β glucanases and glycosyl hydrolases) representing 60% of the shared DE genes across experimental conditions. After 5 days, as for the lipid metabolism, the number of up-regulated

genes decreased with an increase in the down-regulated ones (Figure 3.4A). Down-regulated genes involved in degradation included several chitinases which are enzymes directly involved in digestion. These chitinases showed a dose dependent response with four of the five being differentially expressed only in the HD treatment (Figure 3.4A). Down-regulation was also observed in the carbohydrate biosynthesis with a few synthases only regulated in the HD treatment (Figure 3.4A).

3.3.3. Reproduction and growth

Consistent with a change in energy balance, *A. fundyense* diets affected also the expression of genes involved in reproduction and growth (Figure 3.4B). Transcriptional response was higher at 2 days (20 DEs), with regulation of several chorion peroxidases, vitellogenins and cuticle proteins (Figure 3.4B). In the HD treatment, a pronounced down-regulation was observed for several vitellogenins and cuticle proteins, (Figure 3.4B). At 5 days, there were fewer genes regulated (8 DE) and these genes represented by cuticle proteins, tubulins and myosins (Figure 3.4B).

3.4. Enrichment analysis

Enrichment analysis identified several "enriched" GO terms within the three functional categories: biological process, molecular function and cellular component categories (Tables 3.3 and 3.4). The most striking result of the enrichment analysis was the difference between the two time points: the initial response (2 days) was dominated by processes that were up-regulated while at 5 days the response was characterized by down-regulated processes. Specifically, at 2 days, energy metabolism ("lipid metabolic process") and protein turnover processes were

identified as “enriched” among the up-regulated genes (Table 3.3). Protein turnover involved multiple “enriched” processes such as “RNA biosynthetic process”, “cellular macromolecular biosynthetic process”, “cellular protein metabolic process”, “gene expression”, “translation” and “macromolecule complex” (Table 3.3), many of these are differentially regulated during the “cellular stress response”. The response at 5 days was characterized by down-regulation of major processes including transcriptional regulation (“RNA biosynthetic process”) in the HD treatment, “protein metabolic process” (LD) and “lipid metabolic process” (LD and HD). Some biological processes had changed from up-regulated to down-regulated between the 2- and 5-day time points, in particular in the HD treatment (Table 3.4). Among the down-regulated processes, reproduction was identified as “enriched” although only in the HD treatment (Table 3.4). Absent from the list of “enriched” GO terms were any biological processes associated with the general “response to stimulus” category which includes “detoxification” and “cellular stress response” as “child” terms. One possible explanation is the high stringency of the GO annotation and the enrichment analysis in combination with relatively few DE genes in these specific categories (for example 15 DE genes involved in detoxification).

4. DISCUSSION

The dinoflagellate *A. fundyense* reduces the fitness of the calanoid *C. finmarchicus* as shown by lower egg production and egg viability compared with control females (Chapter 2). In addition, RNA-Seq data showed extensive differences in gene expression profiles measured in females after 2 and 5 days of feeding on the toxic dinoflagellate compared with control females. Functional analysis of the differentially expressed genes was used to start to understand how the dinoflagellate affects the physiology of the copepod. My results show distinct gene expression

profiles between the two time points with few differentially expressed genes shared between 2 and 5 days on the experimental diet.

4.1. Cellular stress response

Stress-defense systems are activated in all eukaryotes with the aim to prevent physiological damage caused by a stressor (Feder and Hoffman, 1999; Van Straallen and Roelofs, 2006; Stillman and Hurt, 2015). The cellular stress response involves large-scale transcriptional activity of a set of stress proteins, which includes HSPs, detoxification enzymes, serine kinases and cell cycle checkpoint proteins (Kültz, 2003, 2005). This response is energetically costly and in order to balance the high energetic costs, simultaneous up-regulation of genes involved in energy-generation (e.g. carbohydrate and lipid degradation) as well as down-regulation of growth-related functions (e.g. RNA metabolism, reproduction and growth) are observed (Kültz, 2003, 2005). Consistent with the cellular stress response, *C. finmarchicus* adult females responded by regulating large number of genes (over 1,000 DE genes in both LD and HD treatments) after 2 days on the *A. fundyense* diets. Up-regulation was observed for the common set of stress response proteins (Van Straallen and Roelofs, 2006) as well as for genes involved in energy metabolism (lipid, carbohydrate). Consistent with this response being energetically costly, I observed down-regulation of genes involved in reproduction and growth. Nevertheless the response was modest with most of the genes showed 2-4 fold change difference in their expression compared with the control. Similar mild responses in differential gene expression have been reported in response to exposure to sub-lethal conditions that commonly occur in the natural environment (Podrabsky and Somero, 2004). Furthermore, in several ecotoxicology studies, when the cladoceran *Daphnia* spp. was exposed to stress conditions that

did not affect survival (e.g. 1/10 LC50), the majority of differentially expressed genes showed a 1- to 4-fold change compared with the control levels (Poynton 2007; 2012). Thus, the modest fold-change for *C. finmarchicus* genes in both LD and HD treatments is not surprising given that the experimental diets did not increase mortality, and that the copepod co-occurs with *A. fundyense* in the Gulf of Maine. However, it is also clear that, by inducing a response that involves over 1,000 differentially expressed genes, *A. fundyense* is affecting the copepod's physiology, and a mild cellular stress response is induced.

A graded response was observed when comparing the number of differentially expressed genes between the two toxic diets. A larger number of genes were differentially regulated in females in the HD than in the LD treatment. This was particularly evident among the down-regulated genes. Interestingly, although there was significant overlap between the two toxic diets (50-65% DE shared), many DE genes in the LD treatment were not differentially expressed in the HD. Similar results were reported for the brown trout *Salmo trutta* when exposed to two doses of an oxidative stressor (glyphosate) (Webster et al., 2015). A fair number of differentially expressed genes were exclusively found in one concentration and not in the other, with fewer then 2% shared between them (Webster et al., 2015). The authors excluded the possibility that this was caused by limited statistical power (e.g. sequencing bias or biological variation).They suggest that although the two treatments differed only for the glyphosate concentration, they may have represented distinct stressors for the organism due to extreme differences in dosage (50-fold; Webster et al., 2015). In my study, I had a 4-fold difference in *A. fundyense* in the diets and in addition, the LD diet included a second alga. Thus, the difference in the treatments was not limited to a single factor. The combination of two algae compared with a unialgal diet might have led to responses that included different sets of DE genes.

4.2. Detoxification and digestive system

Exposure to a broad range of stressors (e.g. temperature, toxic chemicals) is known to trigger an increase in gene expression of enzymes involved in detoxification (phase I, II, III) with the goal to remove the "stress insults" as fast as possible from the organism (Rushmore and Kong, 2002). The importance of detoxification for eukaryotes is exemplified by the high diversification of detoxification enzymes such as cytochrome P450 (CYP P450) and glutathione-S-transferase (GST) families (Chapter 4, Willoughby et al., 2006). In the insect, Willoughby et al. (2006) described detoxification in *Drosophila melanogaster* in response to short-term exposure (4 hr) to natural plant compounds (e.g. caffeine and barbiturate drug phenobarbital). The response involved 10% of the total number of GSTs and cytochrome P450s. All were up-regulated with 2-30 fold change in expression. In contrast, I observed up-regulation of only two GSTs (5%, <3 fold change) while an additional two GSTs (5%) and four CYP450 (4%) were down-regulated. These results suggest that detoxification was not a primary response to the *A. fundyense* diets.

However, a similar transcriptional response, with few detoxification enzymes was reported in the cladoceran *Daphnia* spp., in response to a diet of the cyanobacterium *Microcystis arguinosa* (Asselman et al., 2012). Using gene expression study comparable to our RNA-Seq (microarray with ca. 30,000 genes) the authors reported high transcriptional activity (>2,000 differentially expressed genes) with many genes involved in energy-related processes (e.g. carbohydrate metabolism, protein and ribosome regulation; Asselman et al., 2012). However, they reported only six up-regulated GSTs all with 2-fold difference in expression (Asselman et al., 2012).

Asselman et al., (2012) suggested that the combination of high energetic expenditure and modest detoxification response observed in *Daphnia pulex* feeding on *Microcystis* might be explained by direct inhibition of the digestive system. Asselman et al., (2012) reported the down regulation of many trypsins in *D. pulex*, which could lead to reduction in food assimilation (Asselman et al., 2012). Significant reduction in food assimilation has also been reported in the bivalve *Mytilus chilensis* exposed to a monoalgal diet of dinoflagellate the *Alexandrium catanella* (Fernández-Reiriz et al., 2008). Starting from Day 1 of the experiment, a reduction in the assimilation of proteins, carbohydrate and lipids was observed; this result is correlated to the simultaneous decrease in the enzymatic activity of digestive enzymes such as amylase, cellulase and laminarinase (Fernández-Reiriz et al., 2008).

A. fundyense may be interfering with the copepod digestive system causing low food assimilation. Similar to the *D. pulex* study, I found several amylases (3), cellulases (7) and trypsins (7) differentially expressed at both 2 and 5 days. While some of these trypsins were down-regulated (3/7 DE), particularly at 5 days, other enzymes were up-regulated (all cellulases, 1 amylase and 4 trypsins). Interestingly there was no difference in the fecal pellet production (Chapter 2), which might be expected if food assimilation was affected. Further studies will be needed to test whether of one difference between control and experimental females is food assimilation.

4.3. Cellular homeostasis response

The cellular homeostasis response is an adjustment of the organism's physiology toward a new environmental "reality" by activating a set of stressor-specific adaptations (Kültz, 2005). Here, after 5 days on the *A. fundyense* diets, the physiological response observed in *C.*

finmarchicus was very different compared with the one at 2 days. Fewer than 13% of the differentially expressed genes were shared between time points and within each treatment. The cellular homeostasis response in *C. finmarchicus* was characterized by fewer differentially expressed genes and many of these were involved in energy metabolism. Differentially expressed genes involved in detoxification and cell cycle checkpoint were exclusively down-regulated. Genes involved in lipid and carbohydrate biosynthesis as well as reproduction in particular in the HD treatment were down-regulated suggesting that at 5 days on the dinoflagellate diets, *C. finmarchicus* adult females have less energy available than the control ones. As hypothesized above, one explanation could be that feeding on *A. fundyense* leads to reduced assimilation as suggested by the down-regulation of some trypsins and amylases at both 2 and 5 days. The significance of this down-regulation is suggested by the fact that there were only 5 shared down-regulated genes (LD, HD, 2 and 5 days) and 1 trypsin and 1 cellulase were among them. Difference in food assimilation could explain why in spite comparable ingestion rates, the energy available for growth and reproduction is reduced in females feeding on the toxic diets (Chapter 2). Based on those energy budget considerations, I suggest that *C. finmarchicus* feeds on *A. fundyense* with comparable ingestion rates (μg C fem^{-1} d^{-1}), but food assimilation is reduced. Assuming high metabolic costs to balance this reduction, this could explain the lower investment in egg production observed in Chapter 2 is than explained.

5. CONCLUSIONS

In conclusion, the study provides evidence that the dinoflagellate *A. fundyense* is an environmental stressor for *C. finmarchicus*. Upon ingestion of the dinoflagellate a cellular stress response is activated at 2 days that leads to a very different response at 5 days as suggested bu

the few DE genes. Detoxification did not appear to represent a major component of the physiological response. Differential gene expression profiles suggest differences in the energy balance, which might be explained by lower food assimilation. Less energy available to females on the experimental diets probably contributed to observed lower egg production and egg quality (Chapter 2). Thus, *A. fundyense* seems to act as an energy stressor rather than a toxic stressor that would include a global detoxification response.

Table 3.1. Summary of RNA-sequencing and mapping results for *Calanus finmarchicus* adult females feeding on *Rhodomonas* sp. (C) and *A. fundyense* (LD, HD) diets for 2 and 5 days. Raw reads were quality filtered and trimmed prior to mapping to a *C. finmarchicus* reference transcriptome (Lenz et al., 2014) using Bowtie software. C=control diet, LD= Low dose diet, HD= High dose diet. For each treatment 3 biological replicates were collected (r1, r2, r3).

Samples	Raw reads (#)	Discarded reads (%)	Reads for mapping (#)	Overall alignment rate (#)	Total mapped reads (#)	Mapped reads 1 time (#)	Mapped reads 1 time (%)	Mapped reads > 1time (#)	Mapped reads > 1 time (%)
2 days									
C r1	22,678,446	33	15,256,487	71.4	10,898,526	10,799,129	70.8	99,397	0.6
C r2	15,334,636	34	10,098,479	71.1	7,176,817	7,110,892	70.4	65,925	0.6
C r3	20,164,880	33	13,447,296	70.7	9,511,875	9,419,018	70.7	92,857	0.7
LD r1	24,397,138	35	15,852,474	71.6	11,350,347	11,246,340	70.9	104,007	0.7
LD r2	32,032,448	35	20,922,762	70.9	14,836,541	14,673,615	70.1	162,926	0.8
LD r3	26,732,596	34	17,615,813	71.7	12,635,344	12,513,350	71	121,994	0.7
HD r1	25,371,346	35	16,564,095	71.5	11,840,963	11,735,536	70.8	105,427	0.6
HD r2	36,624,954	35	23,924,809	72.5	17,335,549	17,178,945	71.8	156,604	0.6
HD r3	22,361,678	36	14,346,106	71.6	10,274,526	10,171,571	70.9	102,955	0.7
5 days									
C r1	24,950,224	36	15,972,067	71.4	11,407,748	11,307,484	70.8	100,264	0.6
C r2	33,605,768	35	21,727,284	71.5	15,536,963	15,404,299	70.9	132,664	0.6
C r3	36,757,738	34	24,354,372	71.6	17,447,497	17,294,204	71	153,293	0.6
LD r1	35,824,224	35	23,325,237	71.1	16,583,915	16,443,476	70.5	140,439	0.6
LD r2	22,073,822	34	14,518,531	71.7	10,411,305	10,311,783	71	99,522	0.7
LD r3	30,768,506	35	20,111,901	71.3	14,348,892	14,229,233	70.7	119,659	0.6
HD r1	26,147,888	34	17,172,169	71.4	12,257,785	12,143,290	70.7	114,495	0.7
HD r2	22,435,778	32	15,283,042	71.8	10,952,138	10,851,978	71	100,160	0.7
HD r3	26,186,738	35	17,037,038	71	12,095,361	11,985,806	70.3	109,555	0.6

* % of reads mapping

Table 3.2. List of detoxification enzymes differentially expressed in *C. finmarchicus* feeding on *A. fundyense* diets. 15 putative transcripts involved in detoxification (phases I, II, III) that were differentially expressed in females feeding on *A. fundyense* (LD, HD) at 2 and 5 days compared with adult females feeding on the control diet. For each gene, Accession No. (NCBI), *blastx* annotation and relative fold change in expression (absolute) are listed. The direction of expression (up- down regulated) and the magnitude are indicated by arrows red=up- and green=down-regulated genes). Transcripts are listed in the same ordered as in cluster in Figure 3.3.

Gene-Accession No.	Annotation	2 days		5 days	
		LD	HD	LD	HD
Cytocrome P450					
GAXK01103733	Cytochrome family 11				↓↓
GAXK01141739	Cytochrome p450 2J6-like	↓↓			↓↓
GAXK01079613	Cytochrome p450 20A1-like				↓
GAXK01154455	Cytochrome p450 2J2-like				↓
Glutathione S-transferase					
GAXK01204953	GST Delta-I		↓↓↓		
GAXK01204961	GST Sigma-VI	↑↑↑	↑↑↑		
GAXK01204946	GST Sigma-VII	↑↑			
GAXK01204960	GST Omega-I			↑	
Sulfotransferase					
GAXK01109776	Chondroitin 4-sulfotransferase	↑			
GAXK01135132	Sulfotransferase 1c4	↓			
Thioredoxin					
GAXK01120895	Beta and gamma crystallin	↓↓	↓↓		
GAXK01188041	Nucleoredoxin-like protein 2	↑↑	↑		↑
Glutathione peroxidase					
GAXK01178489	Phospholipid-hydroperoxide glutathione peroxidase			↓	
Multi resistance protein (MXR)					
GAXK01143274	Multidrug resistance protein (abc cassette)		↓		
GAXK01081107	Multidrug resistance protein (abc cassette)	↓↓	↓↓		

Up regulated	Down regulated
↑ < 2.5 fold change	↓ <2.5 fold change
↑↑ 2.5-3.9 fold change	↓↓ 2.5-3.9 fold change
↑↑↑ >4 fold change	↓↓↓ >4 fold change

Table 3.3. GO enrichment analysis of up-and down-regulated genes in *C. finmarchicus* feeding on low (LD) and high dose (HD) treatments of *A. fundyense* for 2 days.

Time/treatment	Term description	GO	Category	FDR
2 days LD				
	Up regulated			
	RNA biosynthetic process	GO:0032774	BP	5.16E-03
	Cellular macromolecule biosynthetic process	GO:0034645	BP	3.01E-03
	Cellular protein metabolic process	GO:0044267	BP	4.73E-02
	Gene expression	GO:0010467	BP	3.01E-03
	Macromolecular complex	GO:0032991	CC	3.01E-02
	Heterocyclic compound binding	GO:1901363	MF	3.01E-02
	Organic cyclic compound binding	GO:0097159	MF	3.01E-02
	Peptidase activity	GO:0008233	MF	3.93E-02
	Transferase activity	GO:0016740	MF	1.10E-02
	Cytoplasmic part	GO:0044444	CC	2.46E-02
	Intracellular membrane-bounded organelle	GO:0043231	CC	3.42E-03
2 days HD				
	Up regulated			
	Lipid metabolic process	GO:0006629	BP	4.62E-02
	RNA biosynthetic process	GO:0032774	BP	1.58E-03
	Translation	GO:0006412	BP	4.62E-02
	Macromolecular complex	GO:0032991	CC	2.25E-02
	Nucleic acid binding	GO:0003676	MF	1.91E-03
	Nucleotide binding	GO:0000166	MF	3.79E-02
	Signal transduction	GO:0007165	BP	1.55E-02
	Cytosol	GO:0005829	CC	2.91E-02
	Extracellular region	GO:0005576	CC	5.75E-03
	Nucleus	GO:0005634	CC	1.23E-03
	Down regulated			
	Cellular metabolic process	GO:0044237	BP	2.48E-02
	Single-organism cellular process	GO:0044763	BP	2.88E-02

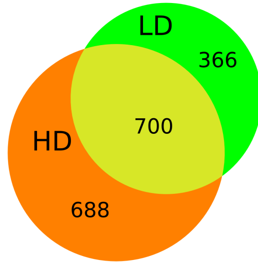
Gene ontology term (GO); Category: Biological process (BP), Molecular function (MF) and Cellular component (CC); false discovery rate (FDR) after Fisher exact test using FDR<5% as cutoff.

Table 3.4. GO enrichment analysis of down-regulated genes in *C. finmarchicus* feeding on low (LD) and high (HD) dose treatments of *A. fundyense* for 5 days.

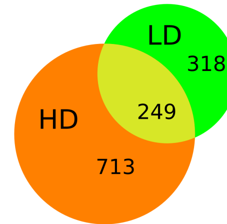
Exposure	Term description	GO term	Category	FDR
5 days LD				
Down				
	Lipid metabolic process	GO:0006629	BP	9.72E-03
	Protein metabolic process	GO:0019538	BP	4.52E-02
	Response to stimulus	GO:0050896	BP	4.52E-02
	Cell communication	GO:0007154	BP	4.52E-02
	Single organism signaling	GO:0044700	BP	3.92E-02
	Nucleobase-containing compound metabolic process	GO:0006139	BP	3.92E-02
5 days HD				
Down				
	Lipid metabolic process	GO:0006629	BP	4.62E-02
	RNA biosynthetic process	GO:0032774	BP	1.58E-03
	Translation	GO:0006412	BP	4.62E-02
	Macromolecular complex	GO:0032991	CC	2.25E-02
	Nucleic acid binding	GO:0003676	MF	1.91E-03
	Nucleotide binding	GO:0000166	MF	3.79E-02
	Signal transduction	GO:0007165	BP	1.55E-02
	Reproduction	GO:0000003	BP	3.84E-02
	Cytosol	GO:0005829	CC	2.91E-02
	Extracellular region	GO:0005576	CC	5.75E-03
	Nucleus	GO:0005634	CC	1.23E-03

Gene ontology term (GO); Category: Biological process (BP), Molecular function (MF) and Cellular component (CC); false discovery rate (FDR) after Fisher exact test using FDR<5% as cutoff.

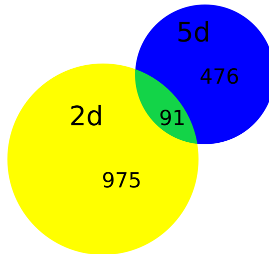
A. 2 days



B. 5 days



C. Low Dose (LD)



D. High Dose (HD)

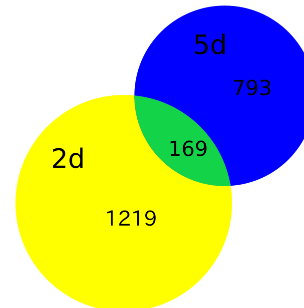


Figure 3.1. Differentially expressed genes in *C. finmarchicus* adult females feeding on *A. fundyense* diets (LD, HD) for 2 days and 5 days. Top panel shows Venn diagram for genes differentially expressed in *C. finmarchicus* adult females feeding on LD and HD diets of *A. fundyense* at 2 days (A) and 5 days (B). Bottom panel shows Venn diagrams for DE genes in *C. finmarchicus* feeding for 2 and 5 days on LD (C) and HD (D) treatments for 2 and 5 days. Here the Venn diagrams show the degree of similarity between the two times of exposure (2 and 5 days). Genes identified as differently expressed using the "exact" test ($p<0.05$) and a multiple comparison correction using Benjamini-Hochberg method (false discovery rate <5%) as implemented by edgeR.

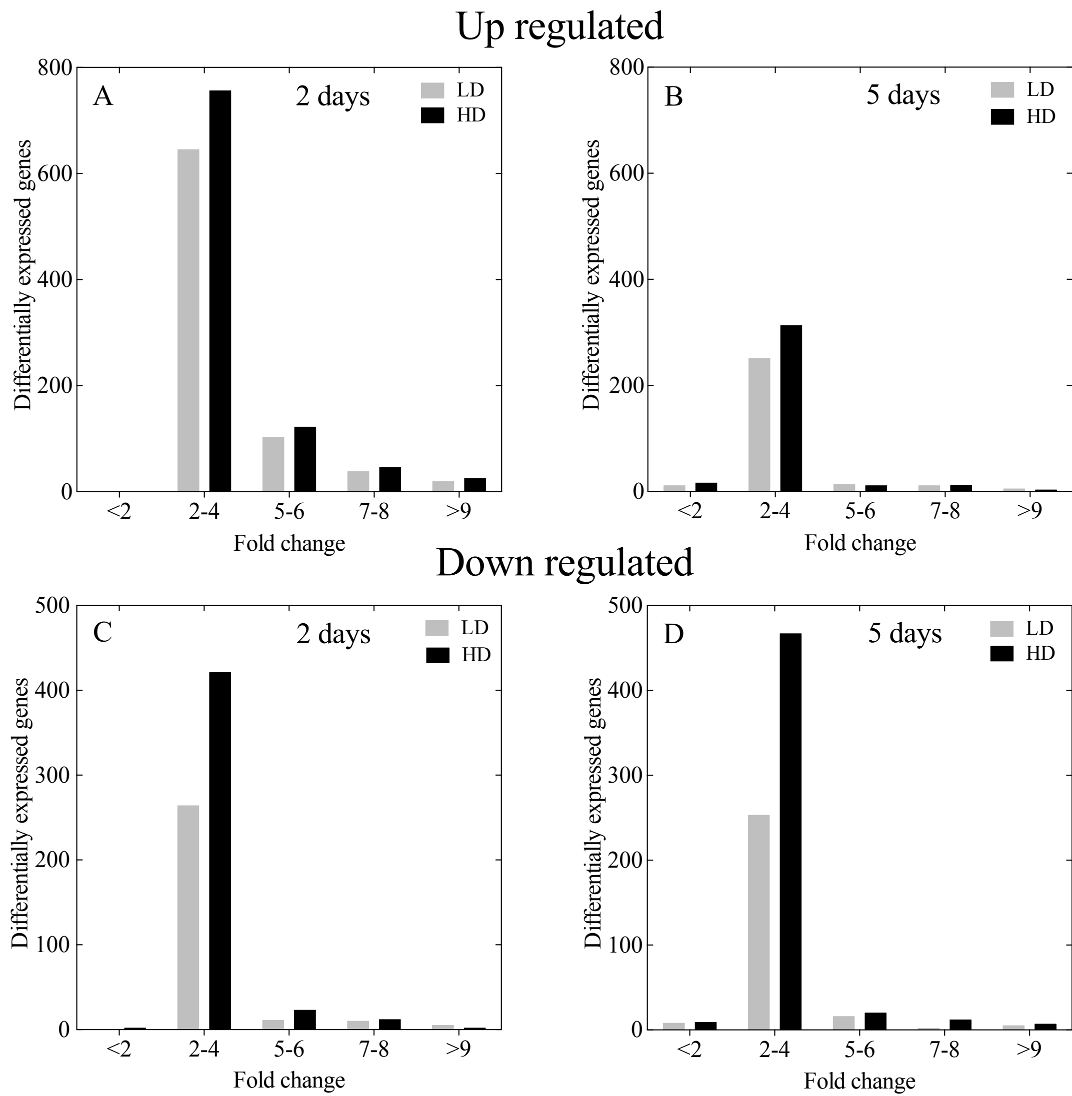


Figure 3.2. Relative expression of differentially expressed genes in *C. finmarchicus* adult females feeding on *A. fundyense* (LD, HD) for 2 and 5 days. Top panel fold change expression (absolute) of up-regulated genes at 2 (A) and 5 days (B). In the lower panel fold change expression (absolute) for the down-regulated genes (note difference in scale) at 2 (C) and 5 days (C).

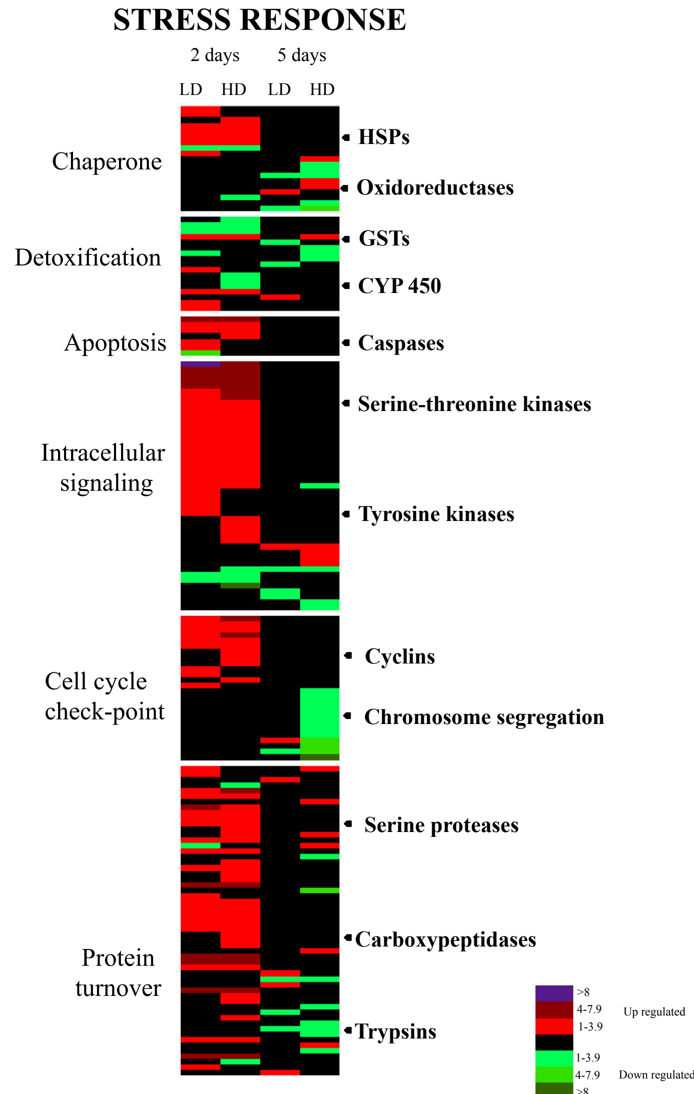


Figure 3.3. *C. finmarchicus* stress proteins. Heat map for stress genes that were differentially expressed in adult females feeding on LD and HD diets of *A. fundyense* at 2 and 5 days. Genes were clustered by functional annotation within the groups (listed on the left): chaperone, detoxification, apoptosis, intracellular signaling, cell cycle checkpoint and protein turnover. Within each cluster, groups of specific genes have been identified (listed on the right). Relative expression rate (absolute fold change) is calculated for females feeding on the toxic diet compared with adult females feeding on the control diet *Rhodomonas* sp. 1st column is 2 days LD; 2nd column 2 days HD; 3rd column 5 days LD; 4th column 5 days HD.

