Control of toxigenic dinoflagellates through parasitism: Implications for host-parasite coevolution

von Yameng Lu

University of Bremen

2016

ALFRED-WEGENER-INSTITUT
HELMHOLTZ-ZENTRUM FÜR POLAR- UND MEERESFORSCHUNG

Universität Bremen

Control of toxigenic dinoflagellates through parasitism: Implications for host-parasite coevolution

Dissertation

in partial fulfillment of the requirements for

Doktor der Naturwissenschaften

-Dr. rer. Nat.-

in the Faculty of Biology and Chemistry (Fachbereich 2) University of Bremen,

Germany

Yameng Lu

2016

Gutachter 1. Prof. Dr. Allan Cembella

2. Prof. Dr. Tilmann Harder

Table of Contents

Aknowledgements

I would like to express my gratitude to Prof. Allan Cembella for giving me the opportunity to work in his lab and to be involved in this very exciting project. His guidance and constant inspiration lead me into the science world. I am also grateful to Uwe John for his supervision and great support through the course of this challenging project. His enthusiasm for new ideas is always contagious and a source of motivation. His understanding and encouragement helped me to overcome difficulties during my study. I would also like to thank my committee members, especially Prof. Tilmann Harder for reviewing this thesis.

I acknowledge the help and support I received from Prof. Stephan Frickenhaus from AWI, Prof. Gernot Glöckner from University of Cologne, Laure Guillou from Biological Station Roscoff (France), and Marco Groth from Jena, the co-authors of each of the manuscripts presented herein. My sincere thanks go to Sylke Wohlrab, for her kind comments and so many discussions.

My most sincere gratitude and appreciation go to the AG Cembella family (Urban Tillmann, Bernd Krock, Annegret Müller, Nancy Kühne, Wolfgang Drebing, Jessica Kegel, Stephanie Westphal, Gerrit Rohner, Carla Flöthe, Julia Busch, Haiyan Ma, Jennifer Hülskötter, Malte Braltmaler). Not only have I received tremendous support from everyone during my PhD study, but also I had the opportunity to know and spend time with the most amazing individuals I have ever met, making this one of the most impressionable life experiences to date!

Finally, I am deeply grateful to my family. My motivating son Ziyuan without whom this thesis would have been completed early, provides me with everyday joy. My beloved husband Chen and parents, without whom this thesis would not have been possible, give me undiminished courage to pursue my dreams.

Summary

Dinoflagellates are among the most important primary producers in the ocean and represent highly diverse life forms. This group practices a wide variety of alternative nutritional modes; there are phototrophic and heterotrophic, as well as mixotrophic, free-living forms, and the obligated symbiotic and parasitic members. Two major lineages of dinoflagellates have been defined - the core and basal groups, the latter including the obligated parasitic syndinean dinoflagellates. Several dinoflagellates belonging to the core dinoflagellates produce dangerous toxins. Among these, *Alexandrium fundyense* is one of the most prominent harmful algal bloom-forming genera with its negative impact on the ecosystem, aquaculture seafood and causing serious hazard to human health. Increasing evidence shows that *Alexandrium* populations can be affected by parasitic attack by the basal syndinean dinoflagellate *Amoebophrya*. However, regulatory mechanisms of the infection of core dinoflagellates by their parasites are largely unknown. The aim of the thesis was to provide insights into the infection dynamics among *Alexandrium* and its parasite *Amoebophrya*, and to better understand the infection processes with implications for host-parasite coevolution.

To investigate the susceptibility of the dinoflagellate to infection by the parasite on an intra-specific level, different populations of the host *Alexandrium* from very distant geographical origins (Alaska, the Gulf of Maine and the North Sea) were provided to the parasite *Amoebophrya*. There was a strong negative effect of parasitism on the development of host populations, but no apparent adaptation of the host *Alexandrium* was observed. Cellular toxin contents were examined, showing that neither toxin concentration nor composition changed within each geographical population. Therefore, the results indicated that the host *Alexandrium* likely does not use toxins as a potential defense strategy against the parasite.

In this thesis, a whole genome sequencing of the parasite *Amoebophrya* was performed for the first time and a transcriptomic dataset from the infection cycle of this

parasite-host system was generated. The basal dinoflagellate *Amoebophrya* has a relatively small genome in size around 90 Mbp. Besides the reduction of genome size, several parasitic features were observed in the genome including loss of duplicated genes and function loss (e.g. inability to generate certain amino acids) that indicates the parasite dependent on the host. Notably, the genome also exhibits novel features. The shikimate and tryptophan synthesis pathways are physically linked that may constitute an unknown mechanism of pathway regulation. Mitochondria are observed, but the mitochondrial genome is completely lost in *Amoebophrya*.

The established cDNA library (>900,000 reads/313 Mbp) consists of 14,455 ESTs. Differentially expressed genes point to general mechanisms in host-parasite recognition and infection. Particular surface lectins are expressed in the parasite *Amoebophrya* at early infection processes, and these lectins likely mediate the attachment to the host cell, followed by processes involved in host recognition, adhesion, and invasion. During maturation, cell division and proliferation related genes reflect fast cell growth of the parasite. These findings indicate the presence of fundamental processes that have remained stable throughout evolution.

By contrast, the host *Alexandrium* reacts differently towards parasite infection and respective parasitic waterborne cues, but both treatments exhibited significant changes in gene expression associated with specific metabolic pathways. A total of 14,882 *Alexandrium* genes were differentially expressed over the whole-parasite infection cycle at three different time points (0, 6 and 96 h). The results from RNA sequencing analyses indicate that parasite infection increases the energy demand of the host, as a large amount of genes involved in photosynthesis, ATP synthesis through glycolysis and fatty acid production were upregulated. The stimulation of signal transduction chains by waterborne cues from the parasite alone could prime the host's defense or induce host's adaptive responses to the parasite activity.

Zusammenfassung

In den Ozeanen sind Dinoflagellaten wichtige und diverse Primärproduzenten. Sie können phototroph, heterotroph oder mixotroph sein und eine symbiontische oder parasitäre Lebensweise aufweisen. Dinoflagellaten lassen sich in zwei Abstammungslinien eingruppieren: die Kerngruppe und die basale Gruppe, welche auch Arten der Ordnung Syndiniales enthält. Einige der Dinoflagellaten der Kerngruppe produzieren gefährliche Toxine. Hierbei ist *Alexandrium* eine der häufigsten Gattungen die schädliche Algenblüten bildet, welche durch die Toxine einen negativen Einfluss auf Ökosysteme und Aquakulturen haben sowie die menschliche Gesundheit gefährden. Es gibt Hinweise darauf, dass *Alexandrium* Populationen durch den basalen Dinoflagellaten *Amoebophrya*, der den Syndiales angehört, kontrolliert werden können. Trotzdem sind regulatorische Prozesse im Infektionszyklus von Kern-Dinoflagellaten durch ihre Parasiten eher unbekannt. Das Ziel dieser Arbeit ist es, Einblicke in die Infektionsdynamiken zwischen *Alexandrium* und seinem Parasiten *Amoebophrya* zu erlangen und die Infektionsprozesse im Hinblick auf Wirt-Parasit-Koevolution besser zu verstehen.

Um die intraspezifische Empfindlichkeit der Dinoflagellaten auf Infektionen durch den Parasiten zu untersuchen wurden verschiedene *Alexandrium*-Populationen von weit auseinanderliegenden geographischen Positionen (Alaska, Golf von Maine und die Nordsee) für eine Infektion durch *Amoebophrya* isoliert. Die Entwicklung der Wirtspopulationen wurde vom Parasitismus sehr negativ beeinflusst, aber eine Adaptation des Wirtes an die Infektionen konnte nicht bestätigt werden. Hierbei gab es weder bei der Toxinkonzentration noch bei der Zusammensetzung verschiedener Toxine einen Unterschied zwischen den Populationen aus verschiedenen geographischen Positionen. Daher wird vermutet, dass der Wirt *Alexandrium* keine Toxine als Abwehrmechanismus gegen den Parasiten verwendet.

In dieser Arbeit wurde zum ersten Mal das komplette Genom des Parasiten *Amoebophrya* sequenziert und charakterisiert. Des Weiteren wurde ein Transkriptomdatensatz über den Infektionszyklus des Wirt-Parasit-Systems erstellt. *Amoebophrya* hat ein relativ kleines Genom mit einer Größe um die 90 Mbp. Neben der Reduktion der Genomgröße wurden weitere parasitäre Eigenschaften festgestellt, wie der Verlust von Genduplikationen

und der Funktionsverlust von Genen (z.B. die Unfähigkeit, bestimmte Aminosäuren zu produzieren) welche darauf hinweisen, dass der Parasit abhängig vom Wirt ist. Bemerkenswert ist die Existenz von neuartigen Eigenschaften im Genom. Die enge Verbindung des Shikimatweges und der Tryptophansynthese weisen auf einen unbekannten Mechanismus der Stoffwechselregulierung hin. Strukturen von Mitochondrien können mikroskopisch in *Amoebophrya* beobachtet werden, aber das Mitochondrien-Genom ist komplett verloren gegangen.

Die neu etablierte cDNA library (>900 000 reads/313 Mbp) besteht aus 14.455 ESTs. Unterschiedlich exprimierte Gene deuten auf generelle Mechanismen für die Wirt-Parasit-Erkennung und Infektion hin. Oberflächenlektine sind im Parasiten *Amoebophrya* besonders während der frühen Infektionsprozesse exprimiert, welche vermutlich die Anheftung an die Wirtszelle regulieren. Dieser Prozess wird von Wirtserkennung, Adhäsion und Invasion in den Wirt gefolgt. Gene, die mit der Reifung, Zellteilung und Proliferation in Verbindung stehen, reflektieren ein schnelles Zellwachstum des Parasiten. Diese Resultate deuten auf fundamentale Prozesse hin, die sich während der Evolution stabil gehalten haben.

Im Gegensatz dazu reagiert der Wirt *Alexandrium* anders auf parasitäre Infektionen und entsprechende durch das Wasser übertragene Signale. Beide Behandlungen resultierten in signifikanten Unterschieden in der Genexpression, welche mit spezifischen metabolischen Stoffwechselwegen assoziiert werden. Insgesamt wurden 14.882 Gene von *Alexandrium* unterschiedlich an drei gemessenen Zeitpunkten des Infektionszyklus‘ (nach 0 h, 6 h und 96 h) exprimiert. Die Resultate der RNA-Sequenzierungen zeigen, dass parasitäre Infektionen den Energiebedarf des Wirtes erhöhen, da ein großer Teil der Gene in Photosynthese, ATP-Synthese durch Glykolyse und Fettsäureproduktion hochreguliert waren. Die Stimulation der Signaltransduktionskette nur durch wasserübertragene Signale des Parasiten konnten die Abwehrmechanismen oder adaptive Reaktionen des Wirtes gegenüber parasitärer Aktivität fördern.