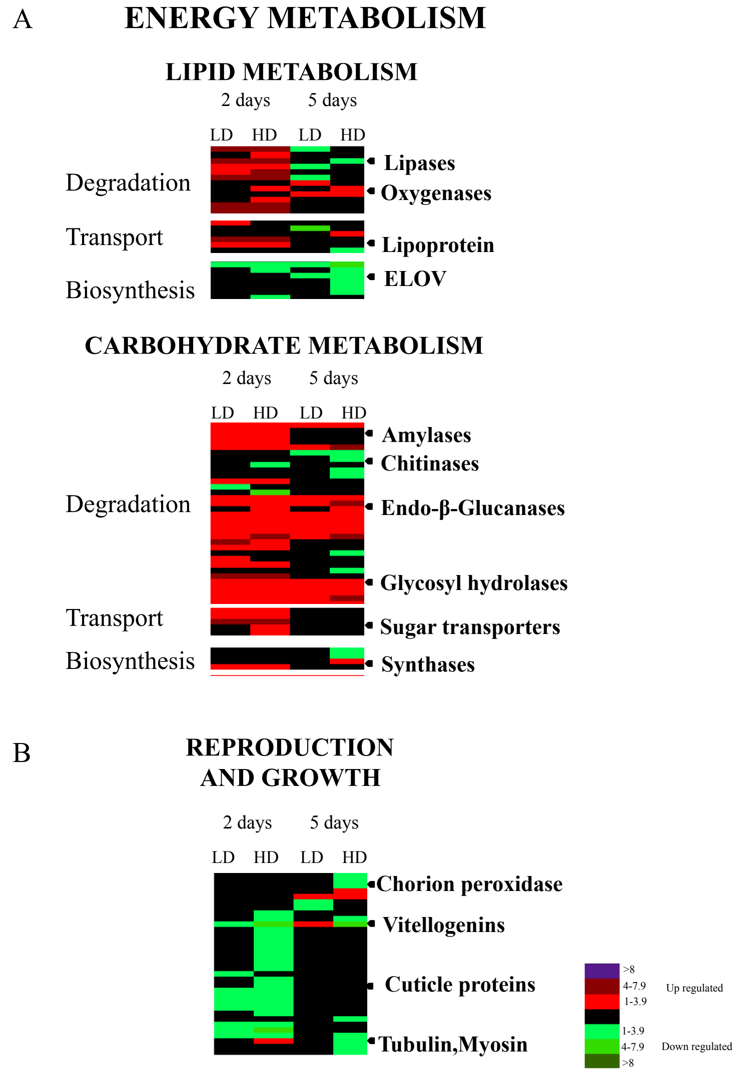


Figure 3.4. ***C. finmarchicus*** **energy and growth related genes.** Heat map genes involved in energy metabolism (A) and reproduction and growth (B) that were differentially expressed genes in *C. finmarchicus* adult females feeding on *A. fundyense* diets (LD, HD) for 2 and 5 days. Genes were clustered by functional annotation within lipid metabolism, carbohydrate metabolism and reproduction and growth. Within lipid and carbohydrate metabolism, genes have been clustered in degradation, transport and biosynthesis (listed on the left). Within each cluster, groups of specific genes have been identified (listed on the right). Relative expression rate (absolute fold change) is calculated for females feeding on the toxic diet compared to adult females feeding on the control diet *Rhodomonas* sp. 1st column is 2 days LD; 2nd column 2 days HD; 3rd column 5 days LD; 4th column 5 days HD

Chapter 4

Glutathione S-transferase (GST) gene diversity in the crustacean *Calanus finmarchicus* – contributors to cellular detoxification

ABSTRACT

Detoxification is a fundamental cellular stress defense mechanism, which allows an organism to survive or even thrive in the presence of environmental toxins and/or pollutants. The glutathione S-transferase (GST) superfamily is a set of enzymes involved in the detoxification process. This highly diverse protein superfamily is characterized by multiple gene duplications, with over 40 GST genes reported in some insects. However, less is known about the GST superfamily in marine organisms, including crustaceans. The availability of two *de novo* transcriptomes for the copepod, *Calanus finmarchicus*, provided an opportunity for an in depth study of the GST superfamily in a marine crustacean. The transcriptomes were searched for putative GST-encoding transcripts using known GST proteins from three arthropods as queries. The identified transcripts were then translated into proteins, analyzed for structural domains, and annotated using reciprocal BLAST analysis. Mining the two transcriptomes yielded a total of 41 predicted GST proteins belonging to the cytosolic, mitochondrial or microsomal classes. Phylogenetic analysis of the cytosolic GSTs validated their annotation into six different subclasses. The predicted proteins are likely to represent the products of distinct genes, suggesting that the diversity of GSTs in *C. finmarchicus* exceeds or rivals that described for insects. Analysis of relative gene expression in different developmental stages indicated low levels of GST expression in embryos, and relatively high expression in late copepodites and adult females for several cytosolic GSTs. A diverse diet and complex life history are factors that

might be driving the multiplicity of GSTs in *C. finmarchicus*, as this copepod is commonly exposed to a variety of natural toxins. Hence, diversity in detoxification pathway proteins may well be key to their survival.

1. INTRODUCTION

The activation of multiple cellular stress defense mechanisms, including an increase in the activity of detoxification enzymes, is key to an organism's ability to survive, and sometimes even thrive, in environments characterized by the presence of toxins and/or pollutants (Kültz, 2005).. In eukaryotes, the cellular detoxification process can be divided into three phases (Xu et al., 2005). In Phase I, reactive/polar groups are enzymatically added to a xenobiotic. In the second phase (Phase II), the modified toxicant is enzymatically conjugated to a polar molecule. In the final phase of the detoxification process (Phase III), efflux transporters that specifically recognize conjugated toxins remove the modified xenobiotic from the cell.

Among the key enzymes for Phase II of the detoxification process are members of the glutathione S-transferase (GST) superfamily (Frova et al., 2006). GSTs are typically small proteins (200-250 amino acids) that are activated in response to oxidative damage and/or exposure to a large variety of toxins (Sheehan et al., 2002). GSTs catalyze the conjugation of reduced glutathione (GSH) to hydrophobic xenobiotics, such as naturally occurring toxins and anthropogenically derived pharmaceuticals and pesticides (Ranson and Hemingway, 2005). The coupling of the xenobiotic to GSH increases the solubility of the toxin, thus facilitating its excretion (Ranson and Hemingway, 2005).

The GSTs are a highly diverse protein superfamily, but can be divided into three distinct classes based on their cellular location, *i.e.*, cytosolic, mitochondrial and microsomal (Sheehan et

al., 2002). The cytosolic class, which is primarily involved in cellular detoxification (Che-Mendoza et al., 2006) contains seven subclasses (Delta, Epsilon, Omega, Sigma, Theta, Mu and Zeta). Six subclasses are found in the insects, which lack members in the subclass Mu (Che-Mendoza et al., 2006). The cytosolic GSTs are all dimeric proteins (homo- or heterodimers) with both subunits originating from the same GST subclass (Habig et al., 1974). Each monomer contains an amino (N)-terminal α/β-domain and a carboxyl (C)-terminal α-helical domain (Che-Mendoza et al., 2006). In all subclasses, the active site, located between the two domains, is composed of two binding sites: the highly conserved G site, which binds reduced GSH, and the highly variable H site (Che-Mendoza et al., 2006).The variability in the H-site allows GSTs to detoxify a variety of "hydrophobic" substrates (Board and Menon, 2013). The catalytic activity of a mature GST is maintained by its dimeric structure, and there is no evidence of any active monomers, which is probably due to structural differences in the G-site between the monomer and the dimer (Abdalla et al., 2002; Fabrini et al., 2009). Members of the Delta and Epsilon subclasses have been implicated in resistance to pesticides, *e.g.*, organophosphates, organochlorines and pyrethroids (Enayati et al., 2005), while the Omega, Theta and Zeta sub-groups appear to be involved in other cellular processes, including protection against oxidative stress (McLellan et al., 1999).

The mitochondrial GSTs, also referred to as Kappa GSTs, are homodimers with a single conserved thioredoxin domain (Frova, 2006;, Board et al., 2013). This functional motif is similar to the N-terminal domain of the cytosolic GSTs, suggesting that these proteins may have similar substrate specificity (Board and Menon, 2013). Kappa GSTs are widely distributed in nature but are absent in insects (Morel and Aninat, 2011). In crustaceans, Kappa GSTs have been predicted from either genomic or transcriptomic sequence data in the daphnid *Daphnia*

pulex (Colbourne et al., 2011) and the copepods *Tigriopus japonicus* (Lee et al., 2008), *Paracyclopina nana* (Lee et al., 2012), *Lepeophtheirus salmonis* (Accession No. ACO11809), *Caligus clemensi* (Accession No. ACO15728) and *Caligus rogercresseyi* (Accession No. ACO10845).

The microsomal GSTs are membrane-associated proteins, primarily localized to the mitochondrion and endoplasmic reticulum (ER), and are involved in eicosanoid and glutathione metabolism (Frova, 2006; Bressell et al., 2005; Jakobsson et al., 2005). This class of GSTs has a single conserved domain, the membrane-associated protein in eicosanoid and glutathione metabolism (MAPEG) domain, which shares high amino acid similarity with the active sites of 5-lipoxygenase-activating protein and leukotriene-C4 synthase, suggesting that they are more distantly related to the cytosolic and mitochondrial GSTs, and may have multiple enzymatic roles that are not exclusively associated with the detoxification response (Bressell et al., 2005; Jakobsson et al., 2005)

Increases in the frequency and magnitude of toxic algal blooms and anthropogenic pollution of marine environments can have devastating impacts on the economies of coastal communities due to the resulting degradation of ecosystems, declines in marine fisheries, and negative impacts on tourism and recreational activities (Islam SM and Tanaka, 2004; Anderson et al., 2008). Although mitigation of the effects of xenobiotics is a high priority, effective management requires an understanding of how toxins and pollutants are transferred through the food chain (Islam SM and Tanaka, 2004). Planktonic copepods are known to play a crucial role in secondary production, potentially serving as vectors in the transfer of toxins to higher trophic levels in marine food webs (Verity and Smetacek, 1996). Alternatively, through biological processes such as detoxification, excretion and fecal pellet production, copepods may be

involved in the removal of xenobiotics from ecosystems (Teegarden et al., 2003). Recently, several investigations have focused on how copepods respond to toxins (Lauritano et al., 2012). In the calanoid copepods *Calanus finmarchicus* and *Calanus helgolandicus*, GSTs have been used as biomarkers of the detoxification response to both natural toxins (phytoplankton toxins) and anthropogenic pollutants (Lauritano et al., 2013; Hansen et al., 2011). Because of limited genomic resources, these studies have depended on single GSTs as biomarkers (Lauritano et al., 2011, 2013). However, given the multiplicity and high diversification of the GST superfamily, these studies may not fully represent the copepods' physiological response to a xenobiotic. Thus, to understand the role of the GSTs in detoxification in marine crustaceans, this protein superfamily must be better characterized. Genomic data from insects, including *Drosophila melanogaster* and *Anopheles gambiae* (Enayati et al., 2005), suggest the presence of 30 or more genes in the GST superfamily. Using insect proteins as queries, just 12 GSTs were identified in the transcriptome of the intertidal copepod *T. japonicus* [15, Roncalli, unpublished]. The identification of only a small number of GSTs in *T. japonicus* raises the question as to whether copepods may exhibit lower GST diversity than insects.

C. finmarchicus, one of the most abundant mesozooplankton species in the North Atlantic Ocean (Dale et al., 2001; Head et al., 2000), is consumed by many economically-important fishes such as cod, mackerel and herring (Beaugrand et al., 2003; Heath and Lough, 2007). Thus, *C. finmarchicus* has been the focus of many ecological studies in the Gulf of Maine, which is well known for frequent blooms of the toxic dinoflagellate, *Alexandrium fundyense* (Teegarden et al., 2003). Recently, a *de novo* reference transcriptome was assembled for *C. finmarchicus* from the Gulf of Maine that included transcripts for six developmental stages (Lenz et al., 2014). It has been estimated that this transcriptome, which was assembled from over 400 million reads

(paired end, 100 bp), includes at least 65% of the complete set of *C. finmarchicus* transcripts (Lenz et al., 2014). This estimate was confirmed by other studies that used the transcriptome to characterize neural signaling molecules in this crustacean (Christie et al., 2013a, 2013b, 2014a, 2014b). Here, this transcriptome was mined for putative GST-encoding transcripts. These data were compared to a second *de novo* transcriptome, generated independently from individuals from a single stage (pre-adult) and originating from the Norwegian Sea (Tarrant et al., 2014). Using known GST protein sequences from insects and other crustaceans as input queries, multiple putative GSTs belonging to the cytosolic, mitochondrial and microsomal classes were identified and characterized from this species. Comparison of the deduced *C. finmarchicus* GSTs with those from the insect *D. melanogaster* and the crustaceans *D. pulex* and *T. japonicus* established that *C. finmarchicus* GST complexity is comparable to those of insects, with the individual proteins showing similarities to both those of insects and of crustaceans. In addition, the relative expression of the putative GST-encoding transcripts was assessed across development. While the relative expression of members of the microsomal and mitochondrial classes was similar in naupliar and copepodite stages, those belonging to several cytosolic subclasses showed low expression in embryos, intermediate expression in early life stages (naupliar and early copepodite stages), and high expression in the pre-adult (late copepodite, CV) and adult stages. Gene diversity was highest for the cytosolic GSTs, specifically in the Delta and Sigma subclasses. These findings are consistent with this gene superfamily playing a critical role in the copepods' physiological response to environmental stressors, and they lay the foundation for future studies on the function of GSTs in *C. finmarchicus* and other copepods.

2. MATERIALS AND METHODS

2.1. *Calanus finmarchicus* transcriptome

Initial searches for *C. finmarchicus* GST-encoding transcripts were performed on the *de novo* assembled transcriptome obtained from animals from the Gulf of Maine; a detailed description of the generation, quality and coverage of this transcriptome can be found in Lenz et al. (Lenz et al., 2014). Briefly, multiplexed gene libraries were generated from RNA collected from six developmental stages: embryo, early nauplius (NI-NII), late nauplius (NV-NVI), early copepodite (CI-CII), late copepodite (CV) and adult female (CVI). Library sequencing was performed using the Illumina HiSeq 2000 platform, generating 415 million, paired-end raw reads (100 base pair long) from the combined samples. These reads were *de novo* assembled using Trinity software generating a total of 206,041 unique transcripts (contigs). The assembled transcripts were submitted to the National Center of Biotechnology Information (NCBI; www.ncbi.nlm.nih.gov) and can be accessed via Bioproject PRJNA236528 (Lenz et al., 2014).

2.2. *In silico* transcriptome mining

Searches of the *C. finmarchicus de novo* assembly for putative GST-encoding transcripts were conducted using the tera-tblastn algorithm of DeCypher Tera-BLASTP on a TimeLogic DeCypher server; detailed descriptions of the search method are provided in Christie et al. (2013,2014) and Lenz et al. (2014). Known GST proteins, the majority from the copepod *T. japonicus* (Lee et al., 2008), were used as the query sequences for all tera-tblastn searches. GST proteins from the insect *D. melanogaster* and the daphnid *D. pulex* were used as queries to search for the cytosolic Epsilon GST subclass (insect specific (Ranson and Hemingway, 2005)) and the mitochondrial Kappa class, respectively. Lastly, the nucleotide sequences of five *C. finmarchicus*

expressed sequence tags (ESTs) previously identified as encoding putative GSTs (Lenz et al., 2012) were used as queries to search the *de novo* transcriptome using the tera-tblastx algorithm. The default parameters of both tera-tblastn and tera-tblastx were used for all searches.

2.3. Protein vetting via reciprocal BLAST and structural motif analyses

To confirm that the putative proteins reported here are true members of the GST superfamily, each was subjected to a well-established vetting protocol that involved both reciprocal BLAST and structural motif analyses; this workflow is described in detail in recent publications (Lenz et al., 2014; Christie et al., 2013a, 2013b, 2014a, 2014b) . In brief, each of the *C. finmarchicus* transcripts identified as encoding a putative GST was fully translated using the ‘‘Translate’’ tool of ExPASy (http:// web.expasy.org/translate/) and then the deduced protein used as the input query for a blastp search of the non-redundant arthropod protein sequences (excluding *C. finmarchicus* proteins) curated at NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Each deduced protein was then aligned with its top blastp protein hit using MAFFT version 7 (Katoh et al., 2002, 2008, 2013), and amino acid identity/similarity between the sequences was calculated. Percent identity between two proteins was defined as the number of identical amino acids present in the alignment (represented by ‘‘*’’ in the MAFFT output) divided by the total number of amino acids in the longest sequence (x100). Amino acid similarity was defined as the number of identical and similar amino acids (the latter represented by the ‘‘:’’ and ‘‘.’’ symbols in the protein alignment) divided by the total number of amino acids in the longest sequence (x100). In the case of partial proteins, amino acid identity and similarity were calculated as described above, but only for the region of overlap.

Protein structural motifs were analyzed using the online program SMART

(http://smart.embl-heidelberg.de/) (Letunic et al., 2009; Schultz et al., 1998). Proteins were screened to confirm that each possessed the complement of structural domains expected for members of their respective GST class/subclass. In all figures showing protein sequences, the functional domains have been highlighted using a common color-coding: GST N-terminal domain, black; GST C-terminal domain, red; microsomal MAPEG domain, green. Proteins described as ‘‘full-length’’ are ones that possessed a stop codon at the 5’ end prior to the first “start” methionine and are flanked on the 3’ end by a second stop codon (or have a “start” methionine that matched the position of the initial “start” methionine in the protein query used for its identification). Proteins described as ‘‘partial’’ lacked a start methionine (referred to here as C-terminal partial proteins), a stop codon (referred to here as N-terminal partial proteins), or both of these features (referred to here as an internal protein fragment).

2.4. Comparison of *Calanus finmarchicus* GST diversity with that of selected insect/crustacean species

The collection of GSTs predicted from *C. finmarchicus* was compared to those from the fruit fly *D. melanogaster* (Saisawang et al., 2012) and the crustaceans *D. pulex* (Colbourne et al., 2010) and *T. japonicus* (Lee et al., 2008). It should be noted, that the proteins available for *T. japonicus* GSTs were derived from transcribed sequences, whereas those from both *D. melanogaster* and *D. pulex* were obtained from genomic data. Thus, the collection obtained for *T. japonicus* may be an incomplete set of GST proteins as not all may have been transcribed at the time of mRNA isolation; while those reported for *D. melanogaster* and *D. pulex* may contain ones that are not actually transcribed in the species in question.

Phylogenetic analysis was performed for GST members of the cytosolic class identified

in *C. finmarchicus* and the cytosolic GSTs from the insect *D. melanogaster* and the crustaceans *D. pulex* and *T. japonicus*. Phylogenetic trees of the cytosolic GSTs were used to establish the relationship among the subclasses in insects (Ranson and Hemingway 2005; Saisawang et al., 2012). Here, the phylogenetic tree was used to support the assignment of predicted GST proteins into subclasses and to establish their relationship to each other and to those from *D. melanogaster*, *D. pulex* and *T. japonicus*. For the construction of an unrooted phylogenetic tree, the publicly available cytosolic GST protein sequences for *D. melanogaster, D. pulex* and *T. japonicus* were downloaded from NCBI using the GenBank accession numbers listed in previous publications (Colbourne et al., 2011; Lee et al., 2008;Saisawang et al., 2012). In addition, for the completeness of the *D. pulex* dataset, GST proteins were also searched for by name ("glutathione S-transferase") and extracted from the genome assembly (daphnia_genes2010_beta3.aa.gz) accessible via wFleaBase (http://wfleabase.org/). Amino acid sequences for GST proteins were aligned using MAFFT software (Katoh et al., 2008, 2013), and resultant alignments were trimmed and corrected manually to remove non-conserved regions and obvious alignment errors. The best-fit likelihood model for each alignment was determined using ProtTest (Abascal, 2005). Phylogenetic reconstruction was performed with MrBayes 3.2 (Ronquist et al., 2003) with four independent runs of four chains each and 10,000,000 generations, using the WAG substitution model of protein evolution (Whelan and Goldman, 2001) and a gamma distribution of rates with four categories. A consensus tree was obtained by discarding the initial 2,500,000 generations as burn-in. Maximum likelihood bootstrap analysis was performed with RAxML 8 (Stamatakis, 2014), with 1,000 bootstrap replicates using the WAG substitution model of protein evolution and a gamma distribution of rates. The unrooted consensus tree from MrBayes was visualized in FigTree v1.3.1 (http://www.tree.bio.ed.ac.uk/software/figtree/) with bootstrap values >50%

reported.

2.5. Expression of GSTs during development

The relative expression of the identified *C. finmarchicus* GSTs was examined across developmental stages (embryo, early nauplius, late nauplius, early copepodite, late copepodite and adult female) as described in earlier publications (Lenz et al., 2014; Christie et al., 2013a, 2013b, 2014a, 2014b). In brief, Illumina reads for six developmental stages obtained in either 2011 (Lenz et al., 2014) or 2012 (Christie et al., 2014b) were mapped against each of the identified *C. finmarchicus* nucleotide sequences using Bowtie software (version 2.0.6; with a setting of 2 mismatches) (Langmead et al., 2009). Prior to the mapping step, reads were quality filtered using FASTX Toolkit software (version 0.013; http://hannonlab.cshl.edu/fastx_toolkit), with a Phred quality score of 20 used as the acceptance cutoff (*i.e.* low quality reads were removed from each dataset). Relative expression was computed for each transcript as reads per kilobase transcript per million reads (RPKM) using a custom written Perl script. Briefly, the total number of reads mapped to each transcript was divided by the total number of mapped reads to the reference transcriptome multiplied by the length of the transcript (Mortazavi et al., 2008)..

2.6. Comparison between two *C. finmarchicus de novo* transcriptomes

In addition to the transcriptome generated from animals obtained from the Gulf of Maine (Lenz et al., 2014), a second *de novo* transcriptome was independently generated by Tarrant and colleagues using material obtained from pre-adult (stage CV) *C. finmarchicus* and publicly deposited (Tarrant et al., 2014)c. For this transcriptome, total RNA was extracted from individuals collected from both surface waters in Trondheim fjord (Norwegian Sea) and from

individuals reared in culture, gene libraries were prepared and sequenced on the Illumina platform (Bioproject No. PRJNA2311645). The ”Norwegian Sea” transcriptome was mined for GST-encoding transcripts using the proteins deduced from the “Gulf of Maine” transcriptome and the *T. japonicus* GSTs as queries. The goal here was to verify the diversity of the putative GSTs in *C. finmarchicus* using a *de novo* transcriptome that had been generated independently, and to compare the predicted GSTs from the two populations. For these BLAST analyses, the searched database of the online program tblastn (National Center for Biotechnology Information, Bethesda, MD; http://blast.ncbi.nlm.nih.gov/Blast.cgi) was set to ‘‘Transcriptome Shotgun Assembly (TSA)’’ and restricted to sequence data from the ‘‘*Calanus finmarchicus* (taxid: 6837)’’, which allowed access to the Norwegian Sea dataset. All hits returned by a given search were translated into proteins and checked manually for homology to the target query as described earlier. Comparisons between sequences included aligning the predicted proteins with their query and determining their percent amino acid identity. When the translated proteins differed in length, percent amino acid identity was determined only for the region of overlap.

3. RESULTS

3.1.Mining of a *Calanus finmarchicus de novo* transcriptome for transcripts encoding glutathione S-transferase proteins

A total of 39 putative GST-encoding transcripts were retrieved from the Gulf of Maine *C. finmarchicus* transcriptome using known GSTs from the crustaceans *T. japonicus* (a copepod) and *D. pulex* (a cladoceran) and the insect *D. melanogaster* as queries (Table 4.1). The putative GST-encoding transcripts identified from *C. finmarchicus* included representatives of all three classes, *i.e.*, cytosolic, mitochondrial and microsomal, with the majority encoding putative

members of the cytosolic class (32 transcripts) in six subclasses (Delta, Theta, Mu, Omega, Sigma and Zeta). Transcripts encoding six microsomal GSTs (subclasses 1 and 3) and one mitochondrial (Kappa) GST were also identified (Table 4.1). Interestingly, the searches using cytosolic GST Delta, Theta and Epsilon subclass members as queries yielded identical sets of *C. finmarchicus* sequences (11 transcripts in total; Table 4.1). Using Delta and Theta GSTs from the copepod *T. japonicus* as queries, the BLAST-generated E-values for the eleven putative GST-encoding transcripts overlapped extensively and ranged from 10^{-68} to 10^{-9}. Not surprisingly, the BLAST-generated E-values for the same list of transcripts were higher using an insect-specific Epsilon GST from *D. melanogaster* as a query, and ranged between 10^{-36} and 10^{-9}. As will be presented later, reciprocal protein BLAST and phylogenetic analyses were used to resolve this apparent conundrum.

It should be noted that the EST database for *C. finmarchicus* contains five sequences annotated as GSTs. Tblastx analysis showed that two of these ESTs (Accession Nos. ES387233 and FG632831) matched two of the putative cytosolic Sigma GSTs identified here, with another (Accession No. FK671334) matching one of the cytosolic Delta sequences (see below), and a fourth (Accession No. ES387262) matching a microsomal GST, with amino acid identity >90% for each of the respective pairs (in bold in Table 4.1). The fifth EST annotated as a GST (Accession No. ES387185) did not generate significant hits from the *Calanus* transcriptome, and a subsequent blastp search of the non-redundant protein database suggests that the protein encoded by this EST may not be a GST. The predicted protein is only 43 amino acids long, and while it is most similar to the C-terminus of a GST from the nematode *Caenorhabditis brenneri* (Accession No. EGT40878), the E-value is very high (10^{-4}).

Class and subclass assignments of the GSTs in Table 4.1 were confirmed by translating each sequence into a predicted protein, followed by reciprocal BLAST and structural analyses.

3.1.1. Cytosolic GSTs

Delta, Epsilon and Theta subclasses

Eight full-length and three partial proteins were predicted from the 11 transcripts putatively identified in the original searches as belonging to either the Delta, Theta or Epsilon GST subclass (Table 4.2). Structural analysis confirmed the presence of GST N-terminal and GST C-terminal domains in all of the predicted full-length proteins. The three partial proteins possessed the expected complement of domains consistent with their incomplete nature (Table 4.2, Figure 4.1A).

Reciprocal BLAST analysis identified ten of the cytosolic GSTs as members of the Delta subclass, with nine of these putative proteins returning Delta GSTs from other copepod species as the top BLAST hit. With respect to these proteins, five, Calfi-Delta-I, Calfi-Delta-II, Calfi-Delta-III, Calfi-Delta-VI and Calfi-Delta-VII, were found to be most similar to Delta GSTs from *L. salmonis*, while two, Calfi-Delta-V and Calfi-Delta-VIII, were most similar to a Delta GST from *T. japonicus* (Table 4.2). The tenth protein, Calfi-Delta-IV, was identified as most similar to a Delta GST from the cladoceran *D. pulex* (Table 4.2). The reciprocal BLAST of the eleventh protein in the GST Delta/Theta/Epsilon list (Table 4.1) identified it as a GST in the Theta subclass, being most similar to a Theta GST protein of the insect *Locusta migratoria* (Table 4.2). None of the 11 GSTs resulting from the Delta/Theta/Epsilon searches (see above) were found to be members of the Epsilon subclass, which is consistent with the hypothesis that this subclass is insect-specific (Shi et al., 2012).

Interestingly, a GST initially identified via transcriptome mining as a cytosolic Zeta subclass GST (Accession No. GAXK01204939) was ultimately determined via reciprocal BLAST analysis to be a member of the Delta subclass; the top hit returned for this protein, Calfi-Delta-XI, was a cytosolic Delta GST from the copepod *L. salmonis* (Table 4.2).

Alignments of each *C. finmarchicus* putative Delta GST with its respective top BLAST hit showed 21%-69% amino acid identity and 54%-89% amino acid similarity for the full-length proteins (Table 4.2, Figure 4.1A). Alignments of the regions of overlap between the partial *C. finmarchicus* sequences and their top BLAST hits revealed 33%-44% identity and 44%-68% similarity in amino acid sequence (Table 4.2, Figure 4.1A). Similarly, alignment of Calfi-Theta-I with its top protein hit showed 45% amino acid identity and 72% amino acid similarity between the two proteins (Table 2). Pairwise alignments of four *C. finmarchicus* Delta GSTs (Calfi-Delta-I, Calfi-Delta-II, Calfi-Delta-III and Calfi-Delta-VI) that had the identical top hit from the copepod *L. salmonis* (Table 4.2) showed that these predicted proteins shared only 27%-50% amino acid identity with each other. Likewise, alignment of Calfi-Delta-V and Calfi-Delta-VIII, which both shared the same *T. japonicus* Delta GST as their top protein hits, showed only 33% amino acid identity between the two proteins. The large differences in amino acid sequence among these *C. finmarchicus* Delta subclass GSTs are consistent with the Trinity software assembly results that placed the transcripts that encode them into unique “comps” which represent transcripts encoded by different genes (Haas et al., 2013), a finding that is consistent with the multiplicity of insect GST genes.

Mu subclass

Five full-length proteins were predicted from the five transcripts identified in the original search as encoding putative members of the Mu subclass (Table 4.2). Each of these proteins possesses the conserved GST N-terminal and GST C-terminal domains (Table 4.2, Figure 4.1B). Reciprocal BLAST analysis confirmed the five proteins as members of the Mu subclass, with four of the proteins returning as top BLAST hits Mu GSTs from other copepod species, and one, Calfi-Mu-II, a Mu GST from the river prawn, a decapod crustacean (Table 4.2). Three of these proteins (Calfi-Mu-III, Calfi-Mu-VI and Calfi-Mu-V) were found to be most similar to the *T. japonicus* Mu GST that was used in the initial search of the transcriptome (Table 4.2).

Alignments of each of the putative *C. finmarchicus* Mu GSTs with its respective top hit revealed 45%-65% amino acid identity and 69%-89% amino acid similarity between the protein pairs (Table 4.2). Pairwise alignments of the three Mu GSTs (Calfi-Mu-III, Calfi-Mu-VI and Calfi-Mu-V) that had the identical top hit from the copepod *T. japonicus* (Table 4.2) showed that these predicted proteins shared 34%-60% amino acid identity, suggesting that different genes encode them.

Omega subclass

One full-length and two partial proteins were predicted from the three transcripts identified in the original search as encoding putative Omega subclass GSTs (Table 4.2). Structural analysis confirmed the presence of GST N-terminal and GST C-terminal domains for the full-length protein, while the two partial proteins possessed only the N-terminal domain (Table 4.2). Results from the reciprocal BLAST analysis identified the three predicted proteins as members of the Omega subclass, returning Omega GSTs from ants as the top BLAST hits (Table 4.2). Alignment of each *C. finmarchicus* putative Omega GST and its top BLAST hit

revealed amino acid identities/similarities ranging from 35%-36% and 64%-68%, respectively (Table 4.2).

Sigma subclass

Ten full-length proteins were predicted from the 10 transcripts putatively identified in the initial search as encoding members of the Sigma subclass (Table 4.2). Structural analysis confirmed the presence of the GST N-terminal and GST C-terminal domains in each protein (Table 4.2).

Reciprocal BLAST analysis confirmed all 10 predicted proteins as members of the Sigma subclass. Seven of the *C. finmarchicus* proteins returned Sigma GSTs from other crustaceans as their top BLAST hits, while three returned Sigma GSTs from insects as the most similar proteins (Table 2). Specifically, five of the *C. finmarchicus* proteins (Calfi-Sigma-I, Calfi-Sigma-II, Calfi-Sigma-VI, Calfi-Sigma-VII and Calfi-Sigma-VIII) were found to be most similar to Sigma GSTs from *D. pulex*, with two (Calfi-Sigma-III and Calfi-Sigma-V) most similar to a Sigma GST from the copepod *T. japonicus* (Table 2). Calfi-Sigma-IV, Calfi-Sigma-IX and Calfi-Sigma-X returned Sigma GSTs from the insects *Apis florea*, *Folsomia candida* and *Megachile rotundata,* respectively, as their top BLAST hits (Table 4.2). Alignments of each *C. finmarchicus* putative Sigma GST with its respective top hit revealed 35%-45% amino acid identity and 51%-76% amino acid similarity between the protein pairs (Table 4.2). Pairwise alignments of the three *C. finmarchicus* Sigma GSTs (Calfi-Sigma-II, Calfi-Sigma-VI and Calfi-Sigma-VIII) that had the identical top BLAST hit (Table 4.2) showed only 36%-42% amino acid identity to each other. Likewise, alignment of Calfi-Sigma-III and Calfi-Sigma-V, both most similar to the same *T. japonicus* Sigma GST, showed 42% amino acid identity between the two proteins.

Zeta subclass

One full-length protein and two partial proteins were predicted from the three transcripts identified in the original search as putatively encoding Zeta subclass GSTs. The partial protein encoded by transcript GAXK01204939 was found to be most similar to a Delta GST, and was assigned to the Delta subclass accordingly (Calfi-Delta-XI, see above). Structural analyses of the two remaining proteins (Calfi-Zeta-I and Calfi-Zeta-II), confirmed the presence of GST N-terminal and GST C-terminal domains in the full-length protein and the GST C-terminal domain in the partial sequence (Table 4.2). Reciprocal BLAST analyses identified these two proteins as members of the Zeta subclass, returning Zeta GSTs from the insects *D. melanogaster* and *Bactrocera dorsalis* as the top hits, respectively (Table 4.2). Alignment of Calfi-Zeta-I with its top hit revealed 54% amino acid identity and 82% amino acid similarity; alignment of the extant sequence of Calfi-Zeta-II and the corresponding portion of its top hit revealed 22% amino acid identity and 32% amino acid similarity (Table 4.2).

3.1.2. Mitochondrial class

A single full-length protein was predicted from the transcript putatively identified as encoding a mitochondrial Kappa GST (Table 4.2). Structural analysis revealed that this protein possesses a mitochondrial GST thioredoxin-like domain, which is typical of mitochondrial GSTs (Table 4.2). Reciprocal BLAST analysis identified the protein as a member of the mitochondrial class, returning a mitochondrial Kappa GST from the copepod *P. nana* as its top BLAST hit (Table 4.2). Alignment of the *C. finmarchicus* mitochondrial Kappa GST with its top hit showed 36% amino acid identity and 61% similarity between the two proteins (Table 4.2).

3.1.3. Microsomal class

mGST-1

One full-length and one partial protein were predicted from the two transcripts identified in the initial search as belonging to the microsomal GST subclass 1 (Table 4.2). Reciprocal BLAST analyses identified both proteins as subclass 1 microsomal GSTs, returning microsomal GST-1s from insects as the top BLAST hits (Table 4.2). Alignment of Calfi-mGST-1-I with its top hit revealed 36% amino acid identity/61% amino acid similarity between the two proteins; 31% amino acid identity/52% amino acid similarity was seen between the known portion of Calfi-mGST-1-II and its top hit (Table 4.2).

Structural analysis identified a single MAPEG domain with the typical four transmembrane regions in both *C. finmarchicus* mGST-1 proteins (Table 4.2; Figure 4.2A). Within the conserved MAPEG region, microsomal GST-1 proteins are characterized by an amino acid pattern that is shared by both arthropods and vertebrates (Bresell et al., 2005, Shi et al., 2012; Holm et al., 2006; , Zhou et al., 2013).The pattern consists of a highly conserved motif of 16 amino acids (VERVRR*X*HLND*X*ENI*X*) where the three *X*s represent variable amino acids (Bresell et al., 2005). The *C. finmarchicus* microsomal GST-1 proteins identified here (Calfi-mGST-1-I and Calfi-mGST-1-II) were aligned with mGST-1 amino acid sequences from other crustaceans, specifically the copepods *C. clemensi*, *C. rogercresseyi*, *L. salmonis* and *T. japonicus,* and the cladoceran *D. pulex* (Figure 4.2B). This alignment showed that the 16 amino acids motif VERVRR*X*HLND*X*ENI*X* was conserved in all crustaceans except for *C. finmarchicus*. In both *C. finmarchicus* sequences, there was a non-conservative substitution in the 9th amino acid of the motif, specifically the stereotypical hydrophobic leucine (L) was

substituted by a hydrophilic glutamine (Q) residue (Figure 4.2B). This amino acid substitution was also present in a protein predicted from the Norwegian Sea transcriptome (Accession No. GBFB01067142; see below). Thus, this observed amino acid substitution is unlikely to be an assembly artifact, and may be *C. finmarchicus*-specific (Figure 4.2B).

mGST-3

Four proteins, three full-length and one partial, were predicted from four transcripts identified as encoding putative members of microsomal GST subclass 3 (Table 4.2). In all four proteins, structural analysis confirmed the presence of the MAPEG domain (Table 4.2). Reciprocal BLAST analysis identified these proteins as microsomal GST subclass 3 members, with each protein returning a crustacean mGST-3 as its top BLAST hit (Table 4.2). Two of the proteins, Calfi-mGST-3-I and Calfi mGST-3-IV were found to be most similar to a mGST-3 from the copepod *Acartia pacifica*, while Calfi-mGST-3-II and Calfi-mGST-3-III were most similar to a mGST-3 from *D. pulex* (Table 4.2). The percent amino acid identity/similarity between each of the *C. finmarchicus* mGST-3 and its top BLAST hit was 42%-65%/73%-84% (Table 4.2). Alignment of the two *C. finmarchicus* mGST3 (Calfi-mGST-3-I and Calfi-mGST-3-IV) that shared the same top hit showed 77% of amino acid identity between the two proteins.

3.2. Glutathione S-transferase diversity in *C. finmarchicus*

The identification of 39 putative *C. finmarchicus* GSTs from the Gulf of Maine transcriptome suggests that the gene complexity found in this copepod species is comparable to that of the insect *D. melanogaster* (40 GST genes) and higher than that of the crustacean *D. pulex* (31 GSTs genes; Table 4.3) (Colbourne et al., 2011; Saisawang et al., 2012). Comparison

between *C. finmarchicus* and *D. pulex* indicates that the number of genes in some subclasses, *i.e.*, the cytosolic Sigma and Theta subclasses, as well as in the microsomal GST-1 group, is very similar (Table 4.3). However, the gene duplication found in the *C. finmarchicus* cytosolic Delta subclass is higher than that reported for *D. pulex*, and is identical to the complexity seen in the insect *D. melanogaster*, which has a total of 11 Delta GST genes (Table 4.3). The complexity of GSTs reported for *T. japonicus*, another member of the Copepoda, is lower than that found for *C. finmarchicus*, although this may be a function of sequencing depth, since the current *T. japonicus* transcriptome data are more limited.

Phylogenetic analysis based on Bayesian likelihood criteria places the deduced *C. finmarchicus* cytosolic GSTs (see above) into distinct clades (Figure 4.3), which are consistent with their classification into different subclasses. Members of the cytosolic subclasses Delta, Omega, Zeta, Mu, Sigma and Theta were identified in the phylogenetic tree with good bootstrap support (>50% for most; Figure 4.3). In the consensus tree, the Delta, Omega, Zeta, Mu and Theta subclasses were each recovered as monophyletic groups with bootstrap support >90% for many. The Sigma subclass was also recovered as monophyletic, with a posterior probability of P> 0.8 (data not shown), and with bootstrap support >50% for most of the branches. The Delta GSTs were recovered as monophyletic with bootstrap support >90%, but nested within the Epsilon GSTs from *D. melanogaster*. Despite this, none of the predicted cytosolic GSTs from *C. finmarchicus*, *T. japonicus*, or *D. pulex* were recovered as most closely related to individual members of the poorly resolved Epsilon subclass, consistent with this subclass being absent in these crustaceans (Shi et al., 2012).

The clustering pattern within individual subclasses varied, but in many cases all, or at least a large subset, of the *C. finmarchicus* GSTs within a subclass were located on a single

branch. For example, in the Delta clade with 11 *C. finmarchicus* GSTs, the majority (nine: Calfi-Delta-V, Calfi-Delta-IX, Calfi-Delta-VIII, Calfi-Delta-I, Calfi-Delta-VI, Calfi-Delta-XI, Calfi-Delta-II, Calfi-Delta-III and Calfi-Delta-X) fell into a single cluster, which was shared with two Delta GSTs from the copepod *T. japonicus* (>90% bootstrap support) (Figure 4.3). The remaining two Delta GSTs (Calfi-Delta-IV and Calfi-Delta-VII) were on separate branches grouped with *D. pulex* GSTs with 50% bootstrap support (Figure 4.3). The second largest diversity of GSTs was found in the Sigma subclass, which grouped into two separate clusters (Figure 4.3), one of which consisted exclusively of *C. finmarchicus* predicted proteins. A single *C. finmarchicus* Sigma GST (Calfi-Sigma-VIII [Cf_S8]) did not cluster with any of the others, and was most similar to a *D. pulex* Sigma GST, which was also located on its own branch (Figure 4.3).

3.3. Expression of GSTs during development

Relative expression of GSTs varied across developmental stages (Figure 4.4), as well as among GSTs. We observed some differences in relative expression between the two years of sample collection, although in general expression patterns were consistent between years (Figure 4.4). Expression levels ranged from very low to high with RPKM values ranging between 1 and 14 (Log_2).

Relative expression in members of the mitochondrial (Calfi-Kappa-I) and microsomal (Calfi-mGST-1-I and Calfi-mGST-3-III) classes was moderately low, but similar across developmental stages except for embryos (Figure 4.4). Relative expression levels of cytosolic GSTs were more variable across life stages, with most GSTs showing low expression in embryos (Figure 4.4). Calfi-Delta-III and Calfi-Sigma-IX were the most highly expressed (RPKM Log_2

between 9 and 11) among the cytosolic GSTs, and peak expression was observed in the adult female and late copepodite stages. In Calfi-Delta-I, expression levels were lower, but showed a similar peak in expression in adult females and late copepodites (Figure 4.4).

3.4. Gulf of Maine vs. Norwegian Sea: Comparison between two *C. finmarchicus de novo* transcriptomes

In addition to the *C. finmarchicus* transcriptome generated from material obtained from the Gulf of Maine (Lenz et al., 2014), there is a second *de novo* transcriptome generated from animals from the Norwegian Sea (tarrant et al., 2014). A total of 39 putative GST-encoding transcripts were retrieved from the Norwegian Sea transcriptome using the GSTs identified in the Gulf of Maine assembly and the known *T. japonicus* GSTs as queries (see section above), confirming a similar diversity in GSTs in the two transcriptomes, and hence two populations.

Comparisons between predicted proteins from the two *C. finmarchicus* populations found good one-to-one correspondence for the majority of the GSTs in the cytosolic, mitochondrial and microsomal classes (Table 4.4). Pairwise alignment of the Gulf of Maine query with its Norwegian Sea hit showed that for 36 putative GSTs there was high amino acid conservation (> 90% identity) between the predicted proteins from the two transcriptomes in their regions of overlap (Table 4.4). This included all predicted GSTs in several cytosolic GST subclasses, *e.g.*, Sigma (10 proteins), Theta (1 protein), Mu (5 proteins), and Omega (3 proteins), as well as the mitochondrial Kappa GST and the microsomal ones in the mGST-3 subclass (4 proteins; Table 4.4). The high amino acid identity found between cytosolic GST members of the two populations is in contrast to the amino acid identity between cytosolic GST members within the same subclass, which is lower (see above).

The second transcriptome not only confirmed the presence of the GSTs, but also provided additional data. In four cases (Calfi-Delta-VIII, Calfi-Delta-IX, Calfi-Sigma-VI and Calfi-mGST-3-IV) the transcripts from the Norwegian Sea transcriptome predicted full-length proteins, while the transcripts identified from the Gulf of Maine assembly encoded only partial ones (Table 4.4). In another three cases, genetic differences between the two transcriptomes were larger than expected. In the first case, there appeared to be an additional Omega transcript in the Norwegian Sea transcriptome (Table 4.5). Protein translation and structural analyses confirmed that the protein was full-length and possessed the typical structural hallmarks (N- and C-terminal domains) of a cytosolic Omega subclass member. A reciprocal BLAST search of the non-redundant arthropod protein database identified its top protein hit as an Omega GST from the insect *L. migratoria* (Accession No. AFK10494). Pairwise comparison of this fourth Omega GST with the other three Omega GSTs (Table 4.4) showed only 30%-37% amino acid identity, suggesting that this transcript represents an additional gene in this subclass. A search of the Gulf of Maine transcriptome using this fourth Omega protein as the query yielded a short nucleotide sequence (504 base pairs) which encoded a partial protein that was 99% identical in sequence to the corresponding portion of the query, confirming the presence of this Omega GST in both transcriptomes (Table 4.5). Thus, *C. finmarchicus* appears to have four genes encoding GSTs in the Omega subclass.

Large differences in amino acid sequences were found for two Delta GSTs, Calfi-Delta-IV and Calfi-Delta-XI, when paired with their top hits in the Norwegian Sea transcriptome (< 90% identity) (Table 4.5). In one case (Calfi-Delta-IV), amino acid identity was only 48% between the Gulf of Maine protein and its Norwegian Sea counterpart (Table 4.5). This level of amino acid identity is similar to the one we observed among different members of the same

subclass (usually 30%-50%), suggesting that this cytosolic GST may be derived from a separate gene, and thus represent a 12th gene in the Delta sub-class. More difficult to interpret is the 88% amino acid identity found between Calfi-Delta-XI and its Norwegian counterpart (Table 4.5); the two proteins did not fall into what had been previously defined as "good one-to-one correspondence" (≥ 90% amino acid identity in region of overlap) but, nevertheless, they shared more then the expected amino acid identity (30%-50%) between subclass members. Thus, if these two Delta GSTs are derived from the same gene, they show significant genetic divergence between the two populations.

In summary, comparison of the two transcriptomes yielded a more complete set of predicted GSTs for *C. finmarchicus.* By combining the two data sets, we have predictions for 36 full-length proteins (88%) and five partial ones. Thirty-nine of these GSTs showed good to excellent amino acid identity (88%-100%) between transcriptomes, and hence populations. Two proteins were found in the Gulf of Maine transcriptome but not in the Norwegian Sea transcriptome. Two additional cytosolic genes were predicted from the Norwegian Sea transcriptome that were absent in the Gulf of Maine dataset, bringing the gene diversity in the Delta subclass to 12 and the Omega subclass to 4 predicted proteins. Based on these two transcriptomes, *C. finmarchicus* is predicted to have a total of 41 GSTs.

4. DISCUSSION

The GSTs belong to a gene superfamily that is present in both prokaryotes and eukaryotes (Habig et al., 1974). In the arthropods, this superfamily is characterized by multiple gene duplications, leading to a diverse set of genes, some of which have been shown to be rapidly evolving in response to natural selection, such as exposure to new insecticides

(Friedman, 2011). Genome sequencing and bioinformatics-based data mining have been a powerful strategy for the discovery and characterization of GSTs. In insects, the number of GST genes varies widely with 13 genes reported in *Apis mellifera*, 23 in *B. mori*, 31 in *A. gambiae*, 40 in *D. melanogaster* and 41 in *T. castaneum* (Shi et al., 2012). Among the crustaceans, the cladoceran *D. pulex* is the only species with a sequenced genome, and its GST superfamily consists of 31 genes (Coulburne et al., 2011). Here, we identified putative GSTs belonging to the cytosolic (34 proteins), microsomal (6 proteins) and mitochondrial (1 protein) classes in the calanoid copepod *C. finmarchicus* by mining two *de novo* transcriptomes using a workflow that included reciprocal BLAST and protein structural analyses. This number is much higher than the twelve GSTs that were identified and classified by *in silico* EST mining in *T. japonicus* (Lee et al., 2008) (Roncalli, unpublished), and the 12 GSTs identified in a search of publicly available ESTs of *C. clemensi* (search completed 12/04/2014; Roncalli, unpublished). However, these may be underestimates given the limited size of the EST databases available for *T. japonicus* and *C. clemensi*. More recently, transcriptome shotgun assemblies have been made available for several copepods, including *L. salmonis* and *C. rogercresseyi*. Searches for "glutathione S-transferase" in these TSA databases on NCBI (search date: 12/04/2014) resulted in 34 transcripts annotated as encoding GST proteins in *L. salmonis* (Bioproject No. PRJNA73429) and 35 in *C. rogercresseyi* (Bioproject No. PRJNA234316). Yang et al. (2014) reported 31 GST proteins in the *de novo* transcriptome of the calanoid *Calanus sinicus*, but to date, these data are not publicly accessible. None of these studies included annotations by GST class or subclass, or protein structural analyses. However, in general, it appears that the number of GST genes in these copepod species exceeds 30 based on automated annotations of TSA data (e.g. Yang et al., 2014).

Although two conserved domains characterize all cytosolic GSTs irrespective of subclass (Sheehan et al., 2002), these proteins are nevertheless highly diverse. In the insects, cytosolic GST members belonging to the same subclass within a species have typically 40%-50% amino acid identity (Hayes and Pulford, 1995).We found a similar pattern in *C. finmarchicus*, where even cytosolic GSTs with identical top hits were quite different from each other, with amino acid identity ranging from 27%-60%, supporting the conclusion that each of the 34 cytosolic GSTs represents a transcript from a separate gene. In contrast, when we compared GSTs obtained from two separate transcriptomes, the predicted proteins were much more similar. Twenty-two (56%) of the predicted proteins were 99%-100% identical, while seventeen showed moderate differences with 88%-98% identity in amino acid sequence in the region of overlap. The transcriptomes were generated from mRNA from individuals from two populations of *C. finmarchicus* (Gulf of Maine and Norwegian Sea) that are separated by over 4,000 km, and these two populations are mostly isolated from each other (Bucklin et al., 1996; Unal et al., 2010). Population genetic studies suggest two to four genetically distinct *C. finmarchicus* populations across the North Atlantic with no direct genetic exchange between the Gulf of Maine and Norwegian Sea (Bucklin et al., 1996; Unal et al., 2010). However, there is evidence for genetic connectivity via the central North Atlantic with genetic exchanges between this *C. finmarchicus* population and the ones in the Labrador Sea/Gulf of Maine in the western Atlantic and the Norwegian Sea in the eastern Atlantic (Unal et al., 2010). Thus, the observed differences in GST protein predictions from the two transcriptomes are not surprising, given that genes in this superfamily are often under natural selection and have been shown to evolve rapidly in other arthropods (Sheehan et al., 2001; Che-Mendoza et al., 2009). However, whether the genetic

variation in *C. finmarchicus* represents differences in function in response to habitat-specific selection has yet to be determined.

Glutathione S-transferases are best known for their role in detoxification of xenobiotics, although other functions have been described (da Fonseca et al., 2010).Given the diversity of environmental toxins and pollutants, and their variable levels of toxicity, it has been hypothesized that the need to metabolize a variety of xenobiotics has driven the expansion of the cytosolic GSTs (da Fonseca et al., 2010). In insects, the subclasses Delta and Epsilon are responsible for the removal of chemical compounds produced by either their food or by pesticides (Ayres et al., 2011; Ortelli et al., 2003). The number of GSTs in the Delta subclass is variable: some species have just a few, *e.g.*, *A. mellifera* (2), *B. mori* (5) and *T. castaneum* (3), while others have over ten, *e.g.*, *Acyrthosiphon pisum* (16), *A. gambiae* (17) and *D. melanogaster* (11). In *C. finmarchicus*, the Delta GST subclass is large with a total of 12 different proteins predicted. If the function of the Delta GSTs in *C. finmarchicus* is similar to that of the insects, extensive gene duplication may have occurred in response to environmental toxins encountered by this copepod. *C. finmarchicus* is a filter feeder and it consumes a variety of microplankton including diatoms, dinoflagellates, flagellates, ciliates and protozoans (Harris et al., 2005; , Stoecker and Capuzzo, 1990). Many common food types such as dinoflagellates and diatoms are known to produce toxic secondary metabolites as defense against predators, competitors and pathogens (Ianora et al., 2012). Although it has been demonstrated that copepods can feed selectively (Koehl and Strickler, 1981) and thus might be able to avoid consuming toxic species, there is good evidence that copepods, including *C. finmarchicus*, ingest toxic species during natural blooms(Turner, 2014).

In the Gulf of Maine, *C. finmarchicus* frequently encounters algal blooms dominated by

the toxic dinoflagellate *Alexandrium fundyense*, known for the production of saxitoxins, which are highly toxic to humans, birds, fishes and marine mammals (Anderson, 1997; 2008; Llewellyn, 2006). *C. finmarchicus* ingests *A. fundyense* with no detrimental effects on its survival (Teegarden et al., 2001, 2008; Campbell et al., 2005; Turner et al., 1983). Spring blooms dominated by diatoms in the genera *Thalassiosira, Skeletonema* and *Chaetoceros* spp. are common in both the Gulf of Maine and the Norwegian Sea (Gettings et al., 2014;, Bratbak et al., 2011). These diatom genera are known for their production of oxylipins, which are toxic at high concentrations to other copepods, such as the congener *C. helgolandicus* (Miralto et al., 1999; d'Ippolito et al., 2009; Fontana et al., 2007). Thus, *C. finmarchicus* inhabiting either the Gulf of Maine or the Norwegian Sea are likely to experience a wide range of natural toxins during their life cycle given a diet that includes phytoplankton species producing a variety of metabolites. The high gene diversity in the Delta GST subclass, which is involved in detoxification, may represent a fitness advantage for *C. finmarchicus*.

Sigma represents the second largest subclass with 10 putative GSTs in *C. finmarchicus*. A similar number of Sigma GSTs are present in the cladoceran *D. pulex* (Colbourne et al., 2011), but the diversity in insects is typically lower and ranges from a single gene (*D. melanogaster* and *A. gambiae*) to six (*A. pisum*) or seven (*T. castaneum*) (Shi et al., 2012). The Sigma GST subclass plays an important role in the protection against oxidative stress in insects (Ranson and Hemingway, 2005). However, it is less clear why this subclass is so diverse in the crustaceans, and the function of individual Sigma GSTs has yet to be investigated even in model species like *D. pulex*. The phylogenetic relationship among the Sigma GSTs (Figure 4.3) showed species-specific clustering of the *D. pulex* Sigma GSTs and the majority (6) of the *C. finmarchicus* Sigma GSTs. Further studies are needed to determine whether high diversity in Sigma GSTs is

common in all crustaceans, and to establish their physiological functions.

In addition to their role in detoxification of exogenous compounds, GSTs play a role during development (Qin et al., 2013). A peak in expression was found in the pre-pupal and pupal stages in Sigma GSTs in the insects *Mayetiola destructor*, *Lucilia cuprina* and *Agrilus planipennis*, presumably in response to an increase in metabolic activity and apoptosis associated with the morphological changes that occur during these periods (Mittapalli et al., 2007; Pal et al., 2012; Rajarapu et al., 2013). Similarly, detoxification from byproducts produced during metamorphosis may explain high relative expression of Delta GSTs in the insects *D. melanogaster*, *A. planipennis* and *Nilaparvata lugens* during the pupal stage (Zhou et al., 2013;,Rajarapu et al., 2013). Copepods, like insects, undergo a significant morphological rearrangement between the 6th naupliar and 1st copepodite stages (Mauchline, 1998). This change in morphology occurs during a molt cycle, and does not involve a pupal stage as in the insects. No significant changes in expression level in either Delta or Sigma GST-encoding transcripts correlated with this transition. Instead, we found highest expression of cytosolic GSTs in the CV and adult female stages. One possible explanation is that in our sample, these late stages were field collected, and thus they had been exposed to a mix of phytoplankton species, while the early developmental stages were laboratory reared on a single algal species (Lenz et al., 2014). The difference in expression may be related to exposure to natural toxins in the field-collected animals.

5. SUMMARY AND CONCLUSIONS

Using two *de novo* assembled transcriptomes, transcripts encoding 41 distinct GST proteins were identified for the copepod *C. finmarchicus*. The deduced proteins included

members of the cytosolic, mitochondrial and microsomal classes, with the highest diversity observed in the cytosolic class. The transcripts/proteins likely represent the products of distinct genes, and if true, the diversity of GST in *C. finmarchicus* exceeds or rivals that described for insects and other crustaceans. The food sources and life history of *C. finmarchicus* are likely factors driving selection for this diversity, as this copepod is commonly exposed to a wide variety of natural toxins, and hence multiplicity in detoxification pathway proteins may well be key to their survival. Characterization of the GST superfamily in *C. finmarchicus* opens opportunities for functional studies of detoxification, and provides a diverse set of biomarkers for this species. These biomarkers will likely be useful for future studies evaluating ecosystem health and organism-environment interactions in the North Atlantic, an area that is regularly challenged by a variety of natural and anthropogenic stressors.

ACKNOWLEDGMENTS

I wish to extend our appreciation to the many colleagues who generously contributed to this study from the initial planning stages to its completion. In particular, we would like to thank Myriam Belanger and Roger Nilsen (Georgia Genomics Facility, University of Georgia), Le-Shin Wu (National Center for Genome Analysis Support, University Information Technology Services, Indiana University), Bradley Jones and Michelle Jungbluth (University of Hawaii at Manoa), and Julian Hartline (www.julianhartline.com).

Table 4.1. Search results from *in silico* mining of a *de novo* transcriptome from *Calanus finmarchicus* using glutathione S-transferase (GST) queries obtained from the copepod *Tigriopus japonicus*, the cladoceran *Daphnia pulex* and the insect *Drosophila melanogaster*. Transcript lengths (base pairs) are given for each transcript identified.