Abbreviations

CoA	coenzyme A
ANOVA	analysis of variance
ATP	adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
cDNA	complementary DNA
cob	cytochrome b
cox1	cytochrome oxidase subunit I
cox3	cytochrome oxidase subunit III
CSD	cold shock protein
DGDG	di-galactosyldiacylglycerol
DNA	deoxyribonucleic acid
DSP	diarrhetic shellfish poisoning
ESTs	expressed sequence tags
EtOH	ethyl alcohol
FAS	fatty acid synthases
GO	Gene Ontology
GPCR	G-protein coupled receptors
GTX	gonyautoxins
HABs	harmful algal blooms
HemB	porphobilinogen synthase
HemC	porphobilinogen deaminase
HemD	uroporphyrinogen III synthase
HemE	uroporphyrinogen III decarboxylase
HemF	coproporphyrinogen oxidase
HLPs	histone-like proteins
HSP	heat shock protein
JRL	jacalin-like lectin
KEGG	The Kyoto Encyclopedia of Genes and Genomes
KO	KEGG Orthology
KOGs	eukaryotic orthologous groups
LC-FD	liquid chromatography with fluorescence detection
LSU	large ribosomal subunit
MALV II	marine alveolate group II
MAPK	mitogen-activated protein kinase
MCMC	Markov chain Monte Carlo
MGDG	mono-galactosyldiacylglycerol

mRNA	messenger ribonucleic acid
NCBI	National Center for Biotechnology information
NGS	next-generation sequencing
PCR	polymerase chain reaction
PDH	pyruvate dehydrogenase
PKS	polyketide synthases
PSP	paralytic shellfish poisoning
PSTs	paralytic shellfish toxins
RAS	Rhizarians-Alveolates-Stramenopiles
rDNA	ribosomal DNA
RNA-seq	RNA sequencing
ROIs	reactive oxygen intermediates
ROS	reactive oxygen species
rRNA	ribosomal ribonucleic acid
SL	trans-spliced leader
SOD	superoxide dismutases
ss-cDNA	single-stranded complementary DNA
SSU rDNA	small subunit ribosomal RNA
STX	saxitoxin
TCA	tricarboxylic acid
TCA cycle	citrate cycle
THF	tetrahydrofuran
tRNA	transfer ribonucleic acid
Uniprot	Universal Protein Resource

1. Introduction

1.1 Evolution and adaptation of parasites

A parasite is defined as a typically small organism that exploits another organism, called their host, as a food source and as a habitat. Unlike a predator, a parasite does not have to kill its host, is generally smaller than its host and lives in or on its host for a certain period (Kuris 1974). In contrast, the parasitoid relationship differs from parasitism in that a parasitoid infects only one host during its lifetime, suppresses further host division and the host is inevitably killed to complete the parasitoid's life cycle, whereas a parasite has effects on the host fitness, influencing its viability only indirectly (Kuris 1974; Park *et al.* 2004).

When the first study of evolutionary biology emerged in the 19th century, parasites were seen as biologically degenerate and contrary to nature (Price 1980). In fact, parasites have evolved from free-living ancestors and obviously undergone important evolutionary changes since transition to a parasitic way of life (Poulin 2011). Parasites are abundant and ubiquitous throughout evolutionary history (Morris 1981), and the number of species of parasites are even more numerous than organisms with a non-parasitic lifestyle (Dobson *et al.* 2008; Jackson 2015). It is also widely appreciated that parasites are prone to rapid evolution and may evolve more rapidly than their host because of their short generation time and co-evolutionary constraints (Kochin *et al.* 2010; Thompson 1998). Hosts and parasites exert reciprocal selective pressures, which may lead to a continuous adaptation and induce evolutionary responses in both parties (Thompson 1998, 1999).

Co-evolution is defined as the process of reciprocal, adaptive genetic change in two or more species, and may occur between any interacting populations, for instance prey and predator, plant and herbivore, host and pathogen, etc (Woolhouse *et al.* 2002). Hosts can select for enhanced infectivity of parasites, while parasites should therefore be expected to induce selection against host resistance. This kind of selection results in co-evolutionary changes governing both host resistance and parasite infectivity (Anderson & May 1982;

Woolhouse *et al.* 2002). Co-evolution may be described by two alternative types of dynamic mechanisms: a persistant "arms race" (accumulated offensive/defensive 'improvements' in both populations) that results in rapid evolution of the genes involved but yields low levels of standing genetic variation (Fig. 1.1a); or the "Red Queen" strategy ("running as fast as you can to stay in the same place" as in the quote from Alice in Wonderland [Lewis Carroll] (Van Valen 1973). The latter strategy, also known as "fluctuating selection dynamics") leads to cyclic dynamics of shifting allele frequencies that results in frequency-dependent selection for rare host and parasite genotypes (Fig. 1.1b) (Anderson & May 1982; Bergelson *et al.* 2001; Thompson 1994; Woolhouse *et al.* 2002). Evidence for co-evolutionary dynamics from natural plankton populations is limited. Nonetheless, recent cross-infections in culture revealed a high potential for *Red Queen* dynamics between the dinoflagellate host *Alexandrium minutum* and parasite *Parvilucifera* (Råberg *et al.* 2014).

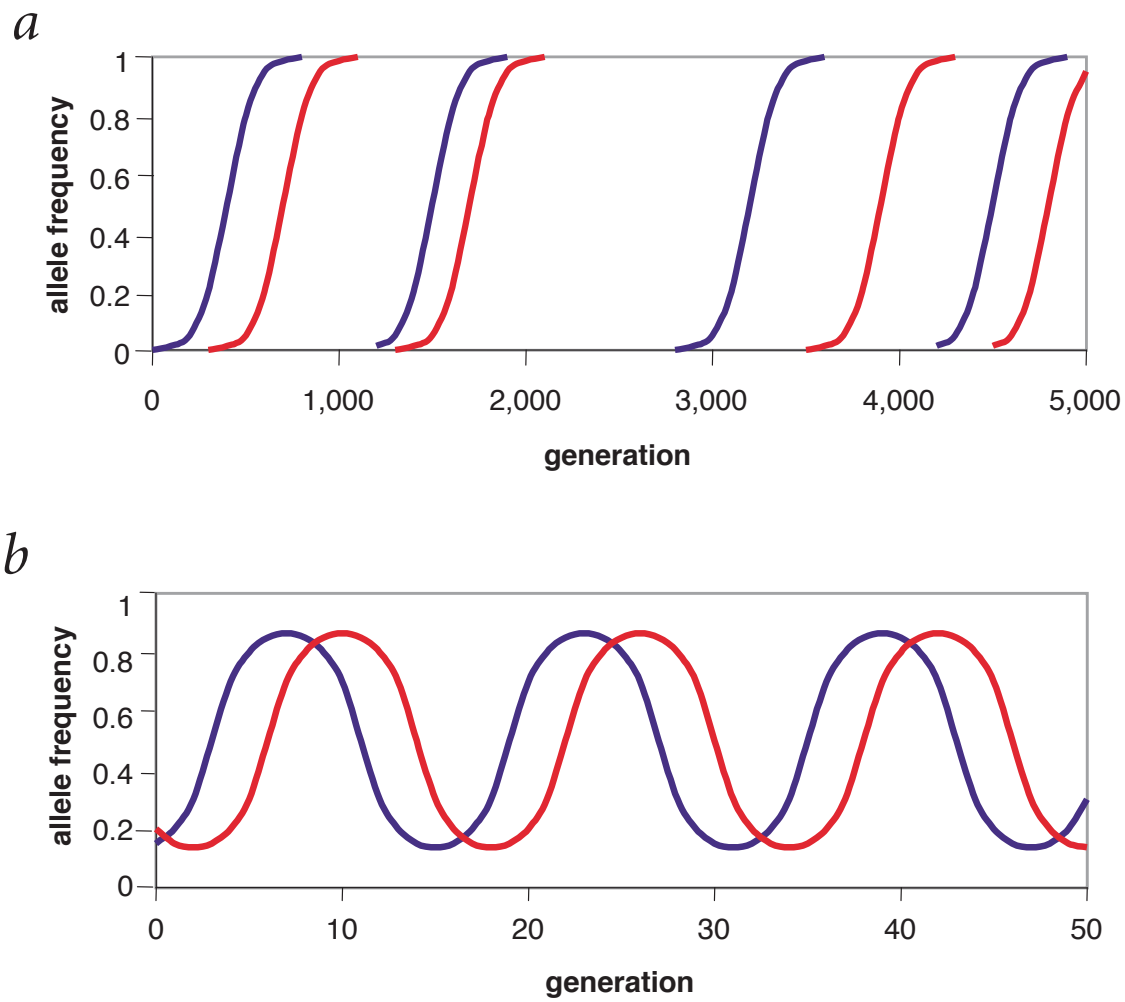


Fig. 1.1: Allele frequency changes driven by co-evolution. a) "arms race" dynamics: a series of selective sweeps by host (blue) and pathogen (red) alleles derived by mutation. b) "Red Queen" dynamics: dynamic polymorphisms in both host (blue) and pathogen (red) acting on existing genetic variation (Woolhouse *et al.* 2002).

Phytoplankton are a diverse primarily unicellular photosynthetic organisms that tend to drift with the currents in marine and fresh waters, although many also have independent motility (Falkowski *et al.* 2004). In terms of numbers of described species, and also biomass contribution to primary production, the major groups of phytoplankton include the diatoms, haptophytes and dinoflagellates. After diatoms, dinoflagellates are the most notable eukaryotic primary producers, especially in marine environments (Anderson *et al.* 2012a). Dinoflagellates display a high diversity in alternative nutritional modes; there are

phototrophic and heterotrophic, as well as mixotrophic, symbiotic and parasitic forms (Hackett *et al.* 2004; Jeong *et al.* 2010; Jeong *et al.* 2005; Miller *et al.* 2012). Dinoflagellates are also known for a wide diversity of alternative life histories upon or within host habitats, although most marine species are free-living and motile in the vegetative stage (Anderson 1998). Symbiotic dinoflagellates such as *Symbiodinium* are essential to reef-building corals (Coffroth & Santos 2005), whereas *Hematodinium* species parasitize crustaceans causing significant damage to commercial fisheries and wild fish stocks (Stentiford & Shields 2005).

Parasites of marine protists were first described in the 19th century, however, their importance was not widely recognised until the late 20th century, when protist parasites were reported to cause epizootics in fishes and invertebrate populations (Chatton 1912; Harvell *et al.* 1999). Accumulating evidence suggests that protist parasites potentially affect a wide range of marine planktonic organisms, ranging from other protists to larger planktonic invertebrates, reviewed by Skovgaard (2014). For instance, protist parasites were found to infect multicellular zooplankton, such as the parasite dinoflagellate *Syndinium* infecting copepods and the parasite ciliate *Collinia* infecting euphasids (Skovgaard 2014; Skovgaard *et al.* 2005). Furthermore, increasing prevalence data indicate that protist parasites play important roles in regulating host populations. Large-scale infections of the parasitic nanoflagellate *Pirsonia* spp. were found in diatom species in coastal waters (Tillmann *et al.* 1999). This parasite is a potentially important degrader of centric diatoms too large to be grazed by many copepods (Kühn 1998).