Class	Subclass	Transcript accession number	Transcript length†
Cytosolic	Delta/Theta/Epsilon*	GAXK01204954	888
		GAXK01204965	921
		GAXK01204950	852
		GAXK01204940	965
		GAXK01204947	1182
		GAXK01204957	902
		GAXK01204968	764
		GAXK01204953	991
		GAXK01073468	401
		GAXK01096295	914
		GAXK01035521	823
	Mu	GAXK01204944	957
		GAXK01204956	741
		GAXK01204948	742
		GAXK01204952	1686
		GAXK01204958	785
	Omega	GAXK01204960	984
		GAXK01016325	1410
		GAXK01164502	894
	Sigma	**GAXK01204949**	1346
		GAXK01204945	839
		GAXK01204943	815
		GAXK01204951	733
		GAXK01204959	955
		GAXK01204946	1022
		GAXK01204964	771
		GAXK01204961	884
		GAXK01204942	756
		GAXK01204966	881
	Zeta	GAXK01204939	1033
		GAXK01204941	2790
		GAXK01084871	359
Mitochondrial	Kappa	GAXK01046934	1108
Microsomal	mGST-1	GAXK01178264	771
		GAXK01081966	347
	mGST-3	GAXK01204963	54
		GAXK01204967	465
		GAXK01204962	7866
		GAXK01204955	680

†Length in nucleotides.
Accession Nos for query proteins: GST Delta *T. japonicus* (ACE81244, ACE81245). GST Theta *T. japonicus* (ACE81253). GST Epsilon *D. melanogaster* (AAF57701). GST Mu *T. japonicus* (ACE81251, ACE81252, ACE81254). GST Omega *T. japonicus* (ACE81246). GST Sigma *T. japonicus* (AAY89316). GST Zeta *T. japonicus* (ACE81250). Mitochondrial GST Kappa *D. pulex* (EFX86155). Microsomal GST-1 *T. japonicus* (ACE81248). Microsomal GST-3 *T. japonicus* (ACE81249) as listed.
*Queries for Delta, Theta and Epsilon subclasses deduced the identical list of 11 transcripts.
C. finmarchicus transcripts whose accession numbers are shown in bold font have sequence support from expressed sequence tag data (see text).

Table 4.2. Annotation of putative glutathione S-transferase-encoding transcripts from Table 4.1, using reciprocal BLAST results and protein domain analysis. BLAST searches were limited to NCBI non-redundant protein database for arthropods (taxid: 6656).

Class – subclass and assigned protein name	Transcript accession No.	Deduced protein length†	Structural domains	Species	Protein accession No.	E-value	% amino acid identity/ similarity
Cytosolic – Delta							
Calfi-Delta-I	GAXK01204953	217*	GSTN, GSTC	*Lepeophtheirus salmonis*	ACO12967	8.9e-44	40/78
Calfi-Delta-II	GAXK01204968	217**	GSTN, GSTC	*Lepeophtheirus salmonis*	ACO12967	8.9e-36	43/77
Calfi-Delta-III	GAXK01204940	218*	GSTN, GSTC	*Lepeophtheirus salmonis*	ACO12967	7.0e-33	39/80
Calfi-Delta-IV	GAXK01204965	220**	GSTN, GSTC	*Daphnia pulex*	EFX81633	2.8e-29	69/70
Calfi-Delta-V	GAXK01204954	221*	GSTN, GSTC	*Tigriopus japonicus*	ACE81245	2.8e-29	57/86
Calfi-Delta-VI	GAXK01204957	237**	GSTN, GSTC	*Lepeophtheirus salmonis*	ACO12967	1.1e-28	38/80
Calfi-Delta-VII	GAXK01204947	262*	GSTN, GSTC	*Lepeophtheirus salmonis*	ADD38060	3.2e-25	59/89
Calfi-Delta-VIII	GAXK01073468	113***	GSTN	*Tigriopus japonicus*	ACE81245	7.5e-27	33/44
Calfi-Delta-IX	GAXK01204950	178***	GSTC	*Caligus clemensi*	ACO15541	3.2e-17	44/68
Calfi-Delta-X	GAXK01035521	201*	GSTN, GSTC	*Caligus rogercresseyi*	ACO15749	2.2e-10	21/54
Calfi-Delta-XI	GAXK01204939	330**	GSTN, GSTC	*Lepeophtheirus salmonis*	ADD38823	1.3e-07	27/54
Cytosolic – Theta							
Calfi-Theta-I	GAXK01096295	262**	GSTN, GSTC	*Locusta migratoria*	AEB91980	1.3e-79	45/72
Cytosolic – Mu							
Calfi-Mu-I	GAXK01204944	222*	GSTN, GSTC	*Caligus clemensi*	ACO15225	2.2e-105	65/89
Calfi-Mu-II	GAXK01204956	222**	GSTN, GSTC	*Macrobrachium nipponense*	AGJ70295	3.1e-95	65/85
Calfi-Mu-III	GAXK01204948	222**	GSTN, GSTC	*Tigriopus japonicus*	ACE81254	1.0e-90	48/69
Calfi-Mu-IV	GAXK01204952	222*	GSTN, GSTC	*Tigriopus japonicus*	ACE81254	9.5e-90	58/82
Calfi-Mu-V	GAXK01204958	222*	GSTN, GSTC	*Tigriopus japonicus*	ACE81254	4.1e-59	45/76
Cytosolic – Omega							
Calfi-Omega-I	GAXK01204960	268*	GSTN, GSTC	*Riptortus pedestris*	BAN21163	5.3e-46	36/64
Calfi-Omega-II	GAXK01164502	235***	GSTN	*Acromyrmex echinatior*	EGI63780	3.2e-32	35/64
Calfi-Omega-III	GAXK01016325	272***	GSTN	*Coptotermes formosanus*	AFZ78680	2.6e-51	36/68
Cytosolic – Sigma							
Calfi-Sigma-I	GAXK01204964	200**	GSTN, GSTC	*Daphnia pulex*	EFX82687	3.1e-44	36/70
Calfi-Sigma-II	GAXK01204942	204*	GSTN, GSTC	*Daphnia pulex*	EFX82672	2.0e-42	40/67
Calfi-Sigma-III	GAXK01204943	216**	GSTN, GSTC	*Tigriopus japonicus*	AAY89316	7.2e-45	40/74
Calfi-Sigma-VI	GAXK01204959	217*	GSTN, GSTC	*Apis florae*	XP_003694330	8.2e-38	35/71
Calfi-Sigma-V	GAXK01204951	218*	GSTN, GSTC	*Tigriopus japonicus*	AAY89316	5.0e-33	38/70
Calfi-Sigma-VI	GAXK01204961	218****	GSTN, GSTC	*Daphnia pulex*	EFX82672	5.1e-39	38/69
Calfi-Sigma-VII	GAXK01204946	225*	GSTN, GSTC	*Daphnia pulex*	EFX63772	2.1e-39	38/76
Calfi-Sigma-VIII	GAXK01204966	239*	GSTN, GSTC	*Daphnia pulex*	EFX82672	1.8e-62	45/75
Calfi-Sigma-XI	GAXK01204949	196*	GSTN, GSTC	*Folsomia candida*	AGZ95070	1.2e-35	35/51
Calfi-Sigma-II	GAXK01204945	223*	GSTN, GSTC	*Megachile rotundata*	XP_00370954	1.1e-37	36/56
Cytosolic – Zeta							
Calfi-Zeta-I	GAXK01204941	225**	GSTN, GSTC	*Drosophila melanogaster*	NP_731358	5.0e-84	54/82
Calfi-Zeta-II	GAXK01084871	113***	GSTC	*Bactrocera dorsalis*	AFI99067	6.1e-27	22/32
Mitochondrial – Kappa							
Calfi-Kappa-I	GAXK01046934	256*	THX	*Paracyclopina nana*	ADV59554	1.2e-65	36/61
Microsomal – mGST-1							
Calfi-mGST-1-I	GAXK01178264	93***	MAPEG	*Drosophila persimilis*	XP_002023020	4.0e-42	36/61
Calfi-mGST-1-II	GAXK01081966	93***	MAPEG	*Ceratitis capitata*	XP_00453691	3.2e-26	31/52
Microsomal – mGST-3							
Calfi-mGST-3-I	GAXK01204963	156**	MAPEG	*Acartia pacifica*	AGN29624	3.2e-60	65/84
Calfi-mGST-3-II	GAXK01204967	156**	MAPEG	*Daphnia pulex*	EFX85348	7.1e-37	46/77
Calfi-mGST-3-III	GAXK01204955	264**	MAPEG	*Daphnia pulex*	EFX85347	2.2e-34	42/73
Calfi-mGST-3-IV	GAXK01204962	145****	MAPEG	*Acartia pacifica*	AGN29624	5.1e-49	61/75

†Length in amino acids.
* Predicted full-length protein flanked by "stop" codons at both N- and C-terminals.
** Putative full-length protein flanked by a "methionine" at the N-terminal, and a "stop" codon at the C-terminal. Identification of full-length is based on presence of expected structural domains and similarity to full-length proteins.
*** Partial protein with either the N-terminal "methionine" or C-terminal "stop" codon missing.
**** Protein originally identified as full-length, but prediction corrected to partial after alignment with their most similar transcript in the Norwegian Sea transcriptome (see text).
**** Protein originally identified as full-length, but prediction corrected to partial after alignment with their most similar transcript in the Norwegian Sea transcriptome (see text)
Abbreviations: GSTN, GST amino (N)-terminal domain; GSTC, GST carboxyl (C)-terminal domain; THX, thioredoxin-like domain; MAPEG, membrane-associated proteins in eicosanoid and glutathione metabolism domain.

Table 4.3. Number of genes in different classes and subclasses of the glutathione S-transferase superfamily in the crustaceans *Calanus finmarchicus*, *Tigriopus japonicus* (Lee et la., 2008) and *Daphnia pulex* (Colbourne et al., 2011) and the insect *Drosophila melanogaster* (Shi et al., 2012).

Class	Subclass	*C. finmarchicus**	*T. japonicus**	*D. pulex***	*D. melanogaster***
Cytosolic	Delta	11	2	4	11
	Theta	1	1	1	4
	Epsilon	0	0	0	14
	Mu	5	3	6	0
	Omega	3	1	1	4
	Sigma	10	1	10	1
	Zeta	2	1	3	2
Mitochondrial	Kappa	1	1	2	0
Microsomal	mGST-1	2	1	2	4
	mGST-3	4	1	2	0
Total GSTs		39†	12	31	40

* Proteins deduced from transcriptome data.
** Proteins deduced from genomic data.
† Predicted from the Gulf of Maine transcriptome only.

Table 4.4. Comparison between putative *Calanus finmarchicus* glutathione S-transferases (GSTs)* identified via transcriptome mining of two *de novo* assemblies representing populations from the Gulf of Maine (Lenz et al., 2014) and the Norwegian Sea (Tarrant et al., 2014).

Class	Subclass	Protein name	Gulf of Maine transcriptome		Norwegian Sea transcriptome		% amino acid identity between proteins
			Transcript accession No.	Deduced protein type	Transcript accession No.	Deduced protein type	
Cytosolic	Delta	Calfi-Delta-I	GAXK01204953	F	GBFB01125513	F	99
		Calfi-Delta-II	GAXK01204968	F	GBFB01087404	F	93
		Calfi-Delta-III	GAXK01204940	F	GBFB01113062	F	98
		Calfi-Delta-V	GAXK01204954	F	GBFB01106634	F	99
		Calfi-Delta-VI	GAXK01204957	F	GBFB01102119	F	93
		Calfi-Delta-VII	GAXK01204947	F	GBFB01126821	F	99
		Calfi-Delta-VIII	GAXK01073468	P	GBFB01085692	F	99
		Calfi-Delta-IX	GAXK01204950	P	GBFB01111155	F	98
		Calfi-Delta-X	GAXK01035521	F	GBFB01031301	P	93
	Theta	Calfi-Theta-I	GAXK01096295	F	GBFB01082889	F	99
	Mu	Calfi-Mu-I	GAXK01204944	F	GBFB01130857	F	100
		Calfi-Mu-II	GAXK01204956	F	GBFB01104663	F	93
		Calfi-Mu-III	GAXK01204948	F	GBFB01069639	F	97
		Calfi-Mu-IV	GAXK01204952	F	GBFB01171394	F	98
		Calfi-Mu-V	GAXK01204958	F	GBFB01086262	P	200
	Omega	Calfi-Omega-I	GAXK01204960	F	GBFB01112247	F	94
		Calfi-Omega-II	GAXK01164502	P	GBFB01122297	P	98
		Calfi-Omega-III	GAXK01016325	P	GBFB01061154	P	99
	Sigma	Calfi-Sigma-I	GAXK01204964	F	GBFB01117919	F	92
		Calfi-Sigma-II	GAXK01204943	F	GBFB01053851	F	99
		Calfi-Sigma-III	GAXK01204943	F	GBFB01057086	F	99
		Calfi-Sigma-IV	GAXK01204959	F	GBFB01080562	F	97
		Calfi-Sigma-V	GAXK01204951	F	GBFB01211675	P	99
		Calfi-Sigma-VI	GAXK01204961	P†	GBFB01107763	F	96
		Calfi-Sigma-VII	GAXK01204946	F	GBFB01147064	F	100
		Calfi-Sigma-VIII	GAXK01204966	F	GBFB01103677	P	98
		Calfi-Sigma-IX	GAXK01204949	F	GBFB01105091	F	99
		Calfi-Sigma-X	GAXK01204945	F	GBFB01125239	F	97
	Zeta	Calfi-Zeta-I	GAXK01204941	F	GBFB01157033	P	100
		Calfi-Zeta-II	GAXK01084871	P	GBFB01237836	P	100
Mitochondrial	Kappa	Calfi-Kappa-I	GAXK01046934	F	GBFB01121887	F	99
Microsomal	mGST-1	Calfi-mGST-1-I	GAXK01178264	F	GBFB01067142	F	100
	mGST-3	Calfi-mGST-3-I	GAXK01204963	F	GBFB01089387	F	99
		Calfi-mGST-3-II	GAXK01204967	F	GBFB01094405	F	100
		Calfi-mGST-3-III	GAXK01204955	F	GBFB01076955	F	100
		Calfi-mGST-3-IV	GAXK01204962	P†	GBFB01082093	F	99

*GST transcripts listed showed >90% amino acid identity between proteins.
†In Table 4.2 these proteins were identified as putative full-length because flanked by a "methionine' at the N-terminal, and a "stop" codon at the C-terminal, and by the conservation of structural domains. However, alignment of each protein with its counterpart from the "Norwegian Sea" transcriptome suggests that they are partial proteins.

Table 4.5. Putative *Calanus finmarchicus* glutathione S-transferases (GSTs) showing differences between the Gulf of Maine (Lenz et al., 2014 and the Norwegian Sea (Tarrant et al., 2014)*de novo* assemblies.

Type of variation	Protein name	Gulf of Maine transcriptome		Norwegian Sea transcriptome		% amino acid identity between proteins
		Transcript accession No.	Deduced protein type	Transcript accession No.	Deduced protein type	
Genetic variation	Calfi-Delta-IV	GAXK01204965	F	GBFB01111154	F	48
	Calfi-Delta-XI	GAXK01204939	F	GBFB01141707	F	88
Additional gene	Calfi-Omega-IV	GAXK01138968*	P	GBFB01119512	F	99

*This transcript was not identified as encoding a GST protein in the original screening of the Gulf of Maine transcriptome, but rather was detected via a query with its Norwegian Sea counterpart.

A. *Daphnia pulex* and *Calanus finmarchicus* Delta GST

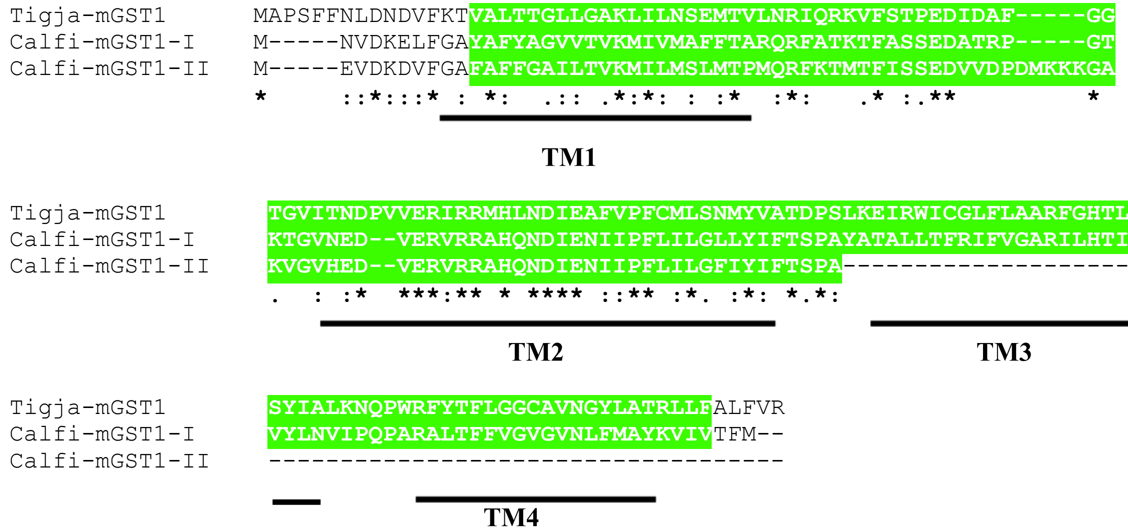
```
Dappu-Delta       MPIDLYYMDVSAPCRAVMMTAKLIGVELNMKMTNLMAGEHMKPEFLAINPQHNVPTVDDK
Calfi-Delta-IV    MAIDLYYMDLSGPCRSVMMTAKMVGVELNLKVLNLMAGEQLTPEFIQINPQHNIPTIVDD
                  *.*******:*.***:******::*****:*: ******::.***: ******:**: *.

Dappu-Delta       GFYLNESRAICAYLINQYAKDDTLYPKEPKVRALVDQRLYFDMGVLYQKFGLVYYPVIMG
Calfi-Delta-IV    GFSLNESRAICGYLVQKYAKDDTLYPKEAQQRALVDQRLYFDMGVLYQRFSKLYYPVMFH
                  ** ********.**:::***********.: *****************:*. :****::

Dappu-Delta       GATKLSESAIKDLAGAIDFLETFLSKTTYAAGDHLTIADVALVASVSTIEAVDKTHLDNH
Calfi-Delta-IV    GATKMEESAKKELDGAIDFLNTFLTANKYAAGDHLTIADICLVATVTTIEGADAQIFDKF
                  ****:.*** *:* ******:***: ..***********:.***:*:***..*   :*:.

Dappu-Delta       PKIQKWLETCKAEIPDYQELNQKGADMFAKWAQGALAKLESA
Calfi-Delta-IV    PKITSWLETCKSELVGYNEANQVGATEFGQMVQAALAKLE--
                  *** .******:*: .*:* ** **  *.: .*.******
```

B. *Tigriopus japonicus* and *Calanus finmarchicus* Mu GST

```
Tigja-Mu          MSNLPVLCYWDIRGLAQPIRLLLEYTGTKFEDKQLVCGPGPNYDKSVWTNEKHKLGLDFP
Calfi-Mu-III      MA--PTLCYWNSRGLGQPIRLLLEFTGTEFEDRKLTMEGAPTYDKSCWTDIKDKLGLDFP
                  *:  *.****: ***.********:***:***::*.   .*.**** **: *.*******

Tigja-Mu          NLPYFVDGDRKITQSNAILRYIARKHNLLGQTEEEQMRVDIMAEQSMDFRNGLVRLCYNQ
Calfi-Mu-III      NLPYFLDGNIKITQSNAILRHIARKNGLCGESVEEKARVDMMADCAMDLRNTIVRLVYNP
                  *****:**: **********:****:.* *:: **: ***:**: :**:** :*** **

Tigja-Mu          SFDQVKDDYLAALVSKLDEFAKFLGDRPWFAGESLTFVDFIMYELLDQHRVLAPEVVNQS
Calfi-MuIII       DYDNLLPPYIERLPSILEPFSKFLAANPWFVGQKVTFVDFPMYELLDQLRNMIPGCLKNY
                  .:*::   *:  * * *: *:***. .***.*:.:***** ******* * : *  :::

Tigja-Mu          PKLVDYLNRFEKLPPIEAYMKSDRFMKRPLNNRMAKFGAEK
Calfi-Mu-III      PNLIKFLDRFEKLPKIAAYMKSSRYMAAPINNKHAKFGG--
                  *:*:.:*:****** * *****.*:*  *:**: ****.
```

Figure 4.1. **Alignment of selected *Calanus finmarchicus* Delta GST (Calfi-Delta-IV) and Mu GST (Calfi-Mu-III) proteins with their top arthropod protein BLAST hits.** (A) Alignment of *D. pulex* Delta GST (Dappu-Delta) (**Accession No. EFX81633**; 222 amino acids long) and Calfi-Delta-IV (220 amino acids long). (B) Alignment of the *T. japonicus* Mu GST (Tigja-Mu; **Accession No. ACE81254**; 221 amino acids long) and Calfi-Mu-III (222 amino acids long). In each panel, ‘‘*’’ located beneath the alignment indicates residues that are identical in the two sequences, while ‘‘:’’ and ‘‘.’’ indicate conservatively substituted (similar) amino acids shared between the protein pairs. Amino acids highlighted in black are the ones predicted by SMART analysis to form the conserved amino (N)-terminal domain (GSTN), amino acids highlighted in red represent the conserved carboxyl (C)-terminal domain (GSTC).

A. *Tigriopus japonicus* and *Calanus finmarchicus* microsomal GST-1



B. Crustacean microsomal GST-1

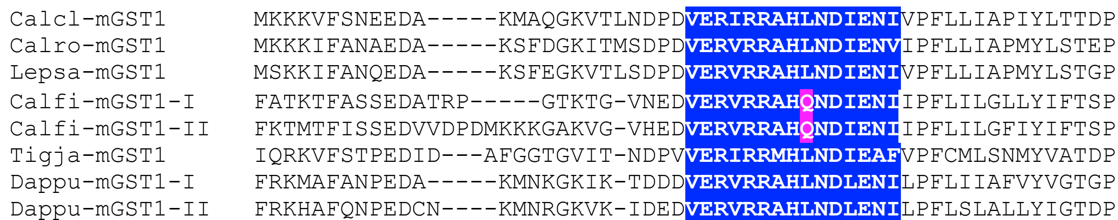


Figure 4.2. **Alignment of *Calanus finmarchicus* microsomal glutathione S-transferase subclass 1 proteins with mGST-1s from other crustaceans**. (A) Alignment of *C. finmarchicus* putative microsomal GST-1 proteins (Calfi-mGST-1-I and Calfi-mGST-1-II) with the *T. japonicus* query used in their discovery (Tigja-mGST-1; Accession No. ACE81248). Highlighted in green are amino acids in the conserved MAPEG structural domain identified using SMART software. The abbreviation "TM" indicates predicted transmembrane regions in the *C. finmarchicus* mGST-1 proteins. The ''*'' located beneath each alignment indicates residues that are identical in the two sequences, while '':'' and ''.'' indicate conservatively substituted (similar) amino acids shared between the protein pairs. (B) Multiple alignments of *C. finmarchicus* microsomal GST-1 proteins (Calfi-mGST-1-I and Calfi-mGST-1-II) with publicly available mGST-1s from the crustaceans *C. clemensi* (Calcl), *C. rogercresseyi* (Calro), *L. salmonis* (Lepsa), *T. japonicus* (Tigja) and *D. pulex* (Dappu). The conserved motif consisting of 16 amino acids (VERVRR*X*HLND*X*ENI*X*, where the three *X*s represent variable residues) is highlighted in blue. The non-conservative substitution found only in *C. finmarchicus* is highlighted in pink.

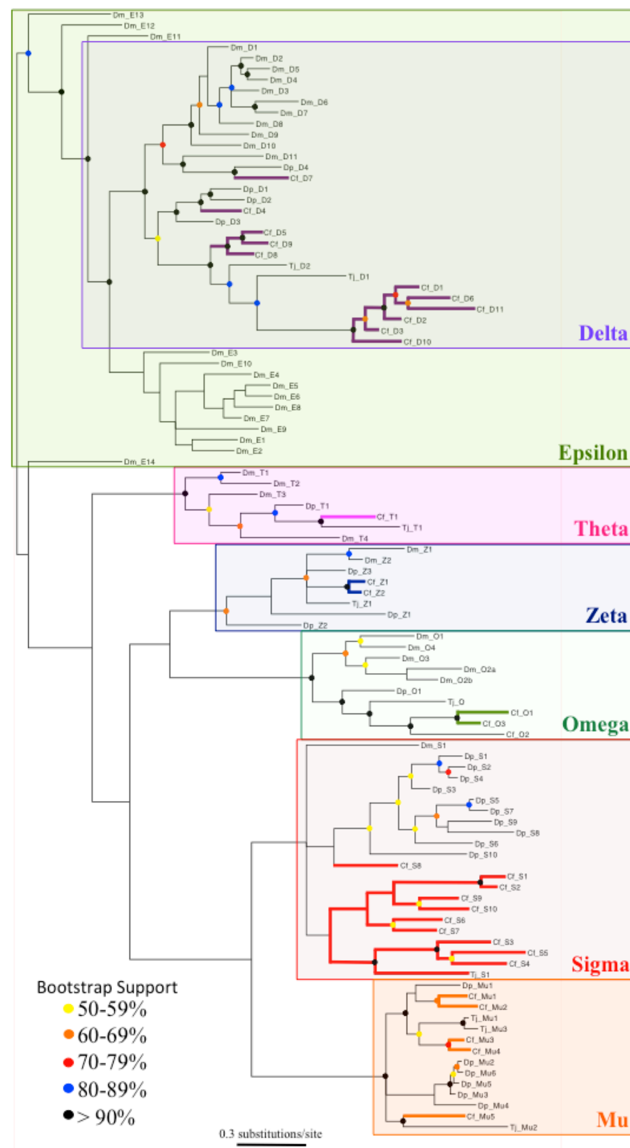


Figure 4.3. Phylogenetic tree for cytosolic GSTs from the crustacean *Calanus finmarchicus* and other selected crustacean and insect species. The consensus Bayesian likelihood tree shows the relationships between cytosolic GSTs from *C. finmarchicus* (Cf, in color) and those from the insect *D. melanogaster* (Dm), the copepod *T. japonicus* (Tj), and the cladoceran *D. pulex* (Dp). The tree was built using an analysis of 10,000,000 generations in MrBayes, excluding the initial 2,500,000 generations as burn-in. Bootstrap values were calculated using RAxML with 1,000 interactions. For 73 branches, Bayesian posterior probabilities were grater than P>0.5, 68% of those with P between 0.9 and 1 (data not shown). 73 branches had bootstrap values greater than 50% (color-coded circles).

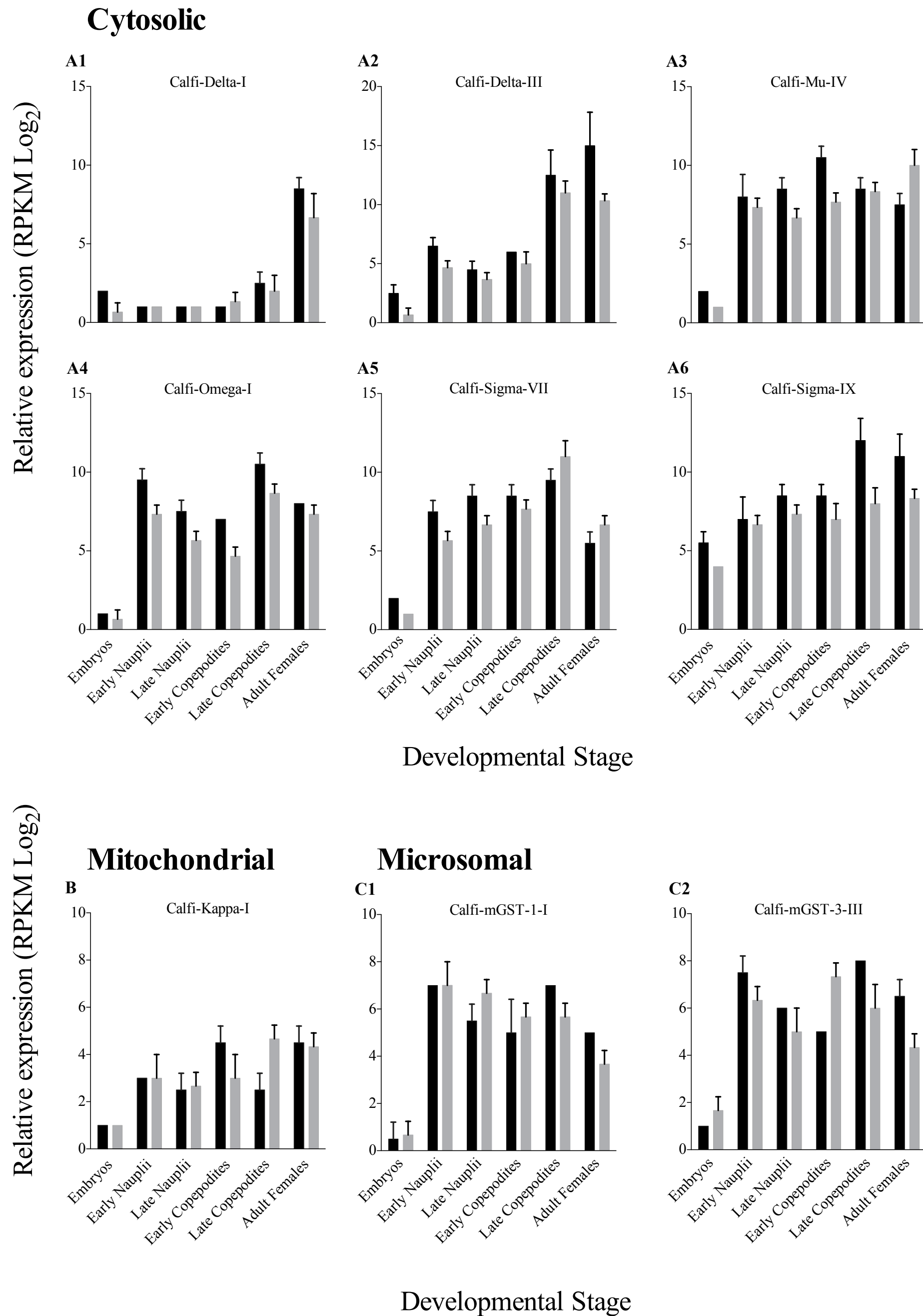


Figure 4.4. Relative expression of selected *Calanus finmarchicus* cytosolic, mitochondrial and microsomal GST-encoding transcripts across six developmental stages. Relative expression measured in 2011 (black bars) and 2012 (grey bars) for nine GSTs are shown for embryos, early nauplii (NI-II), late nauplii (NV-VI), early copepodites (CI-II), late copepodites (CV), and adult females as RPKM (reads per kilobase per million mapped reads) in Log_2. (A) Cytosolic GSTs belonging to the Delta (A1-A2), Mu (A3), Omega (A4) and Sigma (A5-A6) subclasses. (B) Mitochondrial Kappa GST class. (C) Microsomal GST subclass 1 (C1) and subclass 3 (C2). Error bars in 2011 (black) are standard deviations of two technical replicates for each stage, while in 2012 (gray) error bars are standard deviations of three biological replicates.