1.2 Dinoflagellate phylogeny and evolution

Dinoflagellates belong to the superphylum Alveolata, together with apicomplexans and ciliates, in the RAS (abbreviation for Rhizarians-Alveolates-Stramenopiles; alternatively spelled “SAR”) group (Burki *et al.* 2007; Keeling 2013). Apicomplexans are mostly intracellular parasites (exemplified by the malaria parasite *Plasmodium falciparum*) and contain an apicoplast (a non-photosynthetic plastid) (Waller & McFadden 2005), but within

the alveolates, the dinoflagellates and apicomplexans are considered to be more closely related than to ciliates (Fast *et al.* 2002). The ciliates are unicellular heterotrophs notable for their unusual cell biology including nuclear dimorphism and deviation from the universal genetic code (Tourancheau *et al.* 1995). Alveolate species that do not fall within these three phyla are important for inferring ancestral conditions. For example, the non-photosynthetic oyster pathogen *Perkinsus marinus* is sister to dinoflagellates and retains many typical eukaryotic characteristics (Reece *et al.* 1997), whereas the free-living heterotroph *Oxyrrhis marina* is placed just outside of the dinoflagellates, but deviates from the typical eukaryotic state and represents an intermediate dinokaryotic state (Lowe *et al.* 2010; Sano & Kato 2009).

In the present study two major lineages of dinoflagellates were considered (Fig. 1.2): the core and basal syndinean dinoflagellates (Bachvaroff *et al.* 2014; Okamoto *et al.* 2012). Nuclear characters were used to divide these into two major clades:

I. Syndinean dinoflagellates (e.g., parasite *Amoebophrya*) have low chromosome numbers and no obvious gene amplification.
II. Core dinoflagellates (e.g., host *Alexandrium*), formally called the dinokaryotes, share aberrant nuclear characters, including high DNA content, numerous chromosomes condensed during interphase, large-scale gene duplication.

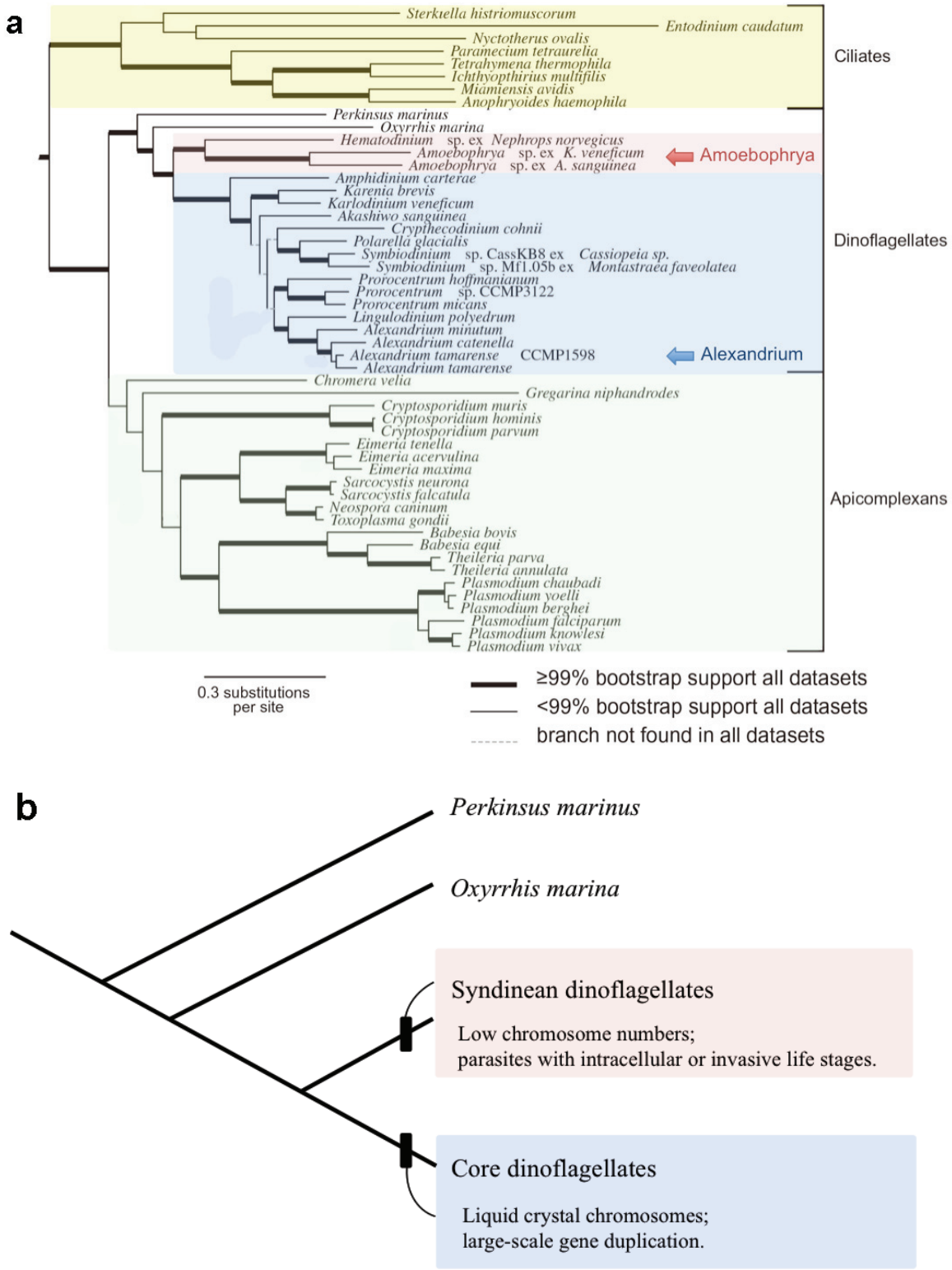


Fig. 1.2: Phylogeny of Alveolates. (a) Phylogenetic relationships between dinoflagellates, apicomplexans and ciliates. (b) Putative evolutionary scheme for the evolution of dinoflagellate characters, based on 73 ribosomal protein-coding genes. The tree is based upon genes found by RAxML using the LG amino acid substitution matrix with gamma correction and 500 rapid bootstraps, according to Bachvaroff *et al.* (2014). Bold branches were found in 100% of bootstrap replicates.

1.3 Organization and structure of the dinoflagellate genome

Dinoflagellates are typically unicellular organisms, although some are also found as chain-forming or colonial forms. The genome is often extremely large (Lin 2011), representing 1- to 80-fold the base-pair complement of the human haploid genome, and is divided among a multitude of chromosomes (Hackett *et al.* 2004). The positive correlation between cell size and genome size suggests that larger celled species, such as *Akashiwo sanguinea*, may have genomes approaching a terabase, but this has not yet been confirmed (Menden-Deuer & Lessard 2000).

Our understanding of the evolution of parasitism was first revolutionized by phylogenetics and now by genomics. Next-generation sequencing (NGS) technologies have finally placed complete genome sequence of dinoflagellates within reach and are helping to elucidate the biology and evolution of these enigmatic organisms (Wisecaver & Hackett 2011).

Early sequencing efforts have focussed on the dinoflagellate with reduced genomes, such as members of the genus *Symbiodinium* that are essential photosynthetic endosymbionts in coral reefs (Baker 2003). A draft assembly of the relatively small genome (1.5-Gbp) of *Symbiodinium minutum* was published in 2013, yet more than half the genome remains unassembled due to the presence of many highly repetitive sequences (Shoguchi *et al.* 2013). A more recent effort with *S. kawagutii* was more successful in that assembly of 0.935-Gbp of the 1.18-Gbp genome was achieved, reconstructing nearly 80% of the genome (Lin *et al.* 2015).

The extreme size of typical free-living dinoflagellate genomes has made them challenging to study, but what little is known suggests that they differ from well-defined eukaryotic model organisms, such as the brassicacean plant *Arabidopsis* and the free-living nematode *Caenorhabditis*, in many fundamental ways (Lin 2011; Wisecaver & Hackett 2011). DNA in core dinoflagellates self-assembles into liquid crystalline state, and chromosomes are

permanently condensed throughout the cell cycle, even through cellular division (Bouligand & Norris 2001; Gautier *et al.* 1986). The interior of the chromosomes in the nucleus is likely too dense to allow transcription, which therefore occurs on peripheral loops that extend from the condensed liquid-crystalline chromosomes (Chan & Wong 2007; Sala-Rovira *et al.* 1991). Histone-like proteins (HLPs) are associated with these loops and are involved in regulation of gene expression (Chan & Wong 2007). Another feature of dinoflagellates is that the genome contains unusual bases with a large portion of methylation, e.g., 12–70% of thymine is replaced by hydroxymethyluracil (Steele & Rae 1980). This unusual base was found in eukaryotes only as the result of oxidative damage of thymine or 5-methylcytosine and is quickly repaired by a DNA glycosylase (Boorstein *et al.* 1989). Core dinoflagellates all show this dinokaryotic nuclear structure; by contrast, *Oxyrrhis* appears to have HLPs but lacks permanently condensed chromosomes (Wisecaver & Hackett 2011). The structure of the nucleus in the Syndiniales has not been thoroughly investigated and is less clear.

Another peculiarity of dinoflagellate genomes is the presence of an invariant 22 bp trans-spliced leader (SL) found at the 5' end of full-length mRNAs (Lidie & van Dolah 2007; Zhang *et al.* 2009; Zhang *et al.* 2007). This trans-splicing has been found throughout diverse dinoflagellates, including in both core and syndinean dinoflagellates, and also in *Oxyrrhis* and in *Perkinsus*, suggesting that this processing arose early in dinoflagellate evolution (Bachvaroff *et al.* 2009; Jaeckisch *et al.* 2011; Joseph *et al.* 2010; Zhang & Lin 2008). The exact function of this mechanism in dinoflagellates is unknown but it may be involved in the resolution of polycistronic mRNAs (Palenchar & Bellofatto 2006) and in mRNA stability or translatability (Lukeš *et al.* 2009; Maroney *et al.* 1995; Satou *et al.* 2006). Recently, a further discovery of relict SL-sequences in tandem repeats has emerged, occurring in cDNAs as well as in the genomic DNA of dinoflagellates (Slamovits & Keeling 2008b). The relict SL-sequences are truncated after nucleotide 7 of the canonical spliced-leader, corresponding to an AG dinucleotide. This pattern indicates that expressed and trans-spliced genes are reverse-transcribed and reintegrated into the genome where they can undergo the next cycle of expression, trans-splicing and reintegration (Jaeckisch *et al.* 2011; Slamovits & Keeling 2008b).

Dinoflagellates also possess many highly duplicated genes that are often arranged into tandem arrays with fairly short intergenic sequences (Bachvaroff & Place 2008; Beauchemin *et al.* 2012; Lin *et al.* 2015; Liu & Hastings 2006; Mendez *et al.* 2015). Bachvaroff *et al.* (2008) determined the genomic structure of 47 genes from *Amphidinium carterae* that varied in their expression level as inferred from their frequency in the cDNA library. More highly expressed genes tended to be found in large tandem gene arrays, whereas genes expressed at a lower level appeared to be encoded by only a single gene and to contain more introns. It is therefore possible that large dinoflagellate genomes are unusually gene-rich for their size, and many of the genes are highly duplicated within the genome. This indicates the greatly complications of genome assembly based upon sequencing from short-reads; high numbers of long-reads with low error rates are required rather than simply large amounts of sequencing data (Wetzel *et al.* 2011).

The genomes of both the mitochondrion and the plastid in alveolates and particularly in dinoflagellates are highly reduced and uniquely organized. The mitochondrial genomes of dinoflagellates and apicomplexans carry only three protein-coding genes (*cob* [*cytochrome b*], *cox1* [*cytochrome oxidase subunit I*] and *cox3* [*cytochrome oxidase subunit III*]), two highly fragmented rRNAs and no tRNAs (Jackson *et al.* 2007; Kamikawa *et al.* 2009; Nash *et al.* 2007; Vaidya & Mather 2009; Waller & Jackson 2009). The *cob* and *cox3* genes are fused as a result of reducing the protein-coding genes into only two in the basal dinoflagellate *Oxyrrhis* (Slamovits *et al.* 2007). In contrast, the ciliates have linear mitochondrial genomes of around 40 kb and contain two rRNAs, seven tRNAs and approximately 50 genes (Burger *et al.* 2000; Gray *et al.* 2004; Pritchard *et al.* 1990). Like the nuclear genome, the mitochondrial genome in dinoflagellates is highly duplicated and the genes are fragmented and rearranged with many aberrant transcripts (Jackson *et al.* 2007; Waller & Jackson 2009). Recently, the gene for nuclear-encoded mitochondrial polymerase of the dinoflagellate *Heterocapsa triquetra* was reported, indicating a requirement for novel accessory factors to ensure the production of functional mRNAs in conjunction with the degenerate nature of the mitochondrial genome in dinoflagellates (Teng *et al.* 2013). Still, because of their large

amounts of inverted repeats, no full-length DNA has been described from the mitochondrial genome. Shoguchi (2015) reported the assembled mitochondrial genome (~326 kb) of the dinoflagellate, *Symbiodinium minutum,* and compared transcriptome between dinoflagellates and malarial parasite, *Plasmodium falciparum.* Small RNAs and noncoding sequences showed similarities and conservation between *S. minutum* and *P. falciparum.*

Plastid genomes in most photosynthetic eukaryotes are single circular chromosomes (~150 kb in length) derived from the genome of a cyanobacterial endosymbiont, containing approximately 100 genes (Lin 2011; Wisecaver & Hackett 2011). In peridinin-containing dinoflagellate plastids, the genome is fragmented into 2- to 3-kb minicircles encoding one to four genes and a non-coding core sequence thought to contain the origin of replication, making it the most reduced plastid genome (Howe *et al.* 2008; Koumandou *et al.* 2004; Zhang *et al.* 1999). Recently, for the first time, complete secondary loss of the plastid organelle, with retention of only few gene of plastid origin, was described with from the dinoflagellate *Hematodinium*, a parasite of marine crustaceans (Gornik *et al.* 2015). The reasons for an extreme reduction of the plastid genome in the dinoflagellates is unclear, but it was shown that this occurred early in their evolutionary history (Saldarriaga *et al.* 2001).

1.4 The dinoflagellate parasite *Amoebophrya*: phylogeny, life cycle and ecology

The Amoebophryidae belong to the Syndiniales (Alveolata), with only one known genus, *Amoebophrya*, but this genus exhibits a high genetic diversity (Alves-de-Souza *et al.* 2012). Corresponding environmental sequences (18S rDNA clone libraries) belonging to Syndiniales cluster into the widespread marine alveolate group II (MALV II), found mainly in the picoplankton size-fraction (< 2 or < 3 μm) (Guillou *et al.* 2008) (Fig. 1.3).

Early studies of *Amoebophrya* infectivity and host specificity suggested that these parasites infect a wide range of host taxa indiscriminately (Coats 1999). However, phylogenetic and culture studies indicated marked host specificity for *Amoebophrya* strains in culture (Coats & Park 2002; Janson *et al.* 2000; Kim & Kim 2007) and significant divergence

in SSU rDNA sequences across strains from different host species (Gunderson *et al.* 2002). *Amoebophrya* strains able to infect more than one host species show a substantially reduced infection success and frequently loose their infectivity over a few generations in alternative hosts (Kim 2006). *Amoebophrya* is now viewed as a species complex consisting of multiple host-specific parasites.

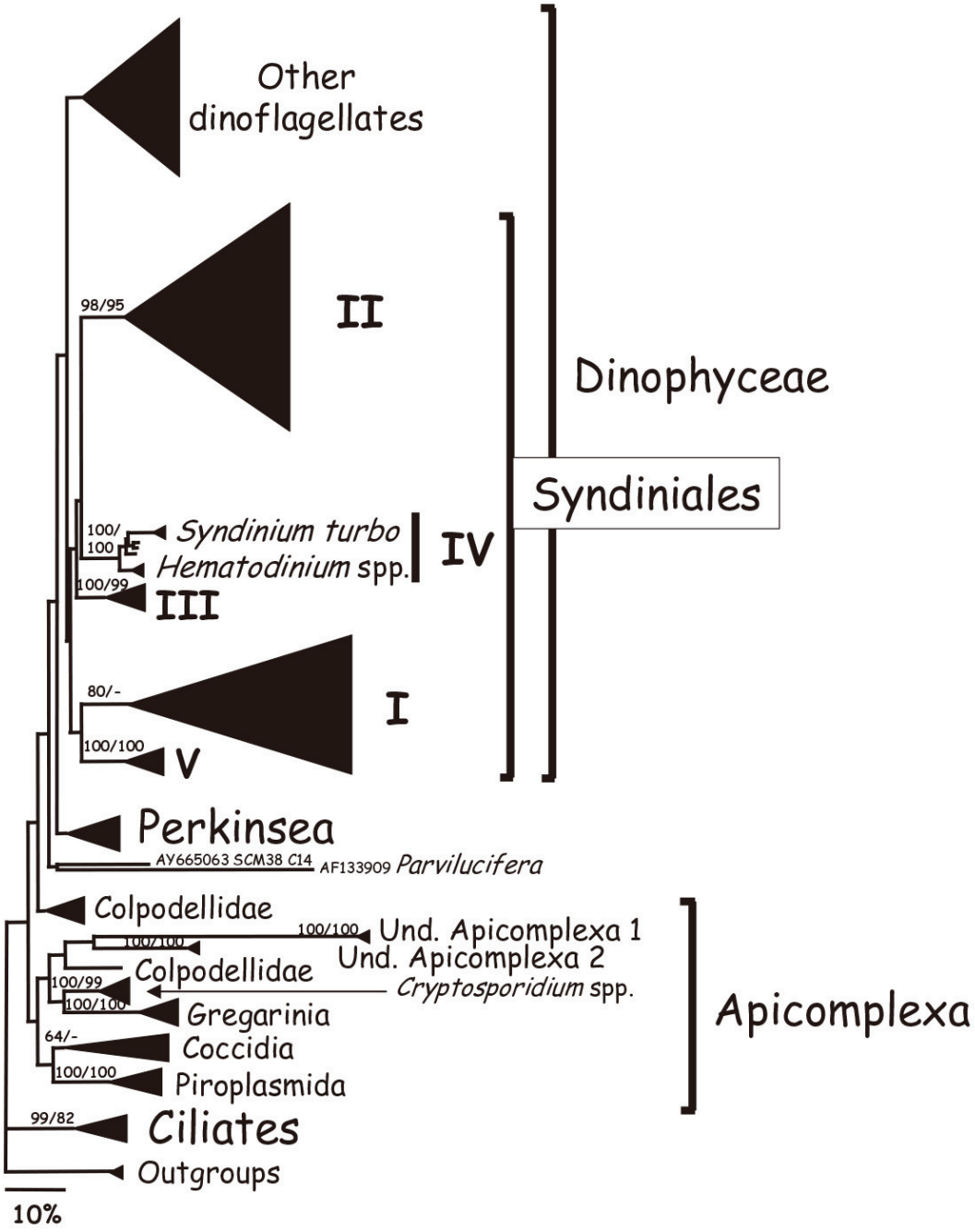


Fig. 1.3: Bayesian phylogeny of alveolates based on analysis of 291 near full-length 18S rRNA gene sequences (Guillou *et al.* 2008).

Amoebophrya parasites have simple direct life-cycles and can be co-cultured with their hosts, which makes them ideal pathogens for exploring host-parasite interactions. An *Amoebophrya* life-cycle taking up to four days for completion is depicted in Fig. 1.4. Infection by *Amoebophrya* is initiated by penetration of the parasitic dinospores into the host cells (Cachon 1964; Miller *et al.* 2012). Once inside the cytoplasm or nucleus (depending on the specific host and parasitic strains), the parasite starts to feed (the trophont stage). The trophont increases in size until sequential nuclear divisions and flagellar replications ultimately form an intracellular and multicellular 'beehive' stage inside the cytoplasm or nucleus of the host cell (the sporocyte). The mature sporocyte ruptures the cell wall of the host, and most develop into a short-lived vermiform stage that soon divides into numerous free-living infectious dinospores (Coats & Bockstahler 1994; Coats & Park 2002).