Chapter 5

Relative expression of Glutathione S-transferase genes in *Calanus finmarchicus* females feeding on the toxic dinoflagellate, *Alexandrium fundyense*, measured by RNA-Seq and quantitative reverse transcription PCR (RT-qPCR)

ABSTRACT

The effect of the toxic dinoflagellate, *Alexandrium fundyense*, on gene expression in the copepod *Calanus finmarchicus* was examined in laboratory experiments. Adult females were fed for 5-days on one of three experimental diets: control (100% *Rhodomonas* spp.), low dose of *A. fundyense* (25% by volume, 75% *Rhodomonas* spp.), and high dose (100% *A. fundyense*). Relative gene expression of three target glutathione S-transferase (GST) genes was measured using RT-qPCR on days 0.5, 1, 2 and 5. In addition, on days 2 and 5 results obtained by RT-qPCR were compared to relative expression determined by RNA-Seq. The level of sensitivity of detection was high and the results from the two independent techniques were in good agreement. The RT-qPCR data confirmed the differential expression for a Delta and a Sigma GST on days 0.5 to 2, but not on day 5. The third target gene, a microsomal GST, was not differentially expressed in either treatment or day. These results indicate a rapid response at the gene expression level after the females are introduced to the experimental diet for the two cytosolic GSTs, and that this response is transient. Relative gene expression patterns in females from the June and July experiments gave very similar results, in spite of differences in overall reproductive rates.

1. INTRODUCTION

In the last decade, high-throughput technology has been used to study the cellular stress response in many organisms including non-model species (Kültz, 2003, 2005 references therein; Somero, 2012). RNA-Seq gene expression profiles, obtained by sequencing millions of short sequence reads and by mapping these reads to a *de novo* reference transcriptome, have been shown to be highly reproducible with relatively little technical variation (Fang and Cui, 2011 and references therein). Nevertheless, quantitative reverse transcription PCR (RT-qPCR) has been considered the "gold standard" for investigating differences in gene expression (Gibson et al., 1996; Foley et al., 1993). Because RNA-Seq is a new technology, it is common practice to use RT-qPCR to validate results from high-throughput transcriptome studies in particular when working with non-model species (Fang and Cui, 2011; Schoville at al., 2012).

RT-qPCR has been one of the most widely used methods of gene expression as it has a large dynamic range with high sensitivity, and is applicable across a wide range of experimental conditions (Costa et al., 2013). While high-throughput sequencing provides relative gene expression on tens of thousands of genes simultaneously, its cost usually limits RNA-Seq data to relatively few treatments, time points and replicates (Chapter 3, Costa et al., 2013). RT-qPCR studies are focused on target genes and include measurements of relative expression for multiple treatments, time points and replicates. For example, in the euryhaline green crab, *Carcinus maenas*, changes in mRNA expression of carbonic anhydrase start to occur after 24 hours from the transfer into low salinity medium, remained elevated through 4 days but decreased by day 7 (Henry et al., 2006). In contrast, carbonic anhydrase protein concentration levels increase only after 48 hours from the transfer and continued to increase through 7 days (Henry et al., 2006). However, the RT-qPCR-based approach requires prior sequence data for the genes of interest, which makes this technology only applicable for targeting genes with known sequences. For

these reasons, the complementary application of both methods represents a powerful way for quantifying transcript numbers. In this study, I used RT-qPCR to measure expression rate of three target genes in the glutathione S-transferase family, at 2 additional time points, in *C. finmarchicus* feeding on *A. fundyense* diets during June and July experiment.

Glutathione S-transferases, which are involved in phase II of the detoxification process, have been extensively used as biomarkers of detoxification in other copepods, such as *Tigriopus japonicus*, *Calanus helgolandicus* and *Calanus glacialis* exposed to stressors including diatoms, heavy metals and oils droplets (Lee et al., 2008; Hansen et al., 2009, 2010, 2011; Lauritano et al., 2012). GSTs were likely candidates for being involved in the detoxification of saxitoxins produced by *A. fundyense*.

The recent identification and characterization of 41 putative GSTs in *C. finmarchicus* (Roncalli et al., 2015) provided an opportunity to investigate relative expression of several genes from this family in more detail. Gene duplication of GSTs is well known among the arthropods, and members from different classes and subclasses differ in their specific role in the regulation of cellular processes. The GST family of proteins in *C. finmarchicus* is particularly diverse, and as hypothesized in Roncalli et al. (2015), GSTs genes belonging to different classes or even different subclasses are likely to differ in their specific function as suggested by their protein structure, phylogenetic placement and relative expression across developmental stages. In this study, the three target GSTs included members of the microsomal and cytosolic class The microsomal GST (mGST3) was selected because it was previously used as a biomarker in *C. finmarchicus* (Hansen et al., 2009, 2010, 2011); one Delta and one Sigma GST (cytosolic class) were chosen because they are representatives of the two most diversified subclasses in *C. finmarchicus* (containing 11 and 10 GSTs respectively; Roncalli et al., 2015). In addition, RNA-

Seq data indicated that these two cytosolic GSTs were differentially expressed at 2 days but not at 5 days (Chapter 3).

In these experiments, *Calanus finmarchicus* adult females were fed for 5-days on one of three experimental diets, control (100% *Rhodomonas* spp.), low dose of the toxic dinoflagellate *Alexandrium fundyense* (25% by volume, 75% *Rhodomonas* spp.), and high dose (100% *A. fundyense*) as described in Chapter 2. The two primary goals for this study were: 1) to compare relative gene expression obtained by RT-qPCR to RNA-Seq; and 2) to investigate the change in expression in females from the June and July experiments at additional time points using three target genes involved in detoxification. In order to validate the accuracy of the RNA-Seq results in Chapter 3, I measured the expression of three glutathione S-transferase (GST) genes with RT-qPCR and compared the results between the two platforms. As mandatory for gene expression studies using RT-qPCR, the expression level was also measured for 3 candidate reference genes (elongation factor 1α, actin, ribosomal 16S) that were not expected to be regulated by the experimental protocols and were used as internal controls (Bustin et al., 2009; Bustin et al., 2010). To address the second goal, relative expression of the target glutathione S-transferases was measured at 0.5, 1, 2 and 5 days in adult females from the June and July experiments fed one of three experimental diets as described in Chapter 2. There was good agreement between the RNA-Seq and RT-qPCR results. In addition, a time and dose- dependent response was observed for the two cytosolic Delta and Sigma GSTs while the microsomal GST did not show significant regulation under our experimental conditions.

2. MATERIALS AND METHODS

2.1 Field collection and maintenance of *Calanus finmarchicus*

C. finmarchicus were collected using a vertical net tow (75 cm diameter, 560 µm mesh) in June and July of 2012 in the Gulf of Maine near Mount Desert Rock (Lat: 44° 2'N; Long: 68°3'W). Copepods were kept overnight at 10ºC and then transferred into experimental jars as described in Chapters 2 and 3.

2.2 Experimental design

Experimental incubations of females used in the RT-qPCR experiments were conducted in parallel with the survival, grazing, reproduction and RNA-Seq experiments described in Chapters 2 and 3. *C. finmarchicus* adult females were incubated with one of the three experimental diets for a total of 5 days with eight females maintained in jars with 800 mL of seawater, which corresponded to the same density as in the experiments of Chapter 2. Six pairs of separate containers (6 total) were set up for each experimental treatment. As described in Chapters 2 and 3, copepods were fed on the non-toxic flagellate *Rhodomonas* sp. (8000 cells $mL^{-1}d^{-1}$) in the control group. In the "low dose" (LD) group copepods were fed 50 cells $mL^{-1}d^{-1}$ *A. fundyense* and 6000 cells $mL^{-1}d^{-1}$ *Rhodomonas sp.*, which correspond to a 25:75 proportion by algal volume. The "high dose" group (HD) was fed 200 cells $mL^{-1}d^{-1}$ of *A. fundyense*. During the incubation period, a new dose of the specified food suspension was added to each jar without regard to food remaining from previous days; daily and copepod survival rates were monitored. The three experimental food suspensions were comparable in terms of carbon content (see Table 2.1, Chapter 2). All experimental animals were kept in a Percival Model I-36VL Incubator System (Percival Scientific, Inc., Perry, IA, USA) at 10ºC on a 14:10 hr light dark cycle. At 0.5, 1, 2 and 5 days, a total of 4 adult females were removed from each of the first members of the jar

pairs and immediately transferred to 0.5 mL RNA later (Ambion) and stored at -80°C until RNA extraction. On days 2 and 5 this process was repeated for the remaining members of the jar pairs; thus, for each sampling time point and treatment (control, low dose and high dose) three biological replicates were collected. One experiment was completed in June, and a second one in July. In the July experiment, samples were not collected at the day 1 time point.

2.3. Gene expression using RT-qPCR

2.3.1. RNA extraction, cDNA synthesis and primer design

Total RNA was extracted from whole *C. finmarchicus* adult females using the QIAGEN RNeasy Mini Kit (QIAGEN Inc., Valencia, CA, USA), in conjunction with a Qiashredder column (QIAGEN Inc.), following the instructions of the manufacturer, with a final elution volume of 30 µl. RNA concentration and quality were checked using an Agilent Model 2100 Bioanalyzer (Agilent Technologies, Inc., Santa Clara, CA, USA).

Poly-A mRNA in 1 µg of total RNA per sample was reverse transcribed using the QuantiTect Reverse Transcription Kit. The resulting cDNAs were used as templates for conventional PCR and real-time quantitative PCR. Specific primers were designed to target the conserved domain for each gene using the software Geneious (v. 6.1). For all genes the amplicon size was lower then 170 bp. Secondary structure and primer accuracy were evaluated using IDT (http://www.idtdna.com) and *in silico* PCR (Bioinfx http://bioinfx.net/cgi-bin/pcr). Each primer was also tested with PCR, optimizing the temperature range (55-63°C) and primer concentrations (100-400 nM). Only primers that generated a single strong band on a 1.5% agarose gel were further considered for the RT-qPCR analysis. The primers used in the RT-qPCR reactions and their optimal temperatures are listed in Supplemental Table 5.1. All the analyses were conducted

according to MIQE guidelines and checklist (Bustin et al., 2009; Bustin et al., 2010).

2.3.2. Reverse-Transcription-Quantitative Real Time Polymerase Chain Reaction (RT-qPCR)

RT-qPCR experiments were performed using the cDNA previously generated (section 2.3.1) in a LightCycler 96 System (Roche) thermal cycler with a PCR volume for each sample of 25 μl, with 12.5 of Fast Start SYBR Green Master Mix (Roche), 2 μl of cDNA template and 1 μl for each oligo (final concentration 400 nM). The RT-qPCR thermal profile was: pre-incubation at 95°C (10 min), followed by 50 cycles of 95°C (30 s), 63°C (1 min) and 72°C (30 s). Melting curve analyses were used to check for gene-specific amplification and the absence of primer dimers by determining the presence of a single peak in the melting-curve. The optimal quantity of template was assessed using serial dilutions, ranging from 1:1 to 1:10000. In addition to the three biological replicates for each treatment and time point, each RT-qPCR reaction was carried out in triplicate to capture intra-assay variability. Each assay included three no-template controls (NTC) for each primer pair. Efficiencies were calculated for each gene using the equation [$E=10^{-1/slope}$]. Two different algorithms were utilized to identify the best reference gene in our experimental design: BestKeeper (Pfaffl et al., 2004) and NormFinder (Andersen et al., 2004). Transcriptional expression of each gene was measured using the REST tool (Relative expression software tool; Pfaffl et al., 2002). For each biological replicate the three technical replicates were averaged for all Cq measurements, and standard deviations calculated for the 3 technical replicates was < 1. For one set of technical replicates, the SD exceeded 1, due to the fact that one of the replicates represented an outlier. The outlier was removed and the remaining two technical replicates were averaged (SD=1). Relative expression was determined for each biological replicate and each target gene using the Relative Expression Software Tool (Pfaffl et al., 2002),

which calculates relative expression as the expression ratio (fold change) between Cq values of a target gene versus Cq values for the reference genes (Pfaffl et al., 2002). To assess the expression rate for the tested genes, I firstly determined the best reference gene from the three genes tested (EFA, 16S, actin), and then normalized expression as described in Pfaffl et al. (2002). Expression rate was quantified as Log_2.

2.4. Gene expression using RNA-Seq technology

As described in Chapter 3, total RNA was obtained from *C. finmarchicus* females on days 2 and 5 of the experiment. A total of 18 RNA-Seq libraries (3 replicates x 3 treatments x 2 time points) were sequenced, and then separately mapped to the *C. finmarchicus* reference transcriptome (96,090 contigs; Lenz et al., 2014) using Bowtie software (version, 2.0.6) (Langmead et al., 2009). Differential gene expression analysis was performed using the BioConductor package edgeR (Robinson et al., 2010). Prior to the statistical test, expression was normalized by the length of the reference transcript and the number of mapped reads in each sample (RPKM: reads per kilobase per million of mapped reads) with a custom Perl script. At days 2 and 5, the number of differentially expressed genes was measured by pairwise comparison of libraries to test: CONTROL vs LD and CONTROL vs HD. Transcripts were identified as differentially expressed using the exact test ($p<0.05$) and a multiple comparison correction with Benjamini-Hochberg method (false discovery rate <5%) implemented by edgeR (Robinson et al., 2010). Expression rate was quantified in units of Log_2 fold (experimental/control) where a value of 0 represents equal expression between experimental condition and control.

2.5. Comparison between RNA-Seq and RT-qPCR expression levels

Comparison between the RNA-Seq and RT-qPCR platforms was done for all six genes represented by three glutathione S-transferases (GSTs) and three candidate reference (EFA, actin, 16S) genes at days 2 and 5. RT-qPCR results were compared with the RNA-Seq data in two ways. First, differential gene expression between the treatments and the control were computed as a ratio for the six genes using normalized RNA-Seq and the normalized RT-qPCR expression data, and testing statistically using a linear regression. Second, the normalized counts (RPKM) from RNA-Seq for each gene-treatment-day combination were compared to the normalized Cq-values (RT-qPCR) using linear regression.

3. RESULTS AND DISCUSSION

3.1. RT-qPCR efficiencies and analysis of candidate reference genes

Amplification of the three GSTs and the candidate reference genes was successful with reaction efficiencies ranging between 1.98 and 2.09 (100% and higher). Changes in expression for the target GSTs were normalized using elongation factor 1α (EFA). This gene was found to be the best reference candidate, because it had the lowest standard deviation (BestKeeper) and the highest stability value (NormFinder) compared with the other two genes (Table 5.1). The expression of elongation factor 1α was normalized using the second best candidate, which according to both metrics was actin (Table 5.1), following MIQE guidelines (Bustin et al., 2009; Bustin et al., 2010).

3.2. Comparison between RT-qPCR and RNA-Seq

RNA-Seq and RT-qPCR results were compared for days 2 and 5 for the July experiment

using all six genes (3 GSTs and 3 candidate reference genes). The results of the two approaches to measuring relative gene expression were in good agreement by comparing the fold-change (Log_2) measured as a statistical difference of normalized counts in RNA-Seq with relative expression obtained by RT-qPCR as shown in Figure 5.1A (R^2=0.896, p<0.0001). The relative expression patterns of the three GSTs confirmed that Delta and Sigma were differentially expressed (down-and up- respectively) in females feeding on *A. fundyense* diets on day 2 but not on day 5 (Figure 5.1A). Fold change expression was highly comparable, ranging between 2.6 to 3.4 fold (Log_2) compared to females feeding on the control diet (Figure 5.1A). The third GST (mGST3) was not differentially expressed on either day, or experimental treatment. In the second approach, expression estimates from RNA-Seq data as normalized counts (RPKM) were compared with expression estimates from RT-qPCR (Cq-values). A least-squares fit through the data was highly significant (R^2=0.879, P<0.0001; Figure 5.1B). Most of the genes showed modest expression levels (5-10 RPKM) (Figure 5.1B), indicating that the RNA-Seq data are robust and supports the conclusion that this technology provides high accuracy in the identification of differentially expressed genes. These results are consistent with other studies that have found excellent agreement between RT-qPCR and RNA-Seq (Gallardo-Escárate et al., 2014: Schoville et al., 2012; Fang and Cui, 2011 and references therein).

The close agreement between RNA-Seq and RT-qPCR results solidifies the conclusions based on the former technique alone. Analysis of the RNA-Seq data at 2- and 5-days resulted in the identification of 16 differentially expressed enzymes (< 1% of the total DEs) involved in detoxification phases I and II (Table 3.2, Chapter 3). This was surprising since a more global detoxification response had been expected in response to the dinoflagellate diets, in particular the HD treatment. The results from the RT-qPCR strengthen the conclusion that the detoxification

response was limited to the differential gene expression of four glutathione S-transferases, four cytochrome P450, two sulfotransferases, two thioredoxins, two multi resistance proteins and a single glutathione peroxidase (Chapter 3).

3.3. Changes in expression of Glutathione S-transferases over time using RT-qPCR

The effect of the dinoflagellate on the relative expression of two glutathione S-transferases was time- and dose- dependent (Figure 5.2). The two cytosolic GSTs (Delta and Sigma; Figures 5.2A, B) were significantly regulated over time while the microsomal GST did not show significant changes compared with the control group (Figure 5.2C). Delta GST was significantly down-regulated already at 12 hr from the introduction of the toxic dinoflagellate into the diet with a similar effect observed in both LD and HD treatments (Figure 5.2A). In the HD treatment significant down-regulation was observed until day 2 in both June and July, but not at day 5. Relative expression in females on the LD treatment was only significantly different at 12 hours, but not at any of the other time points (Figure 5.2A). Thus, the RNA-Seq study that was limited to expression data for day 2 and day 5 missed the very transient regulation of this GST in the LD treatment. Furthermore, these data suggest that down-regulation of the mRNA of this detoxification enzyme occurred very soon after the introduction of the toxic diet, and this response was graded with the magnitude of the response being higher and longer in the HD than the LD treatment. Relative expression of Sigma GST showed an interesting pattern. In the HD treatment, relative expression of this GST increased between the 12 hr and 2 day time points with significant up-regulation occurring only at 2 days (Figure 5.2B). In contrast, significant up-regulation of the Sigma GST in the LD treatment occurred at 12 hrs and 2 days in both June and July experiments (Figure 5.2B). Interestingly, at 1 day, relative expression in June in the LD was

not significantly different, and similarly for the HD. Unfortunately, there was not an equivalent time point during the July experiment. The differential expression of the Delta and Sigma GSTs was modest and ranged between 2.2 to 3.7 fold change compared with females feeding on the *Rhodomonas* sp. diet (Figure 5.2A and B).

Rapid activation of GST Delta and Sigma has been reported in the intertidal copepod *T. japonicus* exposed to trace metals for a total of 4 days (Lee et al., 2008). Both genes were up-regulated after 12 hr and increased their relative expression over time with the highest peak at day 1 and 2, respectively. In our study, differential expression of these GSTs was observed at 12 hr and persisted to 2 days. However, there was no significant increase in the magnitude of the response, and by day 5 expression levels were similar to the control. These results are consistent with RNA-Seq results (Chapter 3): after the immediate response corresponding to a cellular stress response, the copepod adjusts to a new steady-state by day 5 in which detoxification does not appear to have a prominent role. The differences in up- and down- regulation of the GSTs shown here and in Chapter 3 are consistent with the hypothesis that the multiplicity of GST genes have separate functions and can be expected to be activated differentially to any particular stressor.

Agreement in relative expression of the six genes targeted for RT-qPCR between the June and July experiments is shown in Figure 5.3, which is a graph of the fold-change between the experimental treatments and control measured in June and July for days 0.5, 2 and 5. The fold-change measured in the two experiments was very similar as indicated by the correlation coefficient (R^2=0.750, $P<0.0001$). For these genes relative expression was very similar in the two experiments, in spite of the differences in reproduction that were observed between the June and July females in all three treatments (Chapter 2). This consistent response (up-regulation of

Sigma GST, and down-regulation of Delta GST) adds confidence in the use of relative gene expression as a marker for physiology.

4. CONCLUSIONS

Good agreement was found between RT-qPCR and RNA-Seq data validating the application of RNA-Seq technology to measuring differential gene expression across thousands of genes simultaneously. The level of sensitivity of detection was high in both platforms as demonstrated by the modest expression levels measured for the target GST genes. In considering these results and the ones from Chapter 3, it is clear that females show an initial and transient response that includes the up-regulation of a Sigma GST and down-regulation of a Delta GST, that is absent after 5 days on the experimental diets. Changes in relative expression were already present at the first time point (0.5 day), underscoring that there was an immediate response to the introduction of the dinoflagellate diet. Relative gene expression patterns in females from the June and July experiments gave very similar results, in spite of differences in overall reproductive rates. Although detoxification is likely to be part of the copepod's response to the dinoflagellate diet, it does not appear to be the major response, given the modest changes in gene expression as measured by the number of regulated GSTs (4 out of 39 in the reference transcriptome, see Chapter 4) and the magnitude of the differential expression (< 4-fold difference). These conclusions are supported by both RNA-Seq and RT-qPCR results.

Table 5.1. Candidate reference genes (EFA, 16S, ACTIN) in *C. finmarchicus.* For each reference gene, number of samples (n) mean, minimum (min) and maximum (max), standard deviation (SD) and Stability value are listed for both June (A) and July (B) experiments.

A. June

Best Keeper	EFA	16S	ACTIN
N	18	18	18
Mean	18.94	17.32	25.14
Min	16	14.4	22
Max	23	19.4	28
SD	**0.96**	**0.99**	**0.98**
Norm Finder			
Stability value	0.71	0.47	0.64

B. July

Best Keeper	EFA	16S	ACTIN
N	18	18	18
Mean	22.98	17.59	28.04
Min	21.09	16	25
Max	24.3	20	29.8
SD	**0.66**	**0.80**	**0.76**
Norm Finder			
Stability value	0.77	0.46	0.60

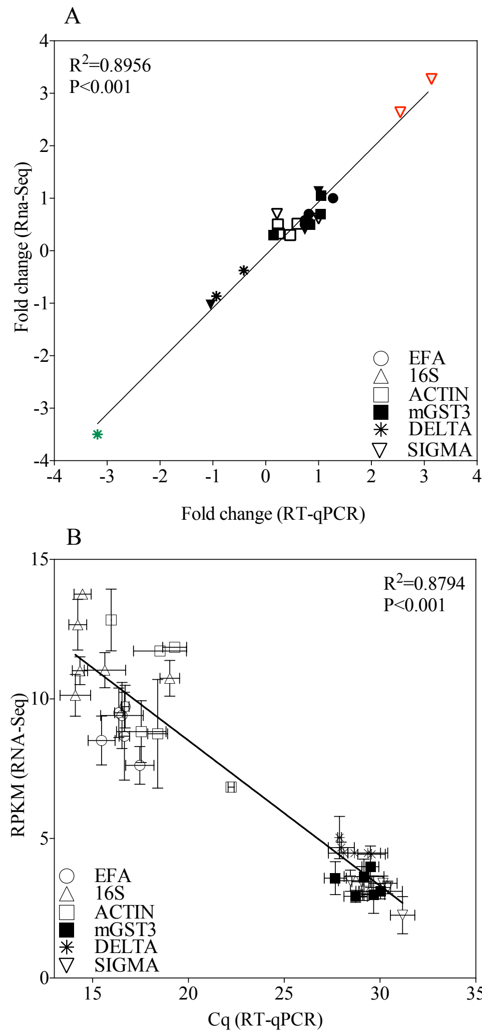


Figure 5.1. Comparison between relative expression measured with RNA-Seq and RT-qPCR. Relative expression measured with RNA-Seq and RT-qPCR for 6 genes including 3 glutathione S-transferase (mGST3, Delta, Sigma) and 3 candidate reference genes (EFA, 16S, Actin) in *C. finmarchicus* adult females feeding on control *Rhodomonas* sp. and two experimental diets (LD and HD of *A. fundyense*) for 2 and 5 days (n=36). In A) expression was measured as Log_2 fold change difference between the experimental and control treatments (n=24) with genes significantly regulated showing fold-change >1. Delta (green) and Sigma (red) were the only genes showing significant differential expression. Delta was down-regulated (green) exclusively in the HD and Sigma up-regulated (red) in LD and HD treatments. Expression for EFA has been normalized with actin. The line is a least-squares linear fit to the data (R^2=0.8956, P<0.0001). In B), relative expression as reads per kilobase per million mapped reads (RPKM) (Log_2) obtained from RNA-Seq were compared with normalized Cq-values from RT-qPCR (n=36). The line is a least-squares linear fit to the data (R^2=0.8794, P<0.0001).

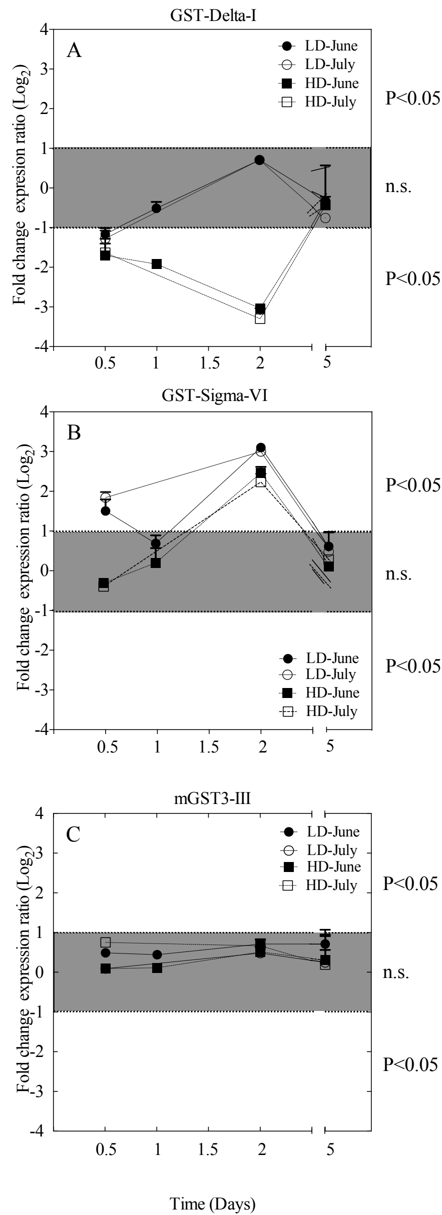


Figure 5.2. Glutathione S-transferase relative expression over time in *C. finmarchicus* feeding on *A. fundyense* diets for 5 days. Relative expression of three GSTs members of the cytosolic Delta-I (A) and Sigma-VI (B) subclasses and the microsomal mGST-3-III (C) was measured at each time point for *C. finmarchicus* adult females feeding on LD and HD of *A. fundyense* in June and July experiments using RT-qPCR. GST expression (Log_2 fold change) was normalized using the elongation factor 1α. The grey area delimits expression level not significantly different from the control diet (Log_2 Fold change <=1**)**. Solid symbols: June experiment; open symbols: July experiment; circles: LD treatment; squares: HD treatment. Day 1 data for July experiment are not available.