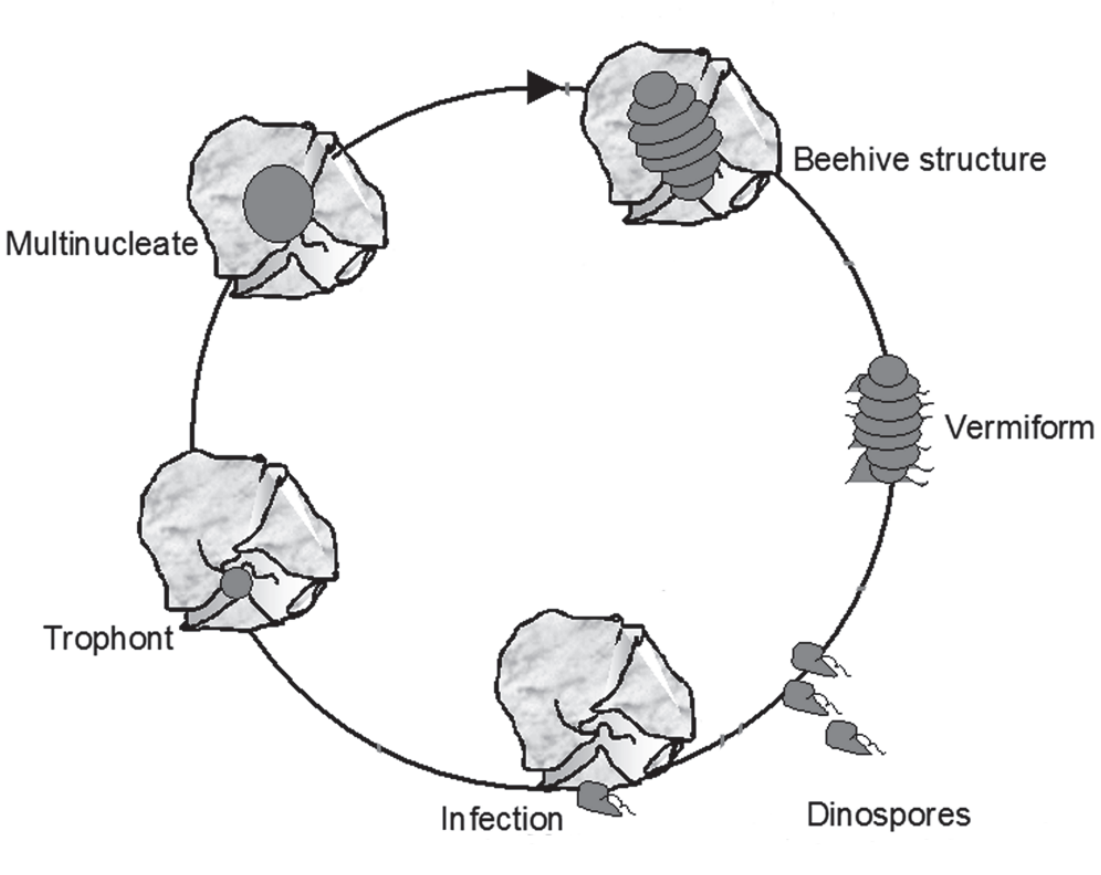


Fig. 1.4: The infection cycle of *Amoebophrya* infecting *Alexandrium*.

The basal syndinean dinoflagellates represented by the genus *Amoebophrya* are heterotrophic parasites that undergo reproduction within the confines of another dinoflagellate host cell. A study by Park *et al.* (2002b) indicated that the development of *Amoebophrya* within host nucleus may disrupt the flow of genetic information involved in plastidial function. A recent study, however, found that the chloroplast membrane galactolipid composition of *Alexandrium* infected by *Amoebophrya* did not cause a phenotypic alteration in the composition of MGDG and DGDG (mono- and di-galactosyldiacylglycerol, respectively), two galactolipids that comprise the majority of photosynthetic membranes (Leblond & Dahmen 2012). Furthermore, *Amoebophrya* did not produce MGDG and DGDG, unlike heterotrophic apicomplexan parasites *Plasmodium falciparum* and *Toxoplasma gondii*, distantly related to dinoflagellates, and that possess MGDG and DGDG as part of vestigial and non-functional plastids (Maréchal *et al.* 2002).

Endoparasitic *Amoebophrya* species have been reported from more than 75 host genera and were found to infect a variety of marine organisms including ciliates, radiolarians, free-living relatives, and even other parasitic dinoflagellates (Cachon 1964; Coats 1999; Kim & Park 2016). Cachon (1964) initially reported that *Amoebophrya ceratii* occurred in Mediterranean host species. *Amoebophrya* parasites were also found in planktonic organisms from the north-west Pacific, North Atlantic, North Sea, Baltic Sea and, recently, even in Korean and Chinese waters (Coats 1999; Fritz & Nass 1992; Kim *et al.* 2004; Li *et al.* 2014). *Amoebophrya* is also the most frequently recorded parasite genus infecting a wide taxonomic range of harmful dinoflagellates (Coats 1999; Park *et al.* 2013). For example, *Alexandrium* spp. (including *A. catenella/pacificum/fundyense* and *A. minutum*) capable of producing paralytic shellfish poisoning (PSP) toxins were parasitized by *Amoebophrya* with prevalences as high as 40% in Sequim Bay and Puget Sound, USA and the Penze estuary, France (Chambouvet *et al.* 2008). Particular interest has been aroused by the fact that since many dinoflagellate species are susceptible to infection by the dinoflagellate *Amoebophrya*, it might be possible to consider biological bloom control of harmful species via these parasites (Chambouvet *et al.* 2008; Coats 1999; Coats *et al.* 1996).

1.5 The dinoflagellate host *Alexandrium*: phylogeny and life history

Fossil evidence shows that dinoflagellates emerged in the Silurian period (400 MYA) or earlier, and major radiation with abundant thecate lineages seemed to occur from the Triassic onward (Fensome *et al.* 1999). One of the most recently diverged lineage is probably *Alexandrium tamarense* species complex, which emerged approximately 23-45 MYA (John *et al.* 2003). Among species within the *Alexandrium* genus, members of the *A. tamarense* species complex are the most extensively studied. Isolates from the *A. tamarense* species complex were originally assigned to *A. tamarense*, *A. fundyense* or *A. catenella* based on morphological characters (Balech 1995). However, many field and culture studies have revealed cells exhibiting morphologies intermediate between these three species (Destombe *et al.* 1992; Gayoso & Fulco 2006; Kim *et al.* 2002; Orlova *et al.* 2007). Phylogenies based on regions in the large ribosomal subunit rDNA demonstrated that the *A. tamarense* species complex comprised five distinct genetic clades named initially based on their presumed geographic distribution, and later renamed as Groups I-V because of the sympatry from same clades (John *et al.* 2003; Lilly *et al.* 2007; Wang *et al.* 2014). Data on morphology, ITS/5.8S genetic distances, ITS2 compensatory base changes, mating incompatibilities, toxicity, the *sxt*A toxin synthesis gene and rDNA phylogenies were analyzed for the *A. tamarense* species complex; this led to redefinition into five species, as follows: Group I, *A. fundyense*; Group II, *A. mediterraneum*; Group III, *A. tamarense*; Group IV, *A. pacificum*; and Group V, *A. australiense* (John *et al.* 2014) (Fig. 1.5).

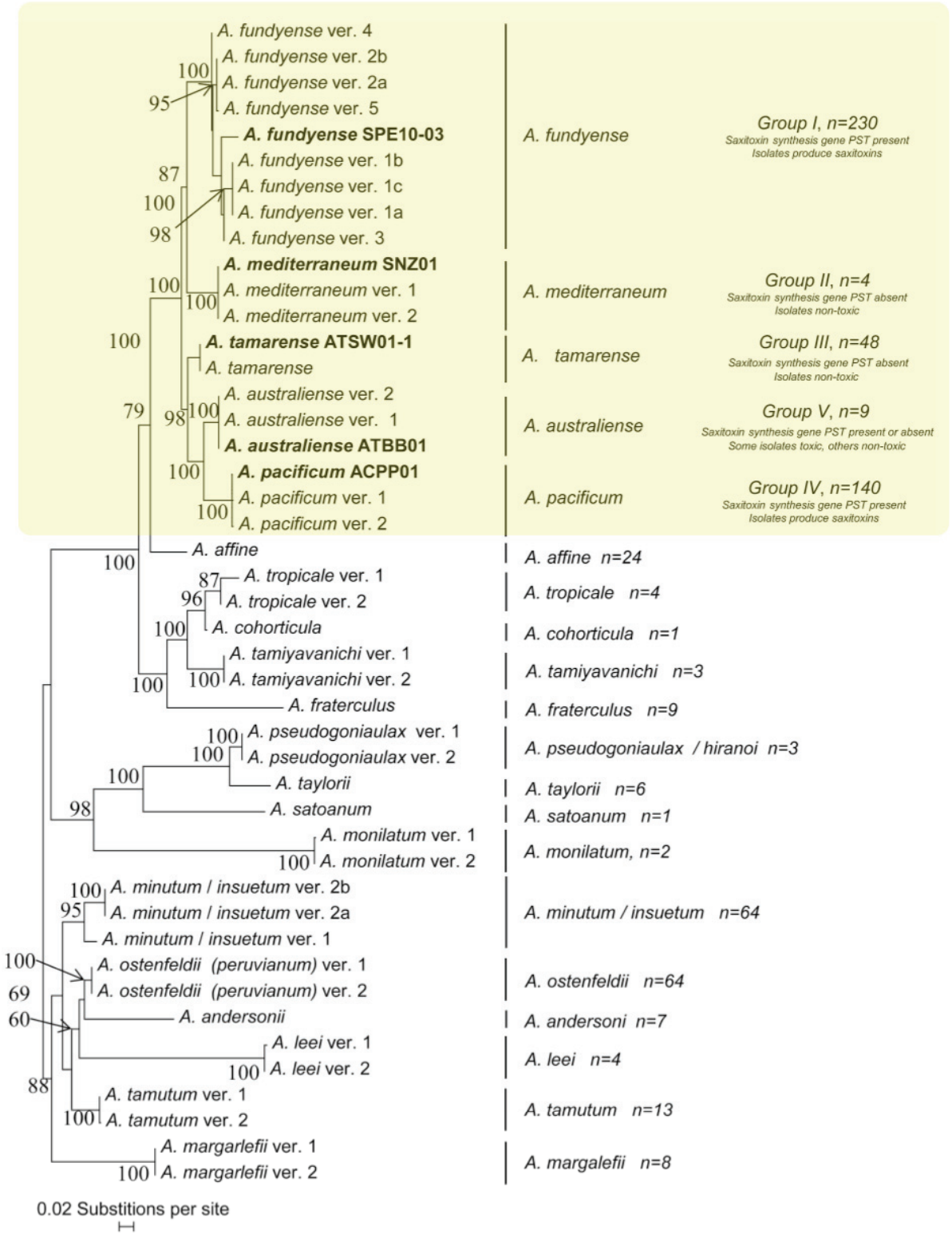


Fig. 1.5: A phylogeny of species of the genus *Alexandrium* based on D1-D2 LSU rDNA (John *et al.* 2014).

The life cycle of most *Alexandrium* species involves an alternation between asexual and sexual reproduction, containing haploid vegetative cells and diploid motile zygotes (Anderson 1998; Wyatt & Jenkinson 1997) (Fig. 1.6). The population of *A. tamarense* is initiated by hatching of diploid planomeiocytes from resting hypnozygotes from sea floor cyst beds at the beginning of the annual growth cycle (Anderson 1998; Wyatt & Jenkinson 1997). Planktonic bloom development is caused by repeat binary fission as a result of the proliferation of motile (vegetative) cells. This asexual process terminates when sexuality begins, such that gametes are formed and fuse to develop swimming zygotes (planozygotes) (Anderson 1998; Brosnahan *et al.* 2010). Vegetative cells can also transform into temporary cysts under stressful conditions such as a sudden change of temperature or salinity, and such temporary cysts can quickly transform back to motile cells when conditions become favourable again (Anderson 1998; Wyatt & Jenkinson 1997). In dinoflagellate populations, it is widely assumed that induction of sexuality is highly variable and depends on the environmental conditions (Anderson *et al.* 1984; Anderson & Lindquist 1985), genotypes present (Figueroa *et al.* 2007) and the time frame (Figueroa *et al.* 2010b). A recent study shows that sexual processes and meiosis occur in a regular diurnal cycle and are light-controlled similar to asexual division (Figueroa *et al.* 2015). These processes constitute a complete life cycle model for dinoflagellates.

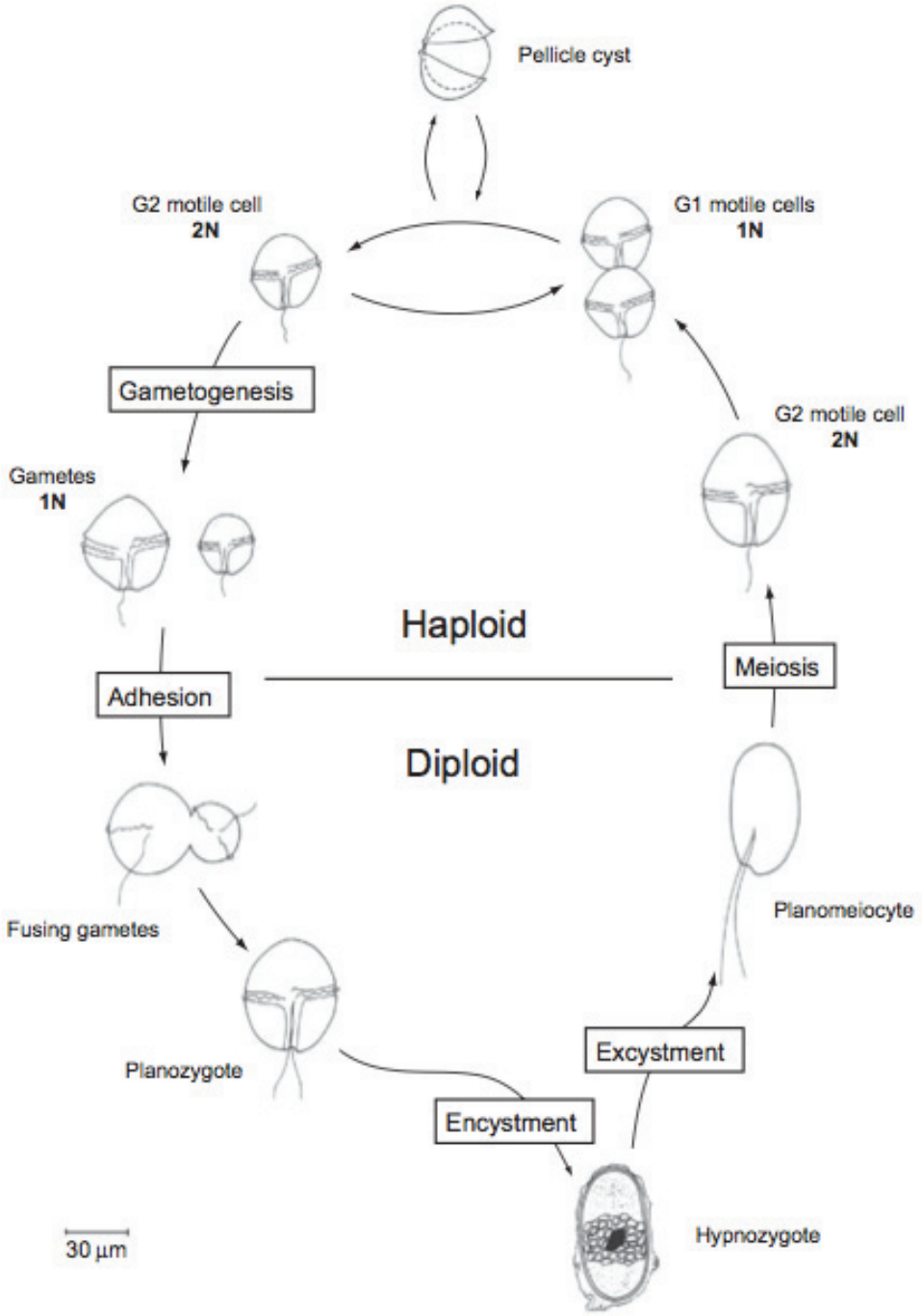


Fig. 1.6: Typical life history of members of the *Alexandrium* species complex (Brosnahan *et al.* 2010)

1.6 Harmful algal blooms and host-parasite ecology

"Red tide" is a common term used to refer to one of a variety of natural phenomena known as harmful algal blooms (HABs), since high densities of algal cells often discolour the surface waters due to their pigmentation. Algal blooms causing water discolouration may yield a wide variety of colours but do not always result in harmful or toxic events. In contrast, HAB species may cause harmful effects even at relatively low cell concentrations (and without visible water discolouration) if they are toxigenic. HABs are defined as algal events resulting in injury to human health or socioeconomic interests, or to components of aquatic ecosystems (Anderson *et al.* 2012b), particularly in coastal or near-shore areas worldwide.

Most HAB phenomena are caused by blooms of microalgae (or phytoplankton), including certain cyanobacteria (also known informally as "blue-green algae"). The term HAB can also apply to certain benthic microalgal aggregations and even to non-photosynthetic species (Landsberg 2002). As wells, HABs comprise harmful blooms of macroalgae ("seaweeds") that destroy habitat or displace indigenous species and deplete oxygen (Anderson 2009).

Marine dinoflagellates are the most prominent microalgae known to form such HABs, particularly of toxigenic species. Toxic dinoflagellate blooms have caused mortalities of marine fishes and invertebrates, and have been defined as serious hazards to human health. Among the best known toxin syndromes caused by ingestion of shellfish contaminated by toxic dinoflagellates are paralytic (PSP) and diarrhetic shellfish poisoning (DSP), often caused by *Alexandrium* spp. and *Dinophysis* spp., respectively (Cembella 2003). In addition to known toxins, which can be vectored through the food web and as a result become dangerous to marine mammals and seabirds, some dinoflagellates produce other bioactive compounds, called allelochemicals, causing cell lysis of plankton competitors and predators (Cembella 2003; John *et al.* 2015; Tillmann & John 2002; Wohlrab *et al.* 2010). Although some degree of allelopathy has been observed among dinoflagellates, the exact purpose and evolution of these compounds is strongly debated, in particular as their production is complex

and influenced by a combination of genetic and environmental factors (Anderson *et al.* 2012b; Cembella 2003; John *et al.* 2015; Wohlrab *et al.* 2010).

The formation of HABs and their ability to sustain high growth rate and biomass over time is strongly dependent on favorable abiotic conditions, such as solar radiation, nutrient concentration, salinity and water mass stability (Shilo 1967; Smayda 2002), typically known as "bottom up" factors. Nevertheless, in many cases, biotic parameters, known as "top down" factors, for instance, avoidance or reduction of competition and grazing (John *et al.* 2015; Tillmann & John 2002) may even be dominant in determining bloom dynamics.

Recent studies demonstrated that active participation of parasitic pathogens can also play important roles in the control of toxic blooms (Chambouvet *et al.* 2008). Members of the genus *Amoebophrya* sp. are known to parasitize heterotrophic or phototrophic dinoflagellates in coastal waters bordering six continents (Fig. 1.7) (Coats & Bockstahler 1994; Coats & Park 2002; Kim *et al.* 2004; Li *et al.* 2014; Park *et al.* 2002b; Park *et al.* 2013). Dinoflagellates infected by *Amoebophrya* have also been found in oligotrophic waters of the Mediterranean Sea, and environmental sequences attributed to the *Amoebophridae* have been obtained from oceanic surface waters and one deep ocean sample (Guillou *et al.* 2008; Siano *et al.* 2010).

Fig. 1.7: Global distribution of PSP toxins and occurrence of dinoflagellates infected by members of the *Amoebophrya* sp. until 2015.

While infections by *Amoebophrya* have been reported for many toxic dinoflagellate species, the dynamics of host toxins following infection has received little attention. The only available study (Bai *et al.* 2007) tracked karlotoxin content over the life cycle of *Amoebophrya* infecting the dinoflagellate *Karlodinium veneficum*. Their results showed no difference in cellular toxin content of infected and uninfected cultures over the infection cycle, suggesting that the parasite inhibits synthesis of new toxin, but does not break down existing toxin. Whether infection by *Amoebophrya* has influence on toxin dynamics of other toxin-producing dinoflagellate species remains unknown.

Our knowledge of parasite infection strategies of marine protists is rather limited, and even less is known about the possible defense mechanisms of the hosts to resist infection. Besides the discussion above that some parasites efficiently infect microalgae by producing toxins (Bai *et al.* 2007), microalgae can resist the pathogens by their capacity to rapidly produce cysts (Toth *et al.* 2004). A study has described the ability of the dinoflagellate *Alexandrium ostenfeldii* to form temporary cysts as a response to waterborne cues from the

parasitic Perkinsozoa *Parvilucifera infectans* (Toth *et al.* 2004). In contrast, parasites are also able to make use of resting stages to overcome unfavorable conditions. A recent study demonstrated *Amoebophrya* sp. can enter into a state of dormancy within the resting cyst of its host, the dinoflagellate *Scrippsiella trochoidea,* and new infective zoospores can then be released in the year following germination of the host (Chambouvet *et al.* 2011a).

2. Aims, hypotheses and outline of the thesis

Within my thesis project I aimed first to sequence the parasite *Amoebophrya* genome (which would have been the first completely sequenced dinoflagellate genome at the time the work was initiated). A further objective was to characterize the infection process of the host-parasite relationship with its co-evolutionary dynamics, and ultimately to define the genetic basis enabling *Amoebophrya* to control toxic *Alexandrium* population dynamics (if at all). Several initial testable hypotheses were proposed as follows:

1) there are notable variations in susceptibility of different *Alexandrium* geographical populations and/or in infectivity of the parasite *Amoebophrya*, as a result of local adaptation;
2) production of known phycotoxins, specifically PSP toxins, by *Alexandrium* is acting as a chemical defensive mechanism against infectivity and is regulated accordingly;
3) parasite characteristics (e.g., small genome size, gene loss, inability to synthesize certain amino acids) and common dinoflagellate features (e.g., highly reduced mitochondrion and plastid genome, trans-splicing, gene duplication) are reflected in the analysis of *Amoebophrya* genome;
4) host-parasite interaction causes significant changes in parasite gene expression over the time course of the infection;
5) exposure to the parasite may produce significant changes in the expression of *Alexandrium* genes associated with specific metabolic pathways, as co-evolutionary feedbacks;
6) waterborne cues from the parasite can induce similar responses and trigger the defense/stress mechanisms of the host.