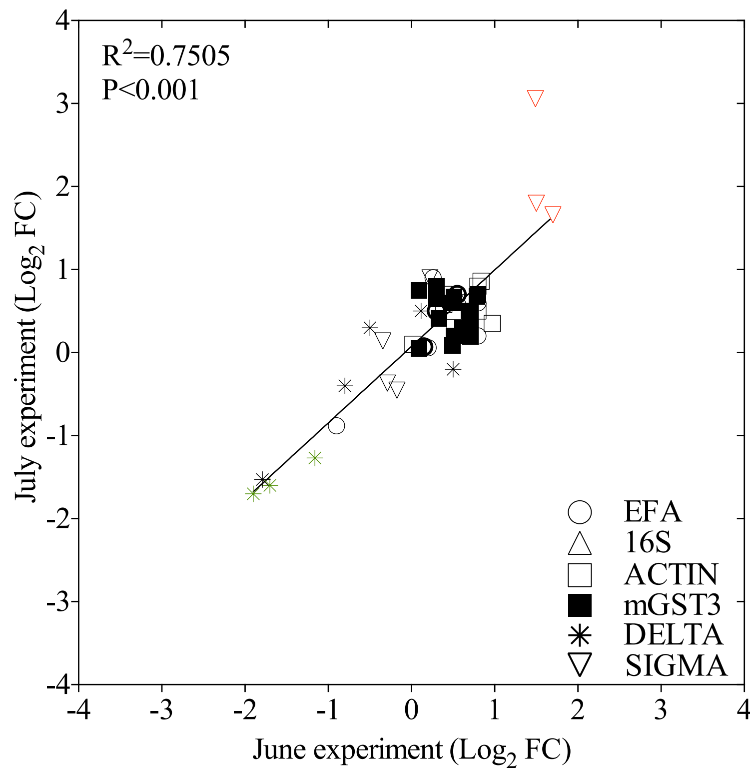


Figure 5.3. Comparison of relative expression levels measured with RT-qPCR for June and July experiments. The line is a least-squares linear fit to the data (R^2=0.7505, P<0.0001). The graph includes expression level (Log_2 fold change) for 6 genes including 3 glutathione S-transferase (mGST3, Delta, Sigma) and 3 candidate reference genes (EFA, 16S, actin) measured in *C. finmarchicus* adult females feeding on LD and HD of *A. fundyense* for 0.5, 1, 2 and 5 days (n=36). Expression levels were calculated at the average of 3 biological replicates for each gene 4 time point 3 treatment combination. Genes with Log_2 fold change higher than 1 are considered significantly expressed. In both June and July, Delta (green) and Sigma (red) were the only genes with significant differential expression compared with the control. Delta was down-regulated (green) exclusively in the HD and Sigma up-regulated (red) in LD and HD treatments. Expression for EFA has been normalized with actin.

Chapter 6

Summary and Conclusions

The predicted increase in harmful algal blooms (HAB) is expected to lead to negative impacts on the economies of coastal communities due to the resulting degradation of ecosystems, declines in marine fisheries, and decline in tourism and recreational activities (Etheridge, 2010; Anderson et al., 2012, 2014; Rossini, 2014). Every summer, HABs of the dinoflagellate *Alexandrium fundyense* occur in the Gulf of Maine (Anderson, 1997; Deeds et al., 2014), causing local fishery closures due to outbreaks of paralytic shellfish poisoning (Kleindinst et al., 2014). During the period in which HABs are most common, the zooplankton population of the Gulf of Maine is dominated by *Calanus finmarchicus* (Davis, 1987; Sherman et al., 1987; Meise and O'Reilly, 1996; Miller et al., 1998; Durbin et al., 2000). Although it was known that the dinoflagellate does not affect the survival of the calanoid (Campbell et al., 2005; Turner and Borkman, 2005; Turner, 2006; Teegarden et al., 2001; 2008; Hassett, 2003), no studies addressed other potential effects of those blooms. In this dissertation, I investigated the effect of *A. fundyense* on the fitness and physiology of the copepod *C. finmarchicus*. These experiments were designed to provide insights into *C. finmarchicus*/*A. fundyense* interaction under natural conditions. The main findings of my dissertation research are:

1. *Calanus finmarchicus* was resistant (in terms of high survival) to the toxic dinoflagellate over a 7-day experiment period even when feeding on high doses (HD) of this alga.

2. In spite of the similar ingestion rates in terms of carbon per female per day, reduced egg production and viability were observed at both high (HD) and moderate (LD) levels of *A. fundyense* in the diet during the course of the experiment. The total production of healthy nauplii was lower in both *A. fundyense* treatments, and significant differences in the HD treatment occurred in both June and July experiments, and in the LD treatment in July.

3. Lower reproductive success by females encountering *A. fundyense* blooms in the Gulf of Maine could potentially affect population growth; a significant reduction in the naupliar recruitment of females during summer could reduce the contribution to the following winter/spring populations (e.g., Miller et al., 1998; Saumweber and Durbin, 2006). However, further studies are needed to assess the magnitude of such an effect.

4. Significant transcriptional responses were measured in *C. finmarchicus* feeding on both *A. fundyense* diets at 2 and 5 days. Gene expression responses to the dinoflagellate differed between the two time points. Similar to the stereotyped cellular stress response (Kültz, 2003, 2005), large numbers of genes were differentially regulated at 2 days and included many genes that are usually regulated in response to stress.

5. At 5 days, fewer genes were differentially expressed, which would be expected for a physiological adjustment to restore the cellular homeostasis to new environmental conditions (i.e., *A. fundyense* diets). The differentially expressed genes at 5 days are indicative of a physiological adjustment to the energy balance that included down-regulation of biological processes associated with growth and reproduction, and down-

regulation of lipid biosynthesis genes. In conclusion, based on those energy budget considerations, it is clear that in spite comparable ingestion rates in μg C fem^{-1} d^{-1} of control and experimental females there are continued metabolic costs associated with the dinoflagellate diets. One possible explanation might be a reduction in food assimilation in the gut, since a number of digestive enzymes such as trypsins, amylases and cellulases were also down-regulated. However, the exact mechanisms remain unknown.

6. Surprisingly, only a small number of detoxification enzymes, which are fundamental in the stress defense mechanism of the cellular stress response (CSR), were regulated in adult females feeding on the toxic treatments. 15 genes (<7% of the number of detox enzymes in the reference transcriptome) involved in the three phases of detoxification were differentially expressed in adult females feeding on the toxic diets. The DE genes included members of the glutathione S-transferase and cytochrome P450 family with the latter being down-regulated. The pattern of global gene expression suggests that the dinoflagellate diet is acting as an environmental stressor that included some detoxification enzymes, but this response was limited to relative few genes, and the magnitude of the difference in expression was modest.

7. A deeper investigation of the glutathione S-transferase family showed high gene diversity with a total of 41 GSTs in *C. finmarchicus, which* rivals or exceeds that described for insects. Phylogenetic analysis and relative gene expression in different developmental stages suggested that the predicted proteins are likely to represent the products of distinct genes whose multiplicity might have been driven by diversity of diet (e.g. natural toxins)

and complexity of life history. However, in response to the *A. fundyense* diets only four out of 39 transcripts in the Gulf of Maine reference transcriptome were differentially expressed on day 2. None were differentially expressed on day 5.

8. Some genetic differences were found when the GST family identified in the transcriptome from the Gulf of Maine population (Lenz et al., 2014) was compared with the one identified in a separate transcriptome generated from individuals from the Norwegian Sea population (Tarrant et al., 2014). Those differences were not surprising given that the two populations are mostly isolated from each other (Bucklin et al., 1996; Unal et al., 2010) and separated by over 4,000 km. However, whether the GST genetic variations represent differences in functions in responses to habitat-specific selection (e.g. harmful algal blooms) has yet to be determined.

9. RNA-Seq data showed good agreement with RT-qPCR results performed on six target genes (three GSTs, three potential reference genes). This study confirmed that RNA-Seq is reliable and powerful tool for investigating global changes in gene expression in the copepod *C. finmarchicus*.

10. Additional data obtained by RT-qPCR indicated that regulation of the two GSTs that did alter their expression is very rapid, and is already present in females after 12 hours on the *A. fundyense* diet. Furthermore, relative gene expression patterns in females from the two experiments were very similar, in spite of other differences between June and July females. This reduces the number of extraneous factors that might contribute to the

pattern observed and supports the conclusion that they represent a response to a toxic alga diet.

Overall, the results of this dissertation support that the dinoflagellate *A. fundyense* is an environmental stressor for the copepod *C. finmarchicus*. The copepod is resistant to the toxic alga but experiences sub-optimal conditions in which physiological adjustments of its energy budget are required. The novelty of the study has been in combining the results obtained from fitness measurements with the gene expression analysis in order to start to understand the physiological response to the dinoflagellate, and relate it to the observed differences in egg production and viability. For instance, the lower egg production in females feeding on both low and high doses of *A. fundyense* might be explained by a reallocation of energy as suggested by a differential expression of genes involved in the regulation of the copepod energy budget. The down-regulation of genes involved in lipid biosynthesis, growth and reproduction characterized the response at 5 days. Reduction in reproduction output has potential negative impacts on the copepod population by decreasing recruitment. The decrease in hatching success remained a mystery. While a decrease in the energy available, in particular in the lipids, may explain the reduced investment in egg production, it still remains unclear why the egg viability is affected by the dinoflagellate. One possible explanation might be a direct interference of the dinoflagellate in development, as suggested in other studies that found that algal metabolites decrease egg viability (Ianora and Miralto, 2010). The broad response in the gene expression profile suggests that an organism's physiology may be affected by what might be viewed as a minor environmental stressor (i.e. no effect on survival) that might be a frequent occurrence in nature. These types of sub-lethal perturbations are the rule rather then the exception in natural

ecosystems (Beaugrand, 2015), and yet they could affect the fitness of key members of the biological community. *A fundyense* is just one example of a stressor that natural populations would expect to encounter in their natural environment. Like *A. fundyense*, these other stressors, may cause global physiological responses similar to what I observed in this study. Furthermore, even if survival is not affected, these stressors may affect the fitness of the organism and contribute to the biogeographic shifts to other environments reported for many zooplankton species (Beaugrand 2015).

Although the current study focused on the effect of *A. fundyense* on the physiology of adult females only, my results suggest that other developmental stages such as copepodites (stages CIV and CV) may be affected similarly. A similar down-regulation of lipid biosynthesis in the pre-adult stage (CV) could interfere with the building of lipid stores that are essential for diapause preparation. However, further studies are required to test this hypothesis. In addition, whether the saxitoxins or other metabolites were responsible for the observed effects on *C. finmarchicus* will require further study. In either case, significant blooms of *A. fundyense* would be expected to have a negative effect on *C. finmarchicus* populations.

APPENDICES

Supplemental Table 2.1. Phytoplankton cell densities for the grazing experiment in July. In the table for each treatment (CONTROL, LD, HD) listed the number of cells found in the "control grazing" container (see text) named here "no grazers", and the number of cell grazed by the copepod after 24 hr incubation named here "with grazers". The number of cells in the "no grazers" container if, higher then the target concentration (*), indicates that the algae was growing during the experimental period. The number of grazed cells have been calculated by the difference between the number of cells in the "no grazers" container and the number of cells left in the experimental containers with the copepods after the 24 h incubation using the Frost equations (see text). For each treatment counts were averaged for the 3 replicates.

Experimental period	CONTROL *Rhodomonas* sp.		LD *Rhodomonas* sp.		LD *A. fundyense*		HD *A. fundyense*	
	"no grazers"	"with grazers"	"no grazers"	"with grazers"	"no grazers"	"with grazers"	"no grazers"	"with grazers"
	cell mL^{-1}	cell mL^{-1}	cell mL^{-1}	cell mL^{-1}	cell mL^{-1}	cell mL^{-1}	cell mL^{-1}	cell mL^{-1}
Day 1	8100	3525	6000	2852	50	25	200	101
Day 2	8800	6713	6600	5164	55	45	220	167
Day 3	8880	6181	6600	4592	58	41	214	150
Day 4	8880	5878	6660	4331	58	38	214	143
Day 5	8800	5939	6660	4389	58	38	212	143
Day 6	8800	6157	6660	4509	54	37	216	150
Day 7	8720	5826	6540	4507	52	34	214	147

*Daily target concentration:
CONTROL: 8000 cell mL^{-1} *Rhodomonas* sp.; LD: 6000 cells mL^{-1} *Rhodomonas* sp. and 50 cell mL^{-1} *A. fundyense*; HD: 200 cells mL^{-1} *A. fundyense*

Supplemental Table 3.1. List of differentially expressed genes involved in response to stress (section 3.3.1, Figure 3) energy metabolism (section 3.3.2, Figure 4A) and reproduction and growth (section 3.3.3, Figure 4B) process. The list includes genes whose expression was significant different in *C. finmarchicus* feeding on *A. fundyense* (LD, HD) for 2 and 5 days compared to the control diet. Genes were recognized as differently expressed using the exact test ($p<0.05$) and a multiple comparison correction with Benjamini-Hochberg method (False discovery rate <5%) implemented by edgeR. Genes were annotated against the non-redundant protein database (nr-NCBI) using *blastx* algorithm with E 10^{-3} as cutoff. For each gene, Accession number (NCBI), annotation result and fold change in expression are listed. The direction of expression is indicated by red (up-regulated genes) and green (down-regulated genes). Transcripts are ordered as clustered by *heatmap.2* as shown in Figures 3 and 4.					
Gene ID.		2 days		5 days	
STRESS RESPONSE		LD	HD	LD	HD
Chaperone					
GAXK01080993	heat shock protein 70	1.95	0	0	0
GAXK01175783	heat shock protein 70	1.83	0	0	0
GAXK01052959	heat shock protein 70	0	2.01	0	0
comp11697_c0_seq1	heat shock protein 70	1.89	1.7	0	0
comp12366_c1_seq1	heat shock cognate 70 protein	2.01	2.02	0	0
GAXK01181466	heat shock protein 70	1.75	1.66	0	0
GAXK01070817	dnaj homolog subfamily b4	2.19	2.33	0	0
Antioxidant					
GAXK01147310	oxidative stress protein	1.92	0	0	0
GAXK01039921	pyridine -disulfide oxidoreductase	0	0	0	1.71
GAXK01150527	uncharacterized oxidoreductase dhs-27-like	0	0	0	-1.75
GAXK01030329	short-chain dehydrogenase reductase sdr	0	0	0	-1.65
GAXK01050421	trans- -enoyl- reductase-like	0	0	-2.05	-2.76
GAXK01020872	animal haem peroxidase family	0	0	0	2.09
GAXK01187041	ascorbate peroxidase	0	0	0	2.01
GAXK01079039	ascorbate peroxidase-like	0	0	2.33	0
GAXK01192041	carbonyl reductase 1	0	-2.13	0	0
GAXK01178489	peroxisomal multifunctional enzyme type 2	0	0	-2.65	-5.73
Detoxification					
GAXK01143274	multidrug resistance-2	0	-1.86	0	0
GAXK01081107	multidrug resistance-2	-2.78	-2.88	0	0
GAXK01120895	beta and gamma crystallin	-3.56	-2.56	0	0
GAXK01188041	nucleoredoxin-2	2.74	1.69	0	1.68
GAXK01178489	glutathione peroxidase	0	0	-1.82	0
GAXK01103733	cytochrome p450 2j2-like	0	0	0	-2.11
GAXK01079613	cytochrome p450 2j6-like	-2.64	0	0	-2.29
GAXK01141739	cytochrome family 11	0	0	0	-3.11
GAXK01154455	cytochrome p450 20a1-	0	0	-1.99	0
GAXK01109776	sulfotransferase	1.91	0	0	0
GAXK01135132	sulfotransferase 1c4	0	-1.64	0	0
GAXK01138359	glutathione s-transferase D	0	-1.88	0	0
GAXK01204953	glutathione s-transferase	2.4	2.58	0	0
GAXK01204961	glutathione transferase O	0	0	1.51	0
GAXK01204960	glutathione s-transferase S	2.5	0	0	0
Apoptosis					
GAXK01124013	caspase nc-like	5.14	4.82	0	0
comp260845_c1_seq1	caspase-1	2.74	2.37	0	0
GAXK01146371	caspase-1	2.12	3.01	0	0
GAXK01088802	caspase	0	2.66	0	0
GAXK01193726	caspase-3-like protein 2	2.7	0	0	0
GAXK01152025	caspase nc-like	1.83	0	0	0

comp1139223_c0_seq1	caspase-1	-4.49	0	0	0
Intracellular signaling					
GAXK01038487	chemotaxis sensory transducer	9.04	7.34	0	0
GAXK01190674	protein kinase	6	6.75	0	0
GAXK01027875	membrane-spanning 4- subfamily member	5.89	5.58	0	0
GAXK01086358	synaptosomal-associated protein 25	5.6	5.47	0	0
GAXK01035412	histamine h2 receptor	5.3	5.79	0	0
GAXK01106098	neurohypophysial n-terminal domain	3.65	4.15	0	0
GAXK01065637	phosphatidylethanolamine-binding protein	2.19	5.16	0	0
GAXK01124915	calcium-binding tyrosine phosphorylation	3.48	3.66	0	0
GAXK01038621	erp protein	3.36	3.79	0	0
GAXK01000917	surface antigen repeat protein	3.33	2.9	0	0
GAXK01174053	pdgf vegf receptor	3.02	2.42	0	0
GAXK01087823	ninein isoform 2	2.97	3.29	0	0
GAXK01095536	neurohypophysial n-terminal domain	2.95	2.99	0	0
GAXK01029328	testis seletal muscle phosphatase	2.56	2.48	0	0
GAXK01014506	sequestosome 1-like	2.44	2.2	0	0
GAXK01037391	scavenger receptor cysteine-rich protein	2.41	2.4	0	0
GAXK01144337	tyrosine kinase receptor	2.4	2.83	0	0
GAXK01071754	kinase c and casein kinase	2.39	2.65	0	0
GAXK01183210	uridine-cytidine kinase 1	2.31	2.18	0	0
GAXK01174906	huntingtin-associated protein-interacting	2.09	1.93	0	0
GAXK01173195	mtss1-like protein	2.04	2.26	0	0
GAXK01037025	tyrosine recombinase-like	1.96	1.88	0	0
GAXK01165291	pip n-terminal domain pip c-terminal	3.71	3.23	0	-2.5
GAXK01060375	phosphatidylinositol-4-phosphate 5-kinase	2.59	0	0	0
GAXK01093312	indian hedgehog homolog a	2.06	0	0	0
GAXK01150261	fyn-related kinase	2.02	0	0	0
GAXK01064897	adenylate cyclase	1.89	0	0	0
GAXK01064554	scavenger receptor cysteine-rich protein	1.79	0	0	0
GAXK01018368	agg2 paralog	0	1.79	0	0
GAXK01161757	tenascin c	0	1.72	0	0
GAXK01163832	adenylate cyclase	0	2.13	0	0
GAXK01117355	ww domain containing protein	0	1.91	0	0
GAXK01134657	phosphatidylethanolamine-binding protein	0	2.38	0	0
GAXK01044304	calmodulin 2	0	0	1.64	2.93
GAXK01004124	serine threonine-protein kinase pim-3	0	0	0	2.16
GAXK01003529	calmodulin	0	0	0	1.95
GAXK01143271	serine threonine-protein kinase chk2	0	0	0	1.61
GAXK01163554	tyrosine kinase receptor	0	-2.43	-2.97	-2.43
GAXK01204109	protein kinase	-1.98	-1.56	0	0
GAXK01072260	guanylate cyclase activator 1a	-2.26	-3.01	0	0
GAXK01038822	protein kinase domain containing protein	0	-9.55	0	0
GAXK01102284	protein kinase domain containing protein	0	0	-2.33	0
GAXK01068902	protein kinase	0	0	-2.64	0
GAXK01032723	kinesin family member 27	0	0	0	-1.64
GAXK01174189	wee1-like cdk tyrosine kinase	0	0	0	-1.7
Cell cycle checkpoint					
GAXK01173300	cyln2	3.96	6.4	0	0
GAXK01173949	cell cycle control 50a	2.58	2.43	0	0
GAXK01033493	annexin a11	3.27	3.04	0	0
GAXK01061806	ankyrin repeat domain-containing protein	3.65	4.62	0	0
GAXK01164011	ankyrin repeat protein	2.9	3.02	0	0

GAXK01001348	ankyrin repeat protein	2.12	2.42	0	0
GAXK01118271	ankyrin unc44	0	1.92	0	0
GAXK01158991	ankyrin repeat domain-containing protein	0	2.14	0	0
GAXK01071650	ankyrin repeat-containing	0	1.9	0	0
GAXK01127967	ankyrin repeat domain-containing protein	2.06	0	0	0
GAXK01008870	ankyrin repeat protein	2.72	0	0	0
GAXK01065956	budding uninhibited by benzimidazoles 3	0	2.21	0	0
GAXK01156205	cell division control 42	1.84	0	0	0
GAXK01200615	speckle-type poz	0	0	0	-1.58
GAXK01061398	epithelial cell transforming sequence 2	0	0	0	-1.78
GAXK01126384	aig1 family protein	0	0	0	-2.06
GAXK01081328	chromosome-associated kinesin	0	0	0	-2.11
GAXK01128023	smc-like protein	0	0	0	-2.15
GAXK01149031	smc n domain protein	0	0	0	-2.64
GAXK01153640	chromosome segregation protein smc	0	0	0	-1.93
GAXK01133334	chromosome segregation protein smc	0	0	0	-2.71
GAXK01148983	chromosome segregation protein smc	0	0	0	-2.81
GAXK01134115	meiosis-specific nuclear structural protein	0	0	3.93	-6.26
GAXK01143458	tyrosine recombinase	0	0	0	-4.8
GAXK01203925	aig1 family protein	0	0	-2.36	-4.28
GAXK01137796	aig1 family protein	0	0	0	-8.89
Protein turnover					
GAXK01030163	ornithine decarboxylase	1.83	0	0	1.5
GAXK01204181	d-lactate dehydrogenase	2.07	0	0	0
GAXK01010695	proline dehydrogenase mitochondrial-like	0	0	2.38	0
GAXK01002818	chorion peroxidase-like	0	-2.21	0	0
GAXK01157768	peptidyl-asp metalloendopeptidase	2.93	4.23	0	0
GAXK01092809	membrane metallo-endopeptidase-like 1	3.41	3.33	0	0
GAXK01013024	membrane metallo-endopeptidase-like 1	0	0	0	2.3
GAXK01071865	meprin a subunit beta	4.54	3.3	0	0
GAXK01058620	procollagen c-endopeptidase enhancer 2	2.92	3.14	0	0
GAXK01083352	matrix metalloproteinase-14-like	2.46	2.73	0	0
GAXK01034829	zinc metalloproteinase nas-13	3.6	2.54	0	0
GAXK01131627	zinc metalloproteinase nas-13	0	3.19	0	0
GAXK01015015	a disintegrin and metalloproteinase	0	2.56	0	2.14
GAXK01080219	a disintegrin and metalloproteinase	1.75	2.24	0	0
GAXK01137057	a disintegrin and metalloproteinase	-2.17	0	0	1.85
GAXK01129294	hatching enzyme	2.25	2.79	0	0
GAXK01053020	interstitial collagenase	0	0	0	-2.02
GAXK01058782	adhesive serine protease	0	1.6	0	0
GAXK01116212	serine protease	2.84	2.71	0	0
GAXK01188611	serine protease	0	1.64	0	0
GAXK01076666	serine protease	0	1.67	0	0
GAXK01081763	serine protease	5.45	5.49	0	0
GAXK01065197	af357226_1cub-serine protease	0	0	0	-4.96
GAXK01135009	blastula protease-10	2.43	0	0	0
GAXK01105326	serine protease easter	1.76	1.83	0	0
GAXK01188932	serine protease easter precursor	1.97	1.94	0	0
GAXK01098506	serine protease 14	3.56	2.9	0	0
GAXK01136255	serine protease 14	2.73	3.08	0	0
GAXK01069221	serine protease 14	2.61	2.69	0	0
GAXK01101808	proclotting enzyme	1.9	1.74	0	0
GAXK01069053	serine protease snake-like	0	2.75	0	0

GAXK01019334	serine proteinase stubble	0	1.98	0	0
GAXK01011395	serine proteinase stubble	0	2.04	0	0
GAXK01030376	serine proteinase stubble	0	0	0	2.16
GAXK01023022	secreted serine	5.29	5.99	0	0
GAXK01031646	transmembrane protease serine 6-like	5.29	6.22	0	0
GAXK01015312	stem cell tumor	2.42	2.92	0	0
GAXK01073101	peptidase s8 and s53 subtilisin kexin	0	0	1.25	0
GAXK01180762	peptidase s8 s53 subtilisin kexin sedolisin	0	0	-2.54	-1.63
GAXK01181808	kallikrein related-peptidase 6 precursor	0	0	2.08	0
GAXK01064074	melanization protease isoform c	6.3	6.01	0	0
GAXK01153694	coagulation factor vii	0	1.98	0	0
GAXK01097615	serine protease inhibitor 4	0	1.81	0	0
GAXK01050132	serine protease inhibitor	0	0	0	-1.66
GAXK01049411	proteinase inhibitor i4 serpin	0	0	-2.1	0
GAXK01134087	trypsin-like proteinase t2a precursor	0	2.18	0	0
GAXK01047165	serine proteinase stubble	0	0	0	-3.84
GAXK01034326	a chain fusarium oxysporum trypsin	0	0	-1.53	-1.51
GAXK01174145	inter-alpha-trypsin inhibitor heavy chain h1	0	0	0	-1.78
GAXK01043485	inter-alpha-trypsin inhibitor heavy chain	1.8	2.42	0	0
GAXK01138601	kunitz bovine pancreatic trypsin inhibitor	0	0	0	1.71
GAXK01112470	carboxypeptidase b	0	0	0	-2.1
GAXK01058659	carboxypeptidase d	5	4.11	0	0
GAXK01145526	carboxypeptidase d	0	-3.62	0	0
GAXK01076981	cysteine protease	2.51	0	0	0
GAXK01068401	papain-like cysteine proteinase	0	0	1.91	0
ENERGY METABOLISM					
Lipid metabolism Degradation	lipase member i	7.02	7.47	-2.35	0
GAXK01163432	member i	0	2.95	0	0
GAXK01172813	pancreatic lipase-related protein 2-like	4.13	4.55	0	-2.65
GAXK01188659	pancreatic lipase-related protein 2-like	2.44	3.46	-1.55	0
GAXK01065401	pancreatic triacylglycerol lipase-like	2.97	5.23	0	0
GAXK01162538	venom phospholipase a1	6.02	6.72	-1.88	0
GAXK01189635	lipase 3	0	0	2.19	0
GAXK01008610	pancreatic triacylglycerol lipase	0	1.85	0	2.83
GAXK01072825	monoglyceride lipase	0	0	2.04	2.62
GAXK01136240	phospholipase d2	0	1.87	0	0
GAXK01086947	phospholipase membrane-associated precursor	5.01	7.48	0	0
GAXK01145780	phosphatidylinositol- phospholipase c	5.31	5.13	0	0
GAXK01074713					
Lipid metabolism Transport					
GAXK01022983	low-density lipoprotein receptor-	3.28	0	0	0
GAXK01153019	sortilin-related receptor-like	0	0	-6.27	0
GAXK01006561	fatty acid-binding protein	0	0	0	1.9
GAXK01164342	fatty acid-binding protein	5.31	5.4	0	0
Lipid metabolism Biosynthesis					
GAXK01161646	elongation of very long chain fatty acids	-2.5	-2.5	-2.48	-6.45
GAXK01066253	acyl- desaturase-like	0	-1.59	0	-3.92