From the ecological perspective, the parasite *Amoebophrya* exerts profoundly negative effects on its host. To investigate the adaptations of both the host and parasite on intra-specific level, in **Chapter I**, three natural host populations of *Alexandrium* were chosen and compared, showing different susceptibility to infection by the parasite *Amoebophrya*. Growth rates and infection percentages were determined to observe and quantify the effect of the parasite on its host. Furthermore, toxin concentration and composition were examined as a potential defense strategy of the host and microsatellite PCR data was analysed to detect genetic adaptations. In addition to investigating a possible correlation between resistance and host genotype, host growth rates, lytic rates and infection percentages were also statistically compared to toxin profiles of *Alexandrium*.

To better understand dinoflagellate evolution and parasitic genome properties, the genome of a strain of parasitic *Amoebophrya* that infects the toxic microalga *Alexandrium* was sequenced and analysed (**Chapter II)**. This analysis provides insights regarding the metabolism of the parasite, several parasitic features and survival strategies linked to this unusual genome.

Few studies have examined the molecular regulation of parasites over a life cycle during infection of the toxic dinoflagellate *Alexandrium*. In **Chapter III**, the transcriptome of the parasite *Amoebophrya* was profiled at three different life history stages: the pure dinospores stage (0 h), the initial infection/penetration stage (6 h), and the maturation stage (96 h). By analysing the expressed sequence tags (ESTs) obtained from different life stages, processes and genes were identified that may be relevant to hosts' susceptibilities. These data increased our understanding of the genetic basis enabling *Amoebophrya* sp. to dominate over toxic algal blooms.

In **Chapter IV**, the investigation of molecular mechanisms was carried on the host *Alexandrium* in response of parasite infection, for which little is currently known about the importance of chemical cues that may prime host responses towards parasites. Analogous to

the previously performed experiment (**Chapter II**), RNAs derived from the treatments at three time points were analysed. RNA sequencing (RNA-seq) was used to compare the transcriptional responses of the host *Alexandrium* to the presence of the parasite *Amoebophrya* and waterborne chemical cues from this parasite, by contrast, to discriminate the results from the response to lysed *Alexandrium* cells. Transcripts that were statistically differentially expressed between treated and untreated *Alexandrium* samples were annotated. Toxins were also examined as a potential defense strategy of the host.

3. List of publications and declaration of contribution

This thesis is organized in four chapters and corresponded to separate research articles of which two are published.

List of publications and declaration of contribution:

Publication I

Lu Y, Braitmaier M, Wang C, John U, Cembella A

Intraspecific variability in the susceptibility of marine dinoflagellate *Alexandrium fundyense* to infection by parasite *Amoebophrya* sp.

To be submitted

The candidate designed the research, performed experiments, analysed the data and prepared the manuscript.

Publication II

John U, **Lu Y**, Wohlrab S, Groth M, Kohli G, Janouškovec J, Keeling P, Felder M, Guillou L, Farhat S, Porcel B, Moustafa A, Valentin1 K, Frickenhaus S, Glöckner G

Eukaryotic-aerobic life is possible without a mitochondrial genome: The parasitic dinoflagellate *Amoebophrya* sp.

To be submitted

The candidate took part in performing the experiments, interpretation of the data and preparation of the manuscript

Publication III

Lu Y, Wohlrab S, Glöckner G, Guillou L, John U (2014)

Genomic insights in processes driving the infection of *Alexandrium tamarense* by the parasitoid *Amoebophrya* sp.

Eukaryotic Cell **13**, 1439-1449

DOI: 10.1128/EC.00139-14

The candidate designed the research, performed experiments, analysed the data and prepared the manuscript.

Publication IV

Lu Y, Wohlrab S, Groth M, Glöckner G, Guillou L, John U (2016)

Transcriptomic profiling of *Alexandrium fundyense* during physical interaction with or exposure to chemical signals from the parasite *Amoebophrya*

Molecular Ecology

DOI: 10.1111/mec.13566

The candidate designed the research, performed experiments, analysed the data and prepared the manuscript.

4. Chapters

4.1 Intraspecific variability in the susceptibility of the marine dinoflagellate *Alexandrium fundyense* to infection by the parasite *Amoebophrya* sp.

4.1.1 Abstract

The parasitic dinoflagellate *Amoebophrya* sp. has been observed to infect the widespread toxic bloom-forming dinoflagellate *Alexandrium fundyense*, and play important roles in control of harmful algal blooms. Little is known about the effects of parasitic infection on the host's population and toxin dynamics. Here, two clones of the parasite *Amoebophrya* from the Gulf of Maine area were provided to different strains of the host *Alexandrium* from three very distant geographical origins (Alaska, the Gulf of Maine, USA and the North Sea). After 94 hours of incubation, a strong negative effect of parasitism on the development of host populations was observed, indicating that the parasite *Amoebophrya* has great potential in controlling blooms of *Alexandrium*. No significant differences in growth rate and lytic rate of the host *Alexandrium* were observed among the three geographical populations, implying that there might no apparent adaptation of the host on an intro-specific level. Also no difference in the parasite prevalence of each geographical population could be identified, but the parasite infection percentages were highly variant on an intra-population level. Cellular toxin contents were examined, showing that neither toxin concentration nor composition changed in each geographical population. Therefore, the results indicated that the host *Alexandrium* might not use toxins as a potential defense strategy against the parasite.

4.1.2 Introduction

Dinoflagellates are important primary producers in the ocean and constitute an integral microbial component of diverse marine food webs. The lethal endoparasitic dinoflagellate *Amoebophrya* can infect a variety of marine organisms including its dinoflagellate relatives (Cachon 1964; Coats 1999; Park *et al.* 2013). *Amoebophrya* species have been recorded from more than 75 host dinoflagellate species and have been identified as important regulatory factors in host population dynamics, with the highest infection percentages occurring near the termination of host blooms (Coats *et al.* 1996; Park *et al.* 2004).

Many toxigenic dinoflagellate species, including those responsible for formation of dense aggregations known as Harmful Algal Blooms (HABs), are among the victims of parasitism by *Amoebophrya*. For example, previous field studies have shown that *Amoebophrya* cells were able to infect the toxigenic dinoflagellate *Alexandrium minutum*, known for the production of potent neurotoxins associated with paralytic shellfish poisoning (PSP), with a prevalence (e.g., the percentage of infected cells in the host population) of up to 40% in Penzé estuary, France (Chambouvet *et al.* 2008). Similarly, the dinoflagellate *Dinophysis norvegica,* associated with the production of polyether diarrhetic shellfish poisoning (DSP) toxins, were found to be infected by *Amoebophrya* at a prevalence as high as 28% in the Baltic and North Sea (Salomon *et al.* 2003a, b).

The parasitic dinoflagellate *Amoebophrya* belongs to the order Syndiniales (Alveolata), all members of which represent endosymbionts and/or parasitoids. The genus displays tendencies regarding host specificity, but populations are globally distributed over a wide range of potential hosts and exhibit a high genetic diversity (Guillou *et al.* 2008). *Amoebophrya* has a simple life cycle lasting approximately 2-3 days in most host species (Coats & Park 2002). Infection by *Amoebophrya* is initiated by penetration of the parasitic dinospores (~5 μm transdiameter) into the host cells (Cachon 1964). Once inside the cytoplasm or nucleus, the parasite starts to feed (the trophont stage). The trophont increases in

size until sequential nuclear divisions and flagellar replications ultimately form an intracellular and multicellular 'beehive' stage inside the cytoplasm or nucleus of the host cell (the sporocyte). The mature sporocyte ruptures the cell wall of the host, and most develop into a short-lived vermiform stage that soon divides into numerous free-living infectious dinospores (Coats & Bockstahler 1994; Coats & Park 2002). Infection by *Amoebophrya* inhibits host reproduction and eventually leads to death of the host.

The potent neurotoxins known collectively as paralytic shellfish toxins (PSTs) are poisonous for many marine fauna and can cause disruptive effects when accumulated via marine food webs, and also serious even lethal consequences for human consumers of seafood contaminated by these toxins (Anderson *et al.* 1994; Cembella 2003). The PSTs comprise the tetrahyropurine alkaloid saxitoxin (STX) and at least 57 described analogues. These toxins can be separated into major structural groups, including the high potency carbamoyl (STX, neoSTX, gonyautoxins [GTX]), the intermediate potency decarbamoyl (dcSTX, dcneoSTX, dcGTX) and low potency *N*-sulfocarbamoyl (B- and C-) toxins (Llewellyn 2006). Recently, relatively hydrophobic forms known as benzoyl derivatives have also been described in the dinoflagellate *Gymnodinium catenatum* but the specific toxicity remains uncertain (Bustillos-Guzmán *et al.* 2015; Vale 2010).

The PSTs are produced among various species and strains of marine dinoflagellates belonging to the genus *Alexandrium*, the naked gymnodinoid species *Gymnodinium catenatum* and the heavily armoured *Pyrodinium bahamense* (Cembella 1998). The capacity for toxin production and the toxin composition is genetically fixed, but varies widely even among populations belonging to the same species (Alpermann *et al.* 2010). In addition, cellular toxin content of dinoflagellates belonging to the genus *Alexandrium*, is known to be affected by a variety of abiotic factors (e.g., temperature, salinity, light, and nutrients) that affect growth rate, reviewed by Cembella (1998), and biotic factors (e.g., defensive response to grazing by predators or competitors) (Selander *et al.* 2006; Wohlrab *et al.* 2010).

By contrast, the dynamics of dinoflagellate cell toxin content and composition influenced by parasitism has received little attention. Bai *et al*. (2007) described a positive correlation between karlotoxin concentration in *Karlodinium veneficum* and infection by *Amoebophrya*; the mean karlotoxin content in infected cultured cells was significantly lower than that of uninfected cells. A recent study by Kim and Park (2016) examined *Amoebophrya* infecting the dinoflagellate *Alexandrium fundyense* and paradoxically showed different results, i.e., that mean cell toxin content for infected cultures was significantly higher than that for uninfected cultures. Little is known about the effect of parasitism on *Alexandrium* toxin production, exposure of different *Alexandrium* strains and analysis of the effects of PST dynamics during parasite infection remain to be explored.

The toxic dinoflagellate *Alexandrium fundyense* (John *et al.* 2014), known as a member of the *Alexandrium tamarense* species complex among other previous designations, is widely globally distributed, with highest prevalence and extensive bloom formation particularly in northern temperate and sub-Arctic waters (Anderson *et al.* 2012a). Furthermore, populations of *A*. *fundyense* from diverse locations have been shown to be susceptible to parasitic infection by *Amoebophrya* (Kim & Park 2016; Velo-Suárez *et al.* 2013), although the ecological consequences on bloom dynamics and fate of toxins remain unknown. The objective of this study was to investigate the susceptibility of host *Alexandrium fundyense* to infection by the parasite *Amoebophrya* on an intra-specific level. To address these issues, host strains were obtained from three very distant geographical locations (Alaska, the Gulf of Maine, USA and the North Sea) to study the differences between different host populations in susceptibility to parasite infection and to search for putative adaptations against infection. Potential shifts in host cell toxin concentration and composition were also examined as a possible inducible defense strategy of the infected dinoflagellates.