GAXK01060815	fatty acid desaturase 2	0	0	-2	-2.18
GAXK01005020	1-acylglycerol-3-phosphate acyltransferase	0	0	0	-1.6
GAXK01196059	isoform a	0	0	0	-2.22
GAXK01033133	long-chain-fatty-acid-- ligase 3	0	0	0	-2.68
GAXK01000971	long-chain-fatty-acid-- ligase 3	0	-1.81	0	0
Carbohydrate metabolism Degradation	isocitrate lyase	2.47	1.89	1.38	3.26
GAXK01124971	dtdp-glucose- -dehydratase	2.69	2.98	0	0
GAXK01120039	prolyl 4-hydroxylase alpha-1 subunit	2.64	2.45	0	0
GAXK01119657	n-formylglutamate amidohydrolase	2.12	2.25	0	0
GAXK01080143	6-phosphogluconolactonase	1.81	2.74	3.47	4.04
GAXK01124810	6-phosphogluconolactonase	0	0	-3.18	-2.93
GAXK01019524	acteylglucosamine pyrophosphorylase	0	0	0	-1.73
GAXK01144515	tubulin polyglutamylase ttll4-like	0	-1.75	0	0
GAXK01179641	probable tubulin polyglutamylase ttll1-like	0	0	0	-1.73
GAXK01000492	alpha-amylase	0	0	0	-2.52
GAXK01093592	hexokinase	3.22	3.5	0	0
GAXK01019025	chitinase a1	-3.06	0	0	0
GAXK01155574	secreted chitinase	0	-5.28	0	0
GAXK01145888	endo-b- -glucanase	3.11	3.71	2.68	2.91
GAXK01130942	beta- -endoglucanase	3.12	3.92	3.96	4.33
GAXK01120018	endoglucanase-type cellulase	0	1.99	0	1.44
GAXK01106967	endo-b- -glucanase	1.84	2.43	2.72	2.34
GAXK01146040	endo-b- -glucanase	3.11	3.71	2.68	2.91
GAXK01130942	endonuclease exonuclease phosphatase family	1.79	2.38	2.46	2.89
GAXK01017480	exo-beta- -glucanase	1.92	2.64	2.01	2.56
GAXK01094122	exo-beta- -glucanase	3.47	4.27	3.94	4.32
GAXK01045446	alcohol dehydrogenase domain protein	5.08	3.43	0	0
GAXK01123964	alcohol dehydrogenase domain protein	2.56	3.21	0	0
GAXK01081348	glyceraldehyde-3-phosphate dehydrogenase	0	0	0	-1.86
GAXK01133121	d-lactate dehydrogenase	2.07	0	0	0
GAXK01204181	protein	2.55	2.38	0	0
GAXK01071144	fructose -bisphosphate aldolase	0	0	0	-1.65
GAXK01175662	phosphoglycerate mutase	5.23	5.75	0	0
GAXK01125514	glycosyl hydrolase family7	3.64	3.81	3.45	3.97
GAXK01155750	glycosyl hydrolase family7	2.88	3.3	2.31	2.52
GAXK01112967	glycosyl hydrolase family7	2.7	3.32	2.7	3.38
GAXK01039246	glycosyl hydrolase family7	2.92	3.86	3.73	4.28
GAXK01111935	glycosyl hydrolase family45	2.39	3.19	1.98	2.62
GAXK01008278	glycoside hydrolase family 16	1.91	2.73	1.65	1.92
Carbohydrate metabolism Transport					
GAXK01150711	sugar transporter	3.07	3.44	0	0
GAXK01032944	sugar transporter	1.94	3.38	0	0
GAXK01097950	sugar transporter	6.17	6.49	0	0
GAXK01200835	sugar transporter	0	2.22	0	0
GAXK01038934	sugar transporter	0	1.78	0	0
Carbohydrate metabolism					

Biosynthesis					
GAXK01052069	glutamine synthetase	0	0	0	-1.89
GAXK01201188	acetoacetyl- synthetase	0	0	0	-2.2
GAXK01132106	succinyl-coa synthetase beta chain	0	0	0	1.91
GAXK01051558	phosphoenolpyruvate carboxykinase	1.7	1.6	0	0
REPRODUCTION AND GROWTH					
GAXK01121850	gametocyte-specific factor 1-like	1.84	1.86	0	0
GAXK01079264	growth factor receptor-bound protein 14	0	0	0	-1.77
GAXK01084117	embryonic muscle myosin heavy chain	0	0	0	-1.58
GAXK01130363	neuroparsin-a precursor	0	0	0	-1.67
GAXK01157316	high affinity nuclear juvenile hormone binding	0	0	0	2.08
GAXK01203015	juvenile hormone-inducible protein	0	0	2.34	3.08
GAXK01049756	ecdysteroid-regulated protein	0	0	-2.11	0
GAXK01024782	juvenile hormone esterase	0	0	-2.17	0
GAXK01120581	farnesoic acid o-methyltransferase	0	-2.14	0	0
GAXK01093948	male sterility domain-containing	0	-1.8	0	-1.92
GAXK01132258	vitellogenin-like protein	-2.79	-4.31	2.27	-4.23
GAXK01171775	vitellogenin-like protein	0	-2.32	0	0
GAXK01129937	vitellogenin-like protein	0	-3.1	0	0
GAXK01055578	integrin alpha-ps2	0	-2.51	0	0
GAXK01095302	bone morphogenetic protein 1b	0	-1.66	0	0
GAXK01037842	bone morphogenetic protein 1b	0	-2.17	0	0
GAXK01119068	bone morphogenic protein 1	0	-2.55	0	0
GAXK01119569	cuticular protein analogous to peritrophins	0	-2.68	0	0
GAXK01013619	flexible cuticle protein 12 precursor	0	-1.8	0	0
GAXK01170220	flexible cuticle protein 12 precursor	-1.61	0	0	0
GAXK01138337	cuticle protein 19	0	-2.23	0	0
GAXK01170219	cuticle protein 6	0	-3.58	0	0
GAXK01168529	cuticle protein 6	-1.91	-2.24	0	0
GAXK01168529	cuticle protein 6	-1.91	-2.24	0	0
GAXK01014239	cuticle protein 6	-1.93	-2.81	0	0
GAXK01014239	cuticle protein 6	-1.93	-2.81	0	0
GAXK01002818	chorion peroxidase-like	0	-2.21	0	0
GAXK01040487	chorion peroxidase	0	0	0	-1.9
GAXK01095530	chorion peroxidase	-3.03	-1.74	0	0
GAXK01111644	flexible cuticle protein 12 precursor	-2.53	-4.88	0	0
GAXK01128951	epididymal sperm-binding protein 1-like	-2.11	-1.57	0	-3.23
GAXK01138250	collagen alpha-1 chain flags: precursor	0	1.73	0	-1.57
GAXK01000492	probable tubulin polyglutamylase ttll1-like	0	0	0	-1.73
GAXK01162698	heavy polypeptide non-muscle	0	0	0	-2.8
GAXK01121850	prelamin-a c	0	1.78	0	0
GAXK01079264	kinetochore protein nuf2	0	0	0	-2.2
GAXK01084117	chromosome segregation protein smc	0	-1.67	0	0
GAXK01130363	cleavage and polyadenylation specific	0	-3.33	0	0
GAXK01157316	fibrillin-like protein	0	-5.26	0	0

Supplemental Table 5.1. List of genes and primers used in quantitative PCR analysis of *Calanus finmarchicus*. For each gene accession number (NCBI), primers (F, R) Amplicon size (bp), PCR efficiency and melting temperature (Tm) are listed.

Gene name	Accession No.	Primer sequence 5'-3'	Amplicon size (bp)	Efficiency (%)	Tm
Genes of interest					
GST-Delta-I	GAXK01204953	F: TCAGGTCACCATCCACAAGC	143	100	62
		R: AGCAGTCCACATGGCTTTGA			60
GST-Sigma-VI	GAXK01204959	F: CCCCTCCCCAGTAGAGCATA	167	98	64
		R: CTTCAACCTGAGAGCCCGAG			64
mGST-3-III	GAXK01204955	F: TCTTGCTCCCTGCTCAGAAT	120	99	60
		R: TTGCGGGCTCTTTGTTAAGT			58
Reference genes					
Actin	comp12286_c0_seq1 *	F: CCCAAGCCTATTGAGGTTCA	124	100	60
		R: CATACTGGGCCTTGGTGTGG			64
EFA	GAXK01169633	F: AATATGGGCGGTGTGACAAT	127	100	58
		R: CTCCGACTCCAAGAACAAGC			62
16S	GAXK01168561	F: CGTCTCTTCTAAGCTCCTGCAC	114	98	68
		R: AAGCTCCTCTAGGGATAACAGC			66

* NCBI submission for the gene actin is in process

REFERENCES

Abascal, F., Zardoya R., Posada D. 2005. ProtTest: selection of best-fit models of protein evolution. Bioinformatics. 21(9): 2104-5.

Abdalla, A.M., Bruns, M.C., Tainer, J.A., Mannervik, B. and Stenberg, G. 2002. Design of a monomeric human glutathione transferase GSTP1, a structurally stable but catalytically inactive protein. Protein Eng. 15(10): 827-34.

Andersen, C.L., Jensen, J.L., Orntoft, T.F. 2004. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Research,* 64, 5245–5250.

Anderson, D., Kulis, D., Sullivan, J., Hall, S.1990. Toxin composition variations in one isolate of the dinoflagellate *Alexandrium fundyense*. Toxicon 28: 885-893.

Anderson, D.M., White, A. W. 1992. Marine biotoxins at the top of the food chain. Oceanus 35(3): 55-61.

Anderson, D.M. 1997. Bloom dynamics of toxic *Alexandrium* species in the northeastern US. Limnology and Oceanography 42 (5 part2), 1009-1022.

Anderson, D.M., Townsend D.W., McGillicuddy,D.J. and Turner, J.T. 2005. The Ecology and Oceanography of Toxic *Alexandrium fundyense* blooms in the Gulf of Maine. Deep-Sea Res. II 52: (19-21): 2365-2876.

Anderson, D.M., Burkholder, J.M., Cochlan W.P., Glibert, P.M,. Gobler, C.J., Heil, C.A., et al. 2008. Harmful algal blooms and eutrophication: Examining linkages from selected coastal regions of the United States. Harmful Algae. 8(1): 39-53.

Anderson D.M., Alpermann, T.J., Cembella, A.D., Collos, Y., Masseret, E., Montresor, M. 2012. The globally distributed genus *Alexandrium*: multifaceted roles in marine ecosystems and impacts on human health. Harmful Algae 14, 10-35.

Anderson, D.M., McGillicuddy, D.J., DeGrasse, Jr., S.L., Sellner, K.G., Bricelj, V.M., Turner, J.T., Townsend, D.W. and Kleindinst, J.L. 2014. Harmful Algae in the Gulf of Maine: Oceanograph, Population Dynamics, and Toxin Transfer in the Food Web. Deep-Sea Res. II 103, 1-376.

Anderson, D.M., Keafer, B.A., Kleindinst, J.L., McGillicuddy, D.J., Martin, J.L., Norton, K., Pilskaln, C.H., Smith, J.L., Sherwood, C.R., Butman, B., 2014. *Alexandrium fundyense* cysts in the Gulf of Maine: long-term time series of abundance and distribution, and linkages to past and future blooms. Deep-Sea Research II 103, 6-26.

Aruda, A.M., Baumgartner, M.F., Reitzel, A.M., Tarrant, A.M., 2011. Heat shock protein expression during stress and diapause in the marine copepod *Calanus finmarchicus*. J. Insect Physiol. 57(5), 665-675.

Asselman, J., De Coninck, D. I., Glaholt, S., Colbourne, J. K., Janssen, C. R., Shaw, J. R., De Schamphelaere, K. A. 2012. Identification of pathways, gene networks, and paralogous gene families in *Daphnia pulex* responding to exposure to the toxic cyanobacterium *Microcystis aeruginosa*. Environmental science & technology, 46(15): 8448-8457.

Astthorsson, O.S., Gislason, A., 1995. Long-term changes in zooplankton biomass in Icelandic waters in spring. ICES J. Mar.Sci: Journal du Conseil 52(3-4), 657-668.

Ayres C., Müeller P., Dyer N., Wilding C., Rigden D., Donnelly M. 2011. Comparative Genomics of the Anopheline Glutathione S-Transferase Epsilon Cluster. PloS One. 6(12).

Båmstedt, U., Nejstgaard, J.C., Solberg, P.T., Høisœter, T., 1999. Utilization of small-sized food

algae by *Calanus finmarchicus* (Copepoda, Calanoida) and the significance of feeding history. Sarsia 84, 19-38.

Beaugrand, G., Brander K.M., Lindley J.A., Souissi S., Reid P.C. 2003. Plankton effect on cod recruitment in the North Sea. Nature. 426: 661–664.

Beaugrand,G. 2015. Marine Biodiversity, Climatic Variability and Global Change. EarthScan Oceans.

Bigelow, H.B., 1926. Plankton of the offshore waters of the Gulf of Maine. Bulletin of the US Bureau of Fisheries 40, 1-509.

Board, P.G., Menon D. 2013. Glutathione transferases, regulators of cellular metabolism and physiology. Biochim Biophys Acta.1830(5):3267-88.

Bratbak, G., Jacquet S., Larsen, A., Pettersson L.H., Sazhin, A.F., Thyrhaug, R. 2011. The plankton community in Norwegian coastal waters-abundance, composition, spatial distribution and diel variation. Cont Shelf Res. 31(14):1500-14.

Bresell, A., Weinander R., Lundqvist, G., Raza, H., Shimoji M, Sun TH, et al. 2005. Bioinformatic and enzymatic characterization of the MAPEG superfamily. FEBS J. 272(7): 1688-703.

Bricelj, V.M., Shumway, S.E., 1998. Paralytic shellfish toxins in bivalve molluscs: occurrence, transfer kinetics and biotransformation. Reviews in Fisheries Science 6: 315-383.

Bucklin A., Sundt R., Dahle G. 1996. Population genetics of *Calanus finmarchicus* in the North Atlantic. Ophelia. 44: 29–45.

Bustin, S.A., Benes, V., Garson, J. A., Hellemans, J.,Huggett, J.,Kubista M., Mueller, R., Nolan, T., Pfaffl, M. W., Shipley, G. L., Vandesompele, J., Wittwer, C. T. 2009. The MIQE

guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem*., 55, 611-622.

Bustin, S.A., Beaulieu, J.F., Huggett, J., Jaggi, R., Kibenge, F.S.B., Olsvik, P.A., Penning, L.C., Toegel, S., 2010. MIQE precise: practical implementation of minimum standard guidelines for fluorescence-based quantitative real-time PCR experiments. BMC Molecular Biology 11.

Campbell, R.G., Runge, J., Durbin, E., 2001. Evidence for food limitation of *Calanus finmarchicus* production rates on the southern flank of Georges Bank during April 1997. Deep Sea Res. II 48(1), 531-549.

Campbell, R.G., Teegarden, G.J., Cembella, A.D., Durbin, E.G., 2005. Zooplankton grazing impacts on *Alexandrium* spp. in the nearshore environment of the Gulf of Maine. Deep-Sea Res. II 52, 2817-2833.

Causton, H.C., Ren B, Koh SS, Harbison CT, Kanin E, Jennings EG, Lee TI, True HL, Lander ES, Young RA. 2001. Remodeling of yeast genome expression in response to environmental changes. Mol. Biol. Cell 12:323–337

Ceballos, S., Ianora, A., 2003. Different diatoms induce contrasting effects on the reproductive success of the copepod *Temora stylifera*. Journal of Experimnental Marine Biology and Ecology 294, 189-202.

Cembella, A.D., 2003. Chemical ecology of eukaryotic microalgae in marine ecosystems. Phycologia 42,420–447.

Che-Mendoza, A. Penilla R.P. and Rodrigues D.A. 2009. Insecticide resistance and glutathione S-transferases in mosquitoes: A review. Afr J Biotech. 8:1386-97.

Chen, L., Zhang, H., Finiguerra, M., Bobkov, Y., Bouchard, C., Avery, D. E., Anderson,P.A.V.,

Lin, S., Dam, H.G., 2015. A novel mutation from gene splicing of a voltage-gated sodium channel in a marine copepod and its potential effect on channel function. J. Exp. Mar. Biol. Ecol. 469, 131–142.

Christie, A.E., Fontanilla, T.M., Nesbit, K.T., Lenz, P.H. 2013. Prediction of the protein components of a putative *Calanus finmarchicus* (Crustacea, Copepoda) circadian signaling system using a *de novo* assembled transcriptome. Comp Biochem Phys D, Genomics Proteomics. 8(3): 165-93.

Christie, A.E., Roncalli, V., Wu, L.S., Ganote, C.L., Doak, T., Lenz, P.H. 2013. Peptidergic signaling in *Calanus finmarchicus* (Crustacea, Copepoda): *in silico* identification of putative peptide hormones and their receptors using a *de novo* assembled transcriptome. Gen Comp Endocrinol. 187:117-35.

Christie, A.E., Fontanilla, T.M., Roncalli, V., Cieslak, M.C., Lenz, P.H. 2014. Diffusible gas transmitter signaling in the copepod crustacean *Calanus finmarchicus*: Identification of the biosynthetic enzymes of nitric oxide (NO), carbon monoxide (CO) and hydrogen sulfide (H2S) using a *de novo* assembled transcriptome. Gen Comp Endocrinol. 202:76-86.

Christie, A.E., Fontanilla, T.M., Roncalli, V., Cieslak, M.C., Lenz, P.H. 2014. Identification and developmental expression of the enzymes responsible for dopamine, histamine, octopamine and serotonin biosynthesis in the copepod crustacean *Calanus finmarchicus*. Gen Comp Endocrinol. 195:28-39.

Colbourne, J.K., Pfrender, M.E., Gilbert, D., Thomas, W.K., Tucker, A., Oakley, T.H., et al. 2011. The ecoresponsive genome of *Daphnia pulex*. Science. 331(6017):555-61.

Colin, S.P., Dam, H.G., 2002. Latitudinal differentiation in the effects of the toxic dinoflagellate

Alexandrium spp. on the feeding and reproduction of populations of the copepod *Acartia hudsonica*. Harmful Algae1,113-125.

Colin, S.P., Dam, H.G., 2007. Comparison of the functional and numerical responses of resistant versus non-resistant populations of the copepod *Acartia hudsonica* fed the toxic dinoflagellate *Alexandrium tamarense*. Harmful Algae 6, 875-882.

Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M., Robles, M., 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21(18), 3674-3676.

Conover, R.,1988. Comparative life histories in the genera *Calanus* and *Neocalanus* in high latitudes of the northern hemisphere. Hydrobiologia 167, 127-142.

Costa, C., Giménez-Capitán, A., Karachaliou, N., & Rosell, R. 2013. Comprehensive molecular screening: from the RT-PCR to the RNA-seq. Translational Lung Cancer Research, *2*(2), 87-91.

d'Ippolito G, Lamari, N., Montresor, M., Romano, G., Cutignano, A., Gerecht, A. et al. 2009. 15S-Lipoxygenase metabolism in the marine diatom *Pseudo-nitzschia delicatissima*. New Phyto.183(4):1064-71.

da Fonseca, R.R., Johnson, W.E., O'Brien, S.J., Vasconcelos, V., Antunes, A. 2010. Molecular evolution and the role of oxidative stress in the expansion and functional diversification of cytosolic glutathione transferases. BMC Evol Biol. 10:281.

Dale, T., Kaartvedt, S., Ellertsen, B., Amundsen, R. 2001. Large-scale oceanic distribution and population structure of *Calanus finmarchicus* in relation to physical food and predators. Mar Biol. 139(3): 561-74.

Dale, B., Edwards, M., Reid, P., 2006. Climate change and harmful algal blooms, Ecology of

harmful algae. Springer, pp. 367-378.

Dam, H.G., 2013. Evolutionary adaptation of marine zooplankton to global change. Annu. Rev. Mar. Sci. 5, 349–370.

Darbyson, E., Swain, D., Chabot, D., Castonguay, M., 2003. Diel variation in feeding rate and prey composition of herring and mackerel in the southern Gulf of St Lawrence. J. Fish Bio. 63, 1235-1257.

Davis, C.S., 1987. Zooplankton life cycles. Georges Bank. MIT Press, Cambridge, MA, 256-267.

Deeds, J.R., Petitpas, C.M., Shue, V., White, K.D., Keafer, B.A., McGillicuddy, D.J., Milligan, P.J., Anderson, D.M., Turner, J.T., 2014. PSP toxin levels and plankton community composition and abundance in size-fractionated vertical profiles during spring/summer blooms of the toxic dinoflagellate *Alexandrium fundyense* in the Gulf of Maine and on Georges Bank, 2007, 2008, and 2010: 1. Toxin levels. Deep-Sea Res. II 103, 329-349.

Demott, W., Muller-Navarra, D., 1997. The importance of highly unsaturated fatty acids in zooplankton nutrition: evidence from experiments with *Daphnia*, a cyanobacterium and lipid emulsions. Freshwater Biology 38(3), 649-664.

Diel, S.,Tande, K., 1992. Does the spawning of *Calanus finmarchicus* in high latitudes follow a reproducible pattern? Marine Biology 113, 21-31.

Doucette, G.J., Turner, J.T., Powell, C.L., Keafer, B.A., Anderson, D.M., 2005. Trophic accumulation of PSP toxins in zooplankton during *Alexandrium fundyense* blooms in Casco Bay, Gulf of Maine, April–June 1998. I. Toxin levels in A. fundyense and zooplankton size fractions. Deep-Sea Research II 52, 2764-2783.

Doucette,G., Cembella, A., Martin, J., Michaud, J., Cole, T., Rolland, R., 2006. Paralytic

shellfish poisoning (PSP) toxins in North Atlantic right whales *Eubalaena glacialis* and their zooplankton prey in the Bay of Fundy, Canada. Mar. Ecol. Progr. Ser. 306, 303-313.

Durbin, E., Garrahan, P., Casas, M., 2000. Abundance and distribution of *Calanus finmarchicus* on the Georges Bank during 1995 and 1996. ICES J. Mar. Sci. 57, 1664-1685.

Durbin, E., Teegarden, G., Campbell, R., Cembella, A., Baumgartner, M.F., Mate, B.R., 2002. North Atlantic right whales, *Eubalaena glacialis*, exposed to paralytic shellfish poisoning (PSP) toxins via a zooplankton vector, *Calanus finmarchicus*. Harmful Algae 1, 243-251.

Dutz, J., 1998. *Alexandrium lusitanicum*: relationship between feeding and egg production. Mar. Ecol. Progr. Ser. 175, 97-107.

Enayati, A., Ranson, H., Hemingway, J. 2005. Insect glutathione transferases and insecticide resistance. Insect Mol Biol. 14(1):3-8.

Etheridge, S.M., 2010. Paralytic shellfish poisoning: seafood safety and human health perspectives. Toxicon 56(2),108-122.

Fabrini, R., De Luca, A., Stella, L., Mei, G., Orioni, B., Ciccone, S., Federici, G. et al. 2009. Monomer-dimer equilibrium in glutathione transferases: a critical re-examination. Biochemistry. 48(43): 10473-82.

Fang, Z., and Cui, X. 2011. Design and validation issues in RNA-seq experiments. Briefings in bioinformatics, bbr004.

Feder, M.E., and Hofmann, G.E.1999. Heat-shock proteins, molecular chaperones, and the stress response: evolutionary and ecological physiology. Annual review of physiology, 61(1), 243-282.

Fernández-Reiriz, M.J., Navarro, J.M., Contreras, A., Labarta, U., 2008. Trophic interactions between the toxic dinoflagellate *Alexandrium catenella* and *Mytilus chilensis*: Feeding

and digestive behavior to long-term exposure. Aqua. Toxicol. 87(4), 245-251.

Fields, D.M., Runge, J.A., Thompson, C., Shema, S.D., Bjelland, R.M., Durif, C.M.F., Skiftesvik, A.B., Browman, H.I., 2015. Infection of the planktonic copepod *Calanus finmarchicus* by the parasitic dinoflagellate, *Blastodinium* spp: effects on grazing, respiration, fecundity and fecal pellet production. Journal of Plankton Research 37(1), 211-220.

Fiksen, Ø., 2000. The adaptive timing of diapause–a search for evolutionarily robust strategies in *Calanus finmarchicus*. ICES Journal of Marine Science 57, 1825-1833.

Finiguerra, M., Avery, D.E., Dam, H.G., 2014. Sodium channel expression in the copepod *Acartia hudsonica* as a function of exposure to paralytic shellfish toxin (PST). Harmful Algae 39, 75-80.

Foley, K.P., Leonard M.W., Engel J.D. 1993. Quantitation of RNA using the polymerase chain reaction. Trends Genet. 9, 380-5.

Fontana, A., d'Ippolito, G., Cutignano, A., Romano, G., Lamari, N., Gallucci, A.M. et al. 2007. LOX-induced lipid peroxidation mechanism responsible for the detrimental effect of marine diatoms on Zooplankton grazers. Chembiochem. 8(15): 1810-18.

Frangópulos, M.,Guisande, C., Maneiro, I., Riveiro, I., Franco, J., 2000. Short-term and long-term effects of the toxic dinoflagellate *Alexandrium minutum* on the copepod *Acartia clausi*. Marine Ecology Progress Series 203, 161-169.

Friedman, R. 2011. Genomic organization of the glutathione S-transferase family in insects. Mol Phylogen Evol. 61(3):924-32.

Frost, B., 1972. Effects of size and concentration of food particles on the feeding behavior of the marine planktonic copepod *Calanus pacificus*. Limnology and Oceanography 17, 805-

815.

Frova, C. 2006. Glutathione transferases in the genomics era: new insights and perspectives. Biomol Eng. 23(4):149-69.

Gallardo-Escárate, C., Valenzuela-Muñoz, V., & Nuñez-Acuña, G. 2014. RNA-Seq analysis using de novo transcriptome assembly as a reference for the salmon louse Caligus rogercresseyi. PloSone, *9*(4), e92239.

Gasch, A.P., Spellman, P.T., Kao, C.M., Carmel-Harel, O., Eisen, M.B., Storz, G., Botstein, D., Brown, P.O., 2000. Genomic expression programs in the response of yeast cells to environmental changes. Mol. Biol. Cell 11(12), 4241-4257.

Gasch, A.P., Werner-Washburne, M., 2002. The genomics of yeast responses to environmental stress and starvation. Funct. Integr. Genomics 2(4-5), 181-192.

Gerecht, A., Carotenuto, Y., Ianora, A., Romano, G., Fontana, A., d'Ippolito, G., Jakobsen, H. H., Nejstgaard, J. C., 2013. Oxylipin production during a mesocosm bloom of *Skeletonema marinoi*. Journal of Experimental Marine Biology and Ecology 446, 159-165.

Gettings R.M., Townsend, D.W., Thomas, M.A., Karp-Boss, L. 2014. Dynamics of late spring and summer phytoplankton communities on Georges Bank, with emphasis on diatoms, *Alexandrium* spp., and other dinoflagellates. Deep-Sea Res Pt II. 103:120-38.

Gibson, U.E., Heid, C.A., Williams P.M.1996. A novel method for real time quantitative RT-QPCR. Genome Res. 6, 995-1001.

Gry, M.,Rimini, R., Strömberg, S., Asplund, A., Pontén, F., Uhlén, M., Nilsson, P., 2009. Correlations between RNA and protein expression profiles in 23 human cell lines. BMC Genomics 10(1), 365.

Guillard, R.,1973. Division rates. In: Stein, J. R. (Ed.), Handbook of Phycological Methods. Cambridge University Press, Cambridge, pp. 289-312.

Guisande,C.,Frangópulos, M., Maneiro, I., Vergara, A.R., Riveiro, I., 2002. Ecological advantages of toxin production by the dinoflagellate *Alexandrium minutum* under phosphorus limitation. Mar. Ecol. Progr. Ser. 225, 169-176.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., et al. 2013. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc. 8(8):1494-512.

Habig, W.H., Pabst, M., Jakoby, W.B. 1974. Glutathione S-transferases: on the first enzymatic step in mercapturic acid biosynthesis. J Biol Chem. 249:7130-9.

Hahn, D.A., Denlinger, D.L., 2011. Energetics of insect diapause. Ann. Rev. Entomol. 56, 103-121.

Hallegraeff, G.M., Anderson, D. M., Cembella, A. D. 2003. Manual of Harmful Marine Microalgae. UNESCO, Paris, 793.

Hammann, S., Tillmann, U., Schröder, M., Vetter, W., 2013. Profiling the fatty acids from a strain of the microalgae *Alexandrium tamarense* by means of high-speed counter-current chromatography and gas chromatography coupled with mass spectrometry. J. Chromatogr. A 1312, 93-103.

Hansen, B.H., Altin D., Booth A., Vang S.H., Frenzel M., Sorheim K.R., et al. 2010. Molecular effects of diethanolamine exposure on *Calanus finmarchicus* (Crustacea: Copepoda). Aquat Toxicol. 99(2), 212-22.

Hansen, B.H., Nordtug T., Altin D., Booth A., Hessen K.M., Olsen A.J. 2009. Gene Expression of GST and CYP330A1 in Lipid-Rich and Lipid-Poor Female *Calanus finmarchicus*

(Copepoda: Crustacea) Exposed to Dispersed Oil. J Toxicol Environ Health A. 72(3-4), 131-9.

Hansen, B.H., Altin D., Rorvik S.F., Overjordet I.B., Olsen A.J., Nordtug T. 2011. Comparative study on acute effects of water accommodated fractions of an artificially weathered crude oil on *Calanus finmarchicus* and *Calanus glacialis* (Crustacea: Copepoda). Sci Total Environ. 409(4), 704-9.

Harris, R.P., Irigoien, X., Head, R.N., Rey, C., Hygum, B.H., Hansen, B.W., et al. 2000. Feeding, growth, and reproduction in the genus *Calanus*. ICES J Mar Sci. 57(6):1708-26.

Hassett, R.P., 2003. Effect of toxins of the 'red-tide'dinoflagellate *Alexandrium* spp. on the oxygen consumption of marine copepods. J. Plankton Res.25, 185-192.

Hayes, J.D., Pulford, D.J. The glutathione S-Transferase supergene family: Regulation of GST and the contribution of the isoenzymes to cancer chemoprotection and drug resistance. Crit Rev Biochem Mol Biol. 1995;30(6):445-600.

Head, E.J.H., Harris, L.R., Campbell, R.W. 2000. Investigations on the ecology of *Calanus* spp. in the Labrador Sea. I. Relationship between the phytoplankton bloom and reproduction and development of *Calanus finmarchicus* in spring. Mar Ecol Prog Ser. 193:53-73.

Heath, M.R. and Lough, R.G. 2007. A synthesis of large-scale patterns in the planktonic prey of larval and juvenile cod (*Gadus morhua*). Fish Oceanogr.16 (2):169-85.

Helland, S., Nejstgaard, J., Humlen, R., Fyhn, H., Båmstedt, U., 2003. Effects of season and maternal food on *Calanus finmarchicus* reproduction, with emphasis on free amino acids. Marine Biology 142, 1141-1151.

Henry, R.P.,Thomason, K.L., & Towle, D.W. 2006. Quantitative changes in branchial carbonic anhydrase activity and expression in the euryhaline green crab, *Carcinus maenas*, in

response to low salinity exposure. J Exp Zool A Comp Exp Biol 305(10), 842-850.

Hirche, H.J., 1990. Egg production of *Calanus finmarchicus* at low temperature. Marine Biology 106, 53-58.

Hirche, H.J., Kwasniewski, S.1997. Distribution, reproduction and development of *Calanus* species in the Northeast Water in relation to environmental conditions. Journal of Marine Systems 10, 299-317.

Holm, P.J., Bhakat, P., Jegerschold, C., Gyobu, N., Mitsuoka, K., Fujiyoshi, Y., et al. 2006. Structural basis for detoxification and oxidative stress protection in membranes. J Mol Bio. 360(5): 934-945.

Ianora, A., Miralto, A., Romano, G. 2012. Antipredatory Defensive Role of Planktonic Marine Natural Products. In: Fattorusso E, Gerwick WH, Taglialatela-Scafati O. Handbook of Marine Natural Products: Springer Dordrecht, Heidelberg; 711-48.

Ianora, A., Poulet, S.A., Miralto, A., 2003. The effects of diatoms on copepod reproduction: a review. Phycologia 42, 351-363.

Ianora, A., Miralto, A., Poulet, S.A., Carotenuto, Y., Buttino, I., Romano, G., Casotti, R., Pohnert, G., Wichard, T., Colucci-D'Amato, L., 2004. Aldehyde suppression of copepod recruitment in blooms of a ubiquitous planktonic diatom. Nature 429, 403-407.

Ianora, A., Turner, J. T., Esposito, F., Carotenuto, Y., d'Ippolito, G., Romano, G., Fontana, A., Guisande, C., Miralto, A. 2004. Copepod egg production and hatching success is reduced by maternal diets of a non-neurotoxic strain of the dinoflagellate *Alexandrium tamarense*. Marine Ecology Progress Series 280, 199-210.

Ianora, A., Miralto, A., 2010. Toxigenic effects of diatoms on grazers, phytoplankton and other microbes: a review. Ecotoxicology 19, 493-511.

Ianora, A., Bastianini, M., Carotenuto, Y., Casotti, R., Roncalli, V., Miralto, A., Romano, G., Gerecht, A., Fontana, A., Turner, J.T., 2015. Non-volatile oxylipins can render some diatom blooms more toxic for copepod reproduction. Harmful Algae 44, 1-7.

Islam, S.M. and Tanaka M. 2004. Impacts of pollution on coastal and marine ecosystems including coastal and marine fisheries and approach for management: a review and synthesis. Mar Pollut Bull. 48(7-8): 624-49.

Jakobsson, P.J., Morgenstern, R., Mancini, J., Ford-Hutchinson, A., Persson, B. 1999. Common structural features of MAPEG - A widespread superfamily of membrane associated proteins with highly divergent functions in eicosanoid and glutathione metabolism. Prot Sci. 8(3): 689-92.

Jansen, S., Riser, C.W., Wassmann, P., Bathmann, U., 2006. Copepod feeding behaviour and egg production during a dinoflagellate bloom in the North Sea. Harmful Algae 5, 102-112.

Johnson, C.L., Runge, J.A., Curtis, K.A., Durbin, E.G., Hare, J.A. et al. 2011. Biodiversity and ecosystem function in the Gulf of Maine: Pattern and role of zooplankton and pelagic nekton. Plos One 6: e16491.

Katoh, K., Misawa, K., Kuma, K., Miyata, T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30(14):3059-66.

Katoh, K., Toh, H. 2008. Recent developments in the MAFFT multiple sequence alignment program. Brief Bioinform. 9(4):286-98.

Katoh, K., Standley, D.M. 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol Biol Evol. 30(4):772-80.

Kleindinst, J.L., Anderson, D.M., McGillicuddy, D.J., Stumpf, R.P., Fisher, K.M., Couture, D.A., Hickey, J.M., Nash, C., 2014. Categorizing the severity of paralytic shellfish

poisoning outbreaks in the Gulf of Maine for forecasting and management. Deep-Sea Res. Part II: Topical Studies in Oceanography 103, 277-287.

Koehl, M.A.R. and Strickler, J.R. 1981. Copepod feeding currents - food capture at low Reynolds number. Limnol Oceanogr. 26(6):1062-73.

Kohen, R.K.,Bayne, B.L.,1989. Towards a physiological and genetical understanding of the energetics of the stress response. Biol. J Linn. Soc. 37(1-2), 157-171.

Kozlowsky-Suzuki, B., Carlsson, P., Rühl et al. 2006. Food selectivity and grazing impact on toxic *Dinophysis* spp. by copepods feeding on natural plankton assemblages. Harmful Algae 5, 57-68.

Kültz, D., 2003. Evolution of the cellular stress proteome: from monophyletic origin to ubiquitous function. J. Exper. Bio. 206(18), 3119-3124.

Kültz D. Molecular and evolutionary basis of the cellular stress response. Annu Rev Physiol. 2005;67:225-57.

Langmead, B., Trapnell, C., Pop, M., Salzberg, S.L., 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10(3), R25.

Lauritano, C., Borra, M., Carotenuto, Y., Biffali, E., Miralto, A., Procaccini, G., et al. 2011. First molecular evidence of diatom effects in the copepod *Calanus helgolandicus*. J Exp Mar Bio Ecol. 2011;404(1-2):79-86.

Lauritano, C., Procaccini, G., Ianora, A. 2012. Gene expression patterns and stress response in marine copepods. Marine Environ Res. 76:22-31.

Lauritano, C., Carotenuto, Y., Procaccini, G., Turner, J.T., Ianora, A. 2013. Changes in expression of stress genes in copepods feeding upon a non-brevetoxin-producing strain of the dinoflagellate *Karenia brevis*. Harmful Algae. 28:23-30.

Lee, K.W., Rhee J-S, Han, J., Park, H.G., Lee, J-S. 2012. Effect of culture density and antioxidants on naupliar production and gene expression of the cyclopoid copepod, *Paracyclopina nana*. Comp Biochem Phys A, Mol Integr Physiol. 161(2): 145-52.

Lee, K.W., Raisuddin, S., Rhee, J.S., Hwang, D.S., Yu, I.T., Lee, Y.M., et al. 2008. Expression of glutathione S-transferase (GST) genes in the marine copepod *Tigriopus japonicus* exposed to trace metals. Aquat Toxicol. 89(3):158-66.

Lee, R.F., Hagen, W., Kattner, G. 2006. Lipid storage in marine zooplankton. Mar. Ecol. Prog. Ser. 307: 273–306.

Lehane, L., 2001. Paralytic shellfish poisoning: a potential public health problem. Med. J. Aust. 175(1), 29-31.

Lenz P.H., Unal E, Hassett RP, Smith CM, Bucklin A, Christie AE, et al. 2012. Functional genomics resources for the North Atlantic copepod, *Calanus finmarchicus*: EST database and physiological microarray. Comp. Biochem. Phys. D, Genomics Proteomics 27(2), 110-123.

Lenz, P.H., Roncalli, V., Hassett, R.H., Wu, L.S., Cieslak, M.C., Hartline, D.K., Christie, A.E. 2014. *De Novo* Assembly of a Transcriptome for *Calanus finmarchicus* (Crustacea, Copepoda) – The Dominant Zooplankter of the North Atlantic Ocean. PloS One 9:(2), e88589.

Letunic, I., Doerks, T., Bork, P. 2009. SMART 6: recent updates and new developments. Nucleic Acids Res.37: 229–232.

Liao, B.Y., Weng, M.P., 2015. Unraveling the association between mRNA expressions and mutant phenotypes in a genome-wide assessment of mice. Proc. Natl. Acad. Sci. 112(15), 4707-4712.

Liu, S., Wang, W.X., 2002. Feeding and reproductive responses of marine copepods in South China Sea to toxic and nontoxic phytoplankton. Maine Biology 140, 595-603.

Llewellyn, L.E. 2006. Saxitoxin, a toxic marine natural product that targets a multitude of receptors. Nat Prod Rep. 23(2):200-22.

Madsen, S.J., Nielsen, T.G., Tervo, O.M., Soderkvist, J., 2008. Importance of feeding for egg production in *Calanus finmarchicus* and *Calanus glacialis* during the Arctic spring. Marine Ecology Progress Series, 353, 177-190.

Martin, J.L., White, A.W., 1988. Distribution and abundance of the toxic dinoflagellate *Gonyaulax excavata* in the Bay of Fundy. Can. J. Fish. Aquat.. Sci. 45, 1968-1975.

Martin, J.L., White, A.W., 1988. Distribution and abundance of the toxic dinoflagellate *Gonyaulax excavata* in the Bay of Fundy. Canadian Journal of Fisheries and Aquatic Sciences 45, 1968-1975.

Mauchline, J., 1998. The Biology of Calanoid Copepods: The Biology of Calanoid Copepods. Academic Press.

McGillicuddy, D.J., Townsend, D.W., Keafer, B.A., Thomas, M., Anderson, D.M., 2014. Georges Bank: a leaky incubator of *Alexandrium fundyense* blooms. Deep-Sea Research II 103, 163-173.

McLellan, L.I. and Wolf, C.R. 1999. Glutathione and glutathione-dependent enzymes in cancer drug resistance. Drug Resist Update.2:153-64.

Meise, C., O'Reilly, J., 1996. Spatial and seasonal patterns in abundance and age-composition of *Calanus finmarchicus* in the Gulf of Maine and on Georges Bank: 1977–1987. Deep-Sea Res.II 43(7), 1473-1501.

Melle, W., Runge, J., Head, E., Plourde, S., Castellani, C., Licandro, P., Pierson, J., Jonasdottir,

S., Johnson, C., Broms, C., 2014. The North Atlantic Ocean as habitat for *Calanus finmarchicus*: environmental factors and life history traits. Progress in Oceanography 129, 244-284.

Menden-Deuer, S.,Lessard, E.J., 2000. Carbon to volume relationships for dinoflagellates, diatoms, and other protist plankton. Limnology and Oceanography 45, 569-579.

Miller, C.B., Lynch, D.R., Carlotti, F., Gentleman, W., Lewis, C.V., 1998. Coupling of an individual-based population dynamic model of *Calanus finmarchicus* to a circulation model for the Georges Bank region. Fish. Oceanogr.7, 219-234.

Miralto, A., Ianora, A., Poulet, S., 1995. Food type induces different reproductive responses in the copepod *Centropages typicus.* Journal of Plankton Research 17, 1521-1534.

Miralto, A., Barone, G., Romano, G., Poulet, S.A., Ianora, A., Russo, G.L., et al. 1999. The insidious effect of diatoms on copepod reproduction. Nature. 402(6758):173-6.

Mittapalli, O., Neal, J.J., and Shukle, R.H.. 2007. Tissue and life stage specificity of glutathione S-transferase expression in the Hessian fly, *Mayetiola destructor*: Implications for resistance to host allelochemicals. J Insect Sci. 7:20.

Morel, F. and Aninat, C. 2011. The glutathione transferase kappa family. Drug Metab Rev. 43:281-91.

Mortazavi, A., Williams, B., McCue, K., Schaeffer, L., Wold, B., 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5(7), 621 - 628.

Nagalakshmi, U., Waern, K., Snyder, M., 2010. RNA-Seq: a method for comprehensive transcriptome analysis. Current Protocols in Molecular Biology, 4.11. 11-14.11. 13.

Niehoff, B., Hirche, H.J., Båmstedt, U., 2000. The reproduction of *Calanus finmarchicus* in the Norwegian Sea in spring. Sarsia 85, 15-22.

Oshlack, A., Robinson, M. D., & Young, M. D. 2010. From RNA-seq reads to differential expression results. Genome biol, 11(12), 220.

Ortelli, F., Rossiter L.C., Vontas J., Ranson H., Hemingway J. 2003. Heterologous expression of four glutathione transferase genes genetically linked to a major insecticide-resistance locus from the malaria vector *Anopheles gambiae*. Biochem J. 373:957-63.

Oshima, Y. 1995. Postcolumn derivatization liquid chromatography method for paralytic shellfish toxins. J. AOAC Int. 78, 528-532.

Pal, R., Sanil N., Clark A. 2012. Developmental studies on the Sigma and Delta-1 glutathione transferases of *Lucilia cuprina*. Comp Biochem Phys D, Genomics Proteomics.7(1):28-34.

Petitpas, C.M., Turner, J.T., Deeds, J.R., Keafer, B.A., McGillicuddy, D.J. Jr., Milligan, P.J., Shue, V., White, K.D., Anderson, D.M., 2014. PSP toxin levels and plankton community composition and abundance in size-fractionated vertical profiles during spring/summer blooms of the toxic dinoflagellate *Alexandrium fundyense* in the Gulf of Maine and on Georges Bank, 2007, 2008, and 2010: 2. Plankton community composition and abundance. Deep-Sea Res. II 103, 350-367.

Pfaffl, M.W., Horgan, G.W., Dempfle, L. 2002. Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucleic Acids Research 30, 9.

Pfaffl, M.W., Tichopad, A., Prgomet, C., Neuvians, T.P. 2004. Determination of stable 589 housekeeping genes, differentially regulated target genes and sample integrity. 590 BestKeeper—Excel-based tool using pair-wise correlations. *Biotechnology Letters* 26, 509–515.

Planque, B., Hays, G.C., Ibanez, F., Gamble, J.C., 1997. Large scale spatial variations in the seasonal abundance of *Calanus finmarchicus*. Deep-Sea Res. I 44, 315-326.

Plourde,S., Runge, J.A., 1993. Reproduction of the planktonic copepod *Calanus finmarchicus* in the Lower St. Lawrence Estuary: relation to the cycle of phytoplankton production and evidence for a *Calanus* pump. Marine Ecology Progress Series 102, 217-227.

Podrabsky, J.E., Somero, G.N., 2004. Changes in gene expression associated with acclimation to constant temperatures and fluctuating daily temperatures in an annual killifish *Austrofundulus limnaeus*. J Exper. Bio.207(13), 2237-2254.

Poulet, S.,Ianora, A., Laabir, M., Breteler, W.K., 1995. Towards the measurement of secondary production and recruitment in copepods. ICES Journal of Maine

Poynton, H.C., Lazorchak, J.M., Impellitteri, C.A., Blalock, B.J., Rogers, K., Allen, H. J., & Govindasmawy, S., 2012. Toxicogenomic responses of nanotoxicity in *Daphnia magna* exposed to silver nitrate and coated silver nanoparticles. Environ. Sci. Technol., 46(11), 6288-6296.

Poynton, H.C., Varshavsky, J.R., Chang, B., Cavigiolio, G., Chan, S., Holman, P.S., et al. 2007. *Daphnia magna* ecotoxicogenomics provides mechanistic insights into metal toxicity. Environmental science & technology, 41(3), 1044-1050.

Prince, E.K., Lettieri, L., McCurdy, K.J., Kubanek, J., 2006. Fitness consequences for copepods feeding on a red tide dinoflagellate: deciphering the effects of nutritional value, toxicity, and feeding behavior. Oecologia, 147, 479–488.

Qin G., Jia M., Liu T., Zhang X., Guo Y., Zhu K.Y., et al. 2013. Characterization and functional analysis of four glutathione S-transferases from the migratory locust, *Locusta migratoria*. PloS One. 8(3):e58410.

Rajarapu S.P., Mittapalli O. 2013. Glutathione-S-transferase profiles in the emerald ash borer, *Agrilus planipennis*. Comp Biochem Phys B, Biochem Mol Biol. 2013;165(1):66-72.

Ranson H. and Hemingway J. 2005. Mosquito glutathione transferases. Methods Enzymol. 401:226-241.

Riesgo, A.,Andrade, S.C., Sharma, P.P., Novo, M., Pérez-Porro, A.R., Vahtera, V., González, V.L., Kawauchi, G.Y., Giribet, G., 2012. Comparative description of ten transcriptomes of newly sequenced invertebrates and efficiency estimation of genomic sampling in non-model taxa. Front. Zool. 9(1), 33-33.

Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26(1), 139-140.

Roff, D.A., 2002. Life history evolution. Sinauer Associates Sunderland.

Roff, D.A., 2007. Contributions of genomics to life-history theory. Nature Rev. Genetics 8(2), 116-125.

Roff, D.A., Fairbairn, D., 2007. The evolution of trade-offs: where are we? J Evol. Biol. 20(2), 433-447.

Roncalli V., Cieslak M.C., Passamaneck Y., Christie A.E., Lenz P.H. 2015. Glutathione S-Transferase (GST) Gene Diversity in the Crustacean *Calanus finmarchicus* – Contributors to Cellular Detoxification. PLoS ONE 10(5): e0123322. doi:10.1371/journal.pone.0123322.

Ronquist, F., Huelsenbeck J.P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics. 19(12):1572-4.

Rossini, G.P. 2014 Toxins and Biologically Active Compounds from Microalgae. CRC Press Italy.

Runge, J., Plourde, S., Joly, P., Niehoff, B., Durbin, E., 2006. Characteristics of egg production of the planktonic copepod, *Calanus finmarchicus*, on Georges Bank: 1994–1999. Deep-Sea Res. II 53, 2618-2631.

Rushmore, T.H., & Tony Kong, A. 2002. Pharmacogenomics, regulation and signaling pathways of phase I and II drug metabolizing enzymes. Current drug metabolism, 3(5), 481-490.

Saisawang C., Wongsantichon J., Ketterman A. J. 2012. A preliminary characterization of the cytosolic glutathione transferase proteome from *Drosophila melanogaster*. Biochem J. 442(1):181-90.

Saumweber, W.J., Durbin, E. G., 2006. Estimating potential diapause duration in *Calanus finmarchicus*. Deep-Sea Research II 53, 2597-2617.

Schoville, S.D., Barreto, F.S., Moy, G.W., Wolff, A., and Burton, R.S. 2012. Investigating the molecular basis of local adaptation to thermal stress: population differences in gene expression across the transcriptome of the copepod *Tigriopus californicus*. *BMC evolutionary biology*, *12*(1), 170.

Schultz, J., Milpetz F., Bork P., Ponting C.P. 1998. SMART, a simple modular architecture research tool: Identification of signaling domains. Proc Nat Acad Sci U.S.A. 95(11):5857-64.

Seixas, P., Coutinho, P., Ferreira, M., Otero, A., 2009. Nutritional value of the cryptophyte *Rhodomonas lens* for *Artemia* sp. Journal of Experimental Marine Biology and Ecology 381(1), 1-9.

Sheehan, D., Foley D.M. and Dowd C.A. 2001. Structure, function and evolution of glutathione

transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily. Biochem J. 360:1–16.

Sherman, K., Smith, W.G., Green J. R.,Cohen, E.B., Berman, M. S., Marti, K.A., Goulet, J.R , 1987. Zooplankton production and the fisheries of the northeastern shelf. In: R. H. Backus (ed.), Georges Bank. MIT Press, Cambridge, MA, pp. 268-282.

Shi, H., Pei L., Gu S., Zhu S., Wang Y., Zhang Y., Li B. 2012. Glutathione S-transferase (GST) genes in the red flour beetle, *Tribolium castaneum*, and comparative analysis with five additional insects. Genomics. 100(5):327-335.

Shumway, S.E., Sherman-Caswell, S., Hurst, J., 1988. Paralytic shellfish poisoning in Maine: monitoring a monster. Journal of Shellfish Research 7, 643-652.

Shumway, S.E., 1990. A review of the effects of algal blooms on shellfish and aquaculture. Journal of the World Aquaculture Society 21: 65-104.

Slater, P.J., Snowdon, C.T., Rosenblatt, J.S., Milinski, M., 1997. Advances in the Study of Behavior. Academic Press.

Somero, G.N., 2012. The physiology of global change: linking patterns to mechanisms. Ann. Rev. Mar. Sci. 4, 39-61.

Sopanen, S., Setälä, O., Piiparinen, j., Erler, K., Kremp, A., 2011. The toxic dinoflagellate *Alexandrium ostenfeldii* promotes incapacitation of the calanoid copepods *Eurytemora affinis* and *Acartia bifilosa* from the northern Baltic Sea. J. Plank. Res. 33, 1564-1573.

Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30(9):1312-3.

Starr, M., Runge, J.A., Therriault, J.C., 1999. Effects of diatom diets on the reproduction of the planktonic copepod *Calanus finmarchicus*. Sarsia 84, 379-389.

Stearns, S.C., 1989. The evolutionary significance of phenotypic plasticity. BioScience, 436-445.

Stillman, J.H.and Armstrong, E., 2015. Genomics Are Transforming Our Understanding of Responses to Climate Change. *BioScience*, biu219.

Stillman, J.H. and Hurt, D. A. 2015. Crustacean genomics and functional genomic responses to environmental stress and infection. Physiology, 4, 420.

Stoecker, D.K., Capuzzo J.M. 1990. Predation on protozoa- its importance to zooplankton. J Plankton Res. 12(5):891-908.

Sun, J., Liu, D., 2003. Geometric models for calculating cell biovolume and surface area for phytoplankton. Journal of Plankton Research 25(11), 1331-1346.

Tang, K.W., Jakobsen, H.H., Visser, A.W. 2001. *Phaeocystis globosa* (Prymnesiophyceae) and the planktonic food web: feeding, growth, and trophic interactions among grazers. Limnol. Oceanogr. 46, 1860–1870.

Tarrant, A.M., Baumgartner, M.F., Verslycke, T., Johnson, C.L., 2008. Differential gene expression in diapausing and active *Calanus finmarchicus* (Copepoda). Mar. Ecol. Prog. Ser. 355, 193–207.

Tarrant, A.M., Baumgartner M.F., Hansen B.H., Altin D., Nordtug T., Olsen A.J. 2014. Transcriptional profiling of reproductive development, lipid storage and molting throughout the last juvenile stage of the marine copepod *Calanus finmarchicus*. Front Zool. 11(1), 91.

Teegarden, G.J., Cembella, A.D., 1996. Grazing of toxic dinoflagellates, *Alexandrium* spp., by adult copepods of coastal Maine: implications for the fate of paralytic shellfish toxins in marine food webs. Journal of Experimental Marine Biology and Ecology 196, 145-176.

Teegarden, G.J., Campbell, R.G., Durbin, E. G., 2001. Zooplankton feeding behavior and particle selection in natural plankton assemblages containing toxic *Alexandrium* spp. Mar. Ecol. Progr. Ser. 218, 213-226.

Teegarden, G.J., Capuano C.L., Barron S.H. and Durbin E.G. 2003. Phycotoxin accumulation in zooplankton feeding on *Alexandrium fundyense* - vector or sink? J Plankton Res. 25(4):429-43.

Teegarden, G.J., 1999. Copepod grazing selection and particle discrimination on the basis of PSP toxin content. Mar. Ecol. Progr. Ser. 181, 163-176.

Teegarden, G.J., Campbell R.G., Anson D.T., Ouellett A., Westman B.A., Durbin E.G. 2008. Copepod feeding response to varying *Alexandrium* spp. cellular toxicity and cell concentration among natural plankton samples. Harmful Algae. 7(1):33-44.

Tillmann, U., Alpermann, T., John, U., Cembella, A., 2008. Allelochemical interactions and short-term effects of the dinoflagellate *Alexandrium* on selected photoautotrophic and heterotrophic protists. Harmful Algae 7, 52-64.

Tillmann, U.,Alpermann, T., John, U., Cembella, A., 2008. Allelochemical interactions and short-term effects of the dinoflagellate *Alexandrium* on selected photoautotrophic and heterotrophic protists. Harmful Algae 7, 52-64.

Tillmann, U., John, U., 2002. Toxic effects of *Alexandrium* spp. on heterotrophic dinoflagellates: an allelochemical defence mechanism independent of PSP-toxin content. Marine Ecology Progress Series 230, 47-58.

Turner, J.T., Anderson T. 1983. Zooplankton grazing during dinoflagellate blooms in a Cape Cod embayment, with observations of predation upon tintinnids by copepods. Mar Ecol. 4:359-74.

Turner, J.T., Tester, P.A., 1997. Toxic marine phytoplankton, zooplankton grazers, and pelagic food webs. Limnol. Oceanogr. 42(5part2), 1203-1213.

Turner, J.T., Borkman, D.G., 2005. Impact of zooplankton grazing on *Alexandrium* blooms in the offshore Gulf of Maine. Deep-Sea Res. II 52, 2801-2816.

Turner, J.T., Doucette, G. J., Keafer, B. A., Anderson, D. M., 2005. Trophic accumulation of PSP toxins in zooplanbkton during *Alexandrium fundyense* blooms in Casco Bay, Gulf of Maine, April-June 1998. II. Zooplankton abundance and size-fractionated community composition. Deep-Sea Research II 52, 2784-2800.

Turner, J.T., 2006. Harmful algae interactions with marine planktonic grazers. In: Granéli, E., Turner, J. T. (Eds.), Ecology of harmful algae. Springer-Verlag, pp. 259-270.

Turner, J.T., 2010. Zooplankton community grazing impact on a bloom of *Alexandrium fundyense* in the Gulf of Maine. Harmful Algae 9, 578-589.

Turner J.T. 2014. Planktonic marine copepods and harmful algae. Harmful Algae. 32:81-93.

Turriff, N., Runge, J. A., Cembella, A. D. 1995. Toxin accumulation and feeding behavior of the planktonic copepod *Calanus finmarchicus* exposed to the red-tide dinoflagellate *Alexandrium excavatum*. Mar. Biol. 123, 55-64.

Unal, E., Bucklin, A., 2010. Basin-scale population genetic structure of the planktonic copepod Calanus finmarchicus in the North Atlantic Ocean. Progr. Oceanogr. 87(1), 175-185.

Uye, S.,Takamatsu, K., 1990. Feeding interactions between planktonic copepods and red-tide flagellates from Japanese coastal waters. Marine Ecology Progress Series 59, 97-107.

Van Straalen, N.M., Roelofs, D., 2012. An introduction to ecological genomics. Oxford University Press.

Verity, P.G., Smetacek V. 1996. Organism life cycles, predation, and the structure of marine

pelagic ecosystems. Mar Ecol Progr Ser.130(1-3):277-93.

Waggett, R.J., Hardison, D.R., Tester, P.A., 2012. Toxicity and nutritional inadequacy of *Karenia brevis*: synergistic mechanisms disrupt top-down grazer control. Mar. Ecol. Prog. Ser. 444, 15–30.

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

Webster, T.M.U., and Santos, E.M. 2015. Global transcriptomic profiling demonstrates induction of oxidative stress and of compensatory cellular stress responses in brown trout exposed to glyphosate and Roundup. BMC genomics, 16(1), 32.

Whelan, S., Goldman N. 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol. 18 (5):691-9.

Wiese, M.,D'agostino, P.M., Mihali, T.K., Moffitt, M.C., Neilan, B.A., 2010. Neurotoxic alkaloids: saxitoxin and its analogs. Mar. Drugs 8(7), 2185-2211.

Willoughby, L., Chung, H., Lumb, C., Robin, C., Batterham, P., & Daborn, P. J.2006. A comparison of *Drosophila melanogaster* detoxification gene induction responses for six insecticides, caffeine and phenobarbital. Insect Biochem Mol Bio, 36(12), 934-942.

Xu, C.,Yong-Tao Li C. and Kong Ah-Ng T. 2005. Induction of phase I, II and III drug metabolism/transport by xenobiotics. Arch Pharm Res. 28(3):249-68.

Yang, Q.,Sun F., Yang Z., Li H. 2014. Comprehensive transcriptome study to develop molecular resources of the copepod *Calanus sinicus* for their potential ecological applications. Biomed Res Int. 2014:493825.

Zhou, W.W., Liang Q.M., Xu Y., Gurr G.M., Bao Y.Y., Zhou X.P., et al. 2013. Genomic

insights into the glutathione S-transferase gene family of two rice planthoppers, *Nilaparvata lugens* (Stal) and *Sogatella furcifera* (Horvath) (Hemiptera: Delphacidae). PloS One. 8(2):e56604.

Zinke, I., Schütz, C.S., Katzenberger, J.D., Bauer, M., Pankratz, M.J., 2002. Nutrient control of gene expression in *Drosophila*: microarray analysis of starvation and sugar-dependent response. EMBO J 21(22), 6162-617**.**