4.1.3 Materials and Methods

Culture origin and protocols

The 30 strains of *Alexandrium fundyense* were isolated from natural plankton populations from Alaska, the North Sea and the Gulf of Maine, USA (Fig. 4.1.1 and Table 4.1.1). The strains were grown in K-medium (Keller *et al.* 1987) prepared from the North Sea water sterilized by filtration through a 0.2-µm pore-size sterile filter (VacuCap; Pall Life Sciences, Dreieich, Germany). The *Alexandrium* cultures were kept in exponential growth by weekly transfers into fresh K-medium. All host strains, including a standard host strain for parasites (Alex 5) (Tillmann *et al.* 2009) belonged to *A. fundyense* (ribotype Group I) (John *et al.* 2014).

The two clones of the parasite *Amoebophrya* sp. were isolated from host *Alexandrium* cells sampled from Salt Pond, MA, USA (Chambouvet *et al.* 2011b) (Fig. 4.1.1 and Table 4.1.1). *Amoebophrya* sp. infecting *Alexandrium* was maintained on a standard host *Alexandrium* strain (Alex5) isolated from the North Sea coast of Scotland (Tillmann *et al.* 2009). All cultures were grown at 15 °C, under cool-white fluorescent lamps (Phillips, Amsterdam, Netherlands) providing a photon flux density of 150 µmol photons m^{-2} s^{-1} on a light:dark cycle of 14:10 h.

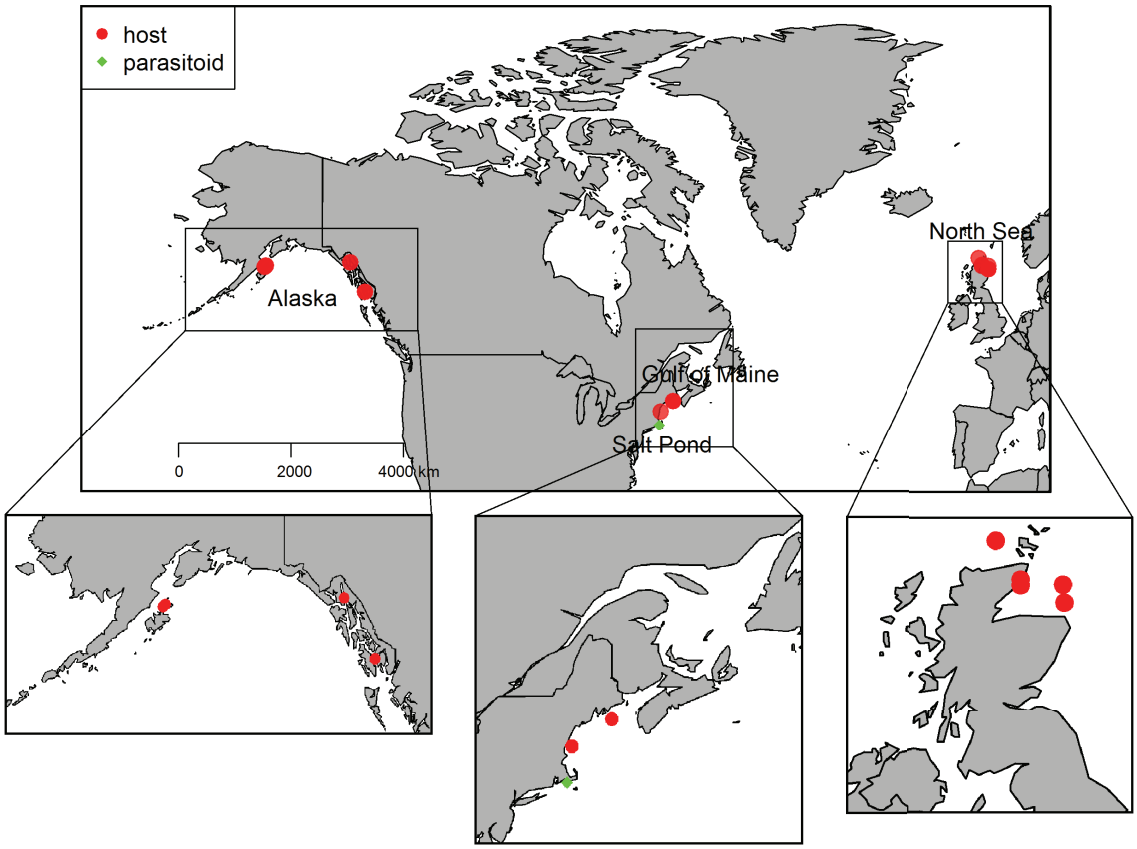


Fig. 4.1.1: Geographical origin of the cultures in the infection experiment. Red dots indicate the origin of the strains of *Alexandrium fundyense*, the green squares indicate the origin of the two clones of *Amoebophrya* sp.

Table 4.1.1: Names and origin of the host and parasite strains in the infection experiments.

Host population	Strain number	Strain ID	Latitude N	Longitude E	Year of isolation
Alaska	A_1	AUK09121111-6	58°37'30.26"	134°65'82.45"	2012
	A_2	AUK09121111-9	58°37'30.26"	134°65'82.45"	2012
	A_3	THB052411-8	55°33'82.73"	131°64'00.59"	2012
	A_4	THB052411-11	55°33'82.73"	131°64'00.59"	2012
	A_5	THB052411-13	55°33'82.73"	131°64'00.59"	2012
	A_6	TRT011311-1	57°78'08.09"	152°39'13.04"	2012
	A_7	TRT011311-2	57°78'08.09"	152°39'13.04"	2012
	A_8	USC010511-11	57°73'48.02"	152°51'13.56"	2012
	A_9	WBS031412-8	57°70'45.9"	152°55'46.03"	2012
	A_10	WBS031412-12	57°75'71.91"	152°47'97.89"	2012
North Sea	NS_1	HE358-A3	57°54'97.22"	1°57'38.89"	2010
	NS_2	HE358-A5	57°54'97.22"	1°57'38.89"	2010
	NS_3	HE358-A6	57°54'97.22"	1°57'38.89"	2010
	NS_4	HE358-A8	57°54'97.22"	1°57'38.89"	2010
	NS_5	HE358-C3	58°63'41.67"	3°60'38.89"	2010
	NS_6	HE358-E4	57°54'97.22"	1°57'38.89"	2010
	NS_7	HE358-E6	57°54'97.22"	1°57'38.89"	2010
	NS_8	HE358-F6	57°74'94.44"	2°00'02.78"	2010
	NS_9	HE358-G2	57°75'19.44"	2°75'75"	2010
	NS_10	HE358-H1	58°20'63.89"	2°75'88.89"	2010
Gulf of Maine	GOM_1	2000CB-08D5rep02	43°41'44"	69°53'41"	2000
	GOM_2	F14 7/3/08 KL	42°16'49"	70°47'03"	2005
	GOM_3	H15 7/3/08 KL	43°41'57"	69°53'43"	2005
	GOM_4	D3 7/3/08 KL	42°41'39"	70°29'07"	2005
	GOM_5	I3 7/3/08 KL	44°58'58"	66°49'30"	2005
	GOM_6	I8 7/3/08 KL	44°58'58"	66°49'30"	2005
	GOM_7	HT 140-G5	44°04'72"	67°46'61"	2001
	GOM_8	HT 140-G7	44°04'72"	67°46'61"	2001
	GOM_9	HT 140-G10	44°04'72"	67°46'61"	2001
	GOM_10	GTCA 29	43°	70°19'	1985
Parasite clone					
Salt Pond (MA)	Parasite 1	ATSPSF7-5	41°32'34.83"	70°37'36.78"	2003
	Parasite 2	GOM (HQ337038)	41°32'34.83"	70°37'36.78"	2003

Fixation and counting methods

Host cells (1 ml) were fixed in the culture flasks with Lugol's iodine solution (10 g potassium iodide, 5 g iodine, 100 ml distilled water) with a final concentration of 2% (Tillmann *et al.* 2009), and three 100 or 200 µl aliquots were counted under an inverted microscope (Zeiss Axiovert 200M, Oberkochen, Germany) after sedimentation in counting chambers. The number of cells counted was always more than 400 per sample. The growth rate of *Alexandrium* was calculated according to the following equation:

$$\mu = \frac{1}{t_2 - t_1} ln \frac{N_2}{N_1}$$

,where µ is the growth rate (d^{-1}), t is the sampling day, and N_1 and N_2 are the abundances of *Alexandrium* at t_1 and t_2, respectively.

The lytic rate of *Alexandrium* was calculated using the equations of Frost (1972)

$$I = \frac{N_2 - N_1}{(t_2 - t_1)(k - g)} gV$$

where I is the lytic rate (cells ml^{-1} h^{-1}), and V is the volume (ml) of the culture. The growth constant for algal growth, *k*, was calculated from

$$N_{c2} = N_{c1} e^{k(t_2 - t_1)}$$

where N_{c1} and N_{c2} are the abundances of *Alexandrium* in the control cultures at t_1 and t_2, respectively. The lysing coefficient, *g*, was calculated from

$$N_2 = N_1 e^{(k-g)(t_2 - t_1)}$$

Parasite samples (1 ml) were fixed with formaldehyde (10% $CaCO_3^-$ buffered formaldehyde; 2% final concentration) (Coats & Park 2002), and three 100 or 200 µl aliquots were counted in duplicate by inverted microscope after sedimentation in chambers. Infection percentages of the parasites (i.e., the percentage of infected host cells) were determined by detecting the natural auto-fluorescence (Zeiss Axiovert 200M, Oberkochen, Germany) of the parasites (Coats & Bockstahler 1994), where the parasites showed distinctive green auto-fluorescence, and the hosts showed clear red auto-fluorescence. Only well-maturing parasites and therefore only obvious infections were counted.

Infection experiments

Triplicate cultures (30 ml) of *A. fundyense* at a concentration of approximately 3 x 10^3 cells ml^{-1} were prepared in tissue culture flasks one day prior to the parasite exposure treatment. Triplicate cultures of each strain containing the host only served as no-parasite controls (negative controls). The standard host *Alexandrium* strain (Alex5) served as positive control, also in triplicate. Infection of the host culture was established following the methods of Coats and Park (Coats & Park 2002). Infective parasite dinospores for the experiment were harvested from an infected host culture by gravity filtration through a 10 µm-pore-size mesh. The dinospores were examined under the microscope to ensure that they were actively swimming and that no host cells remained. The dinospores were immediately inoculated into the *Alexandrium* cultures at a parasite:host cell ratio of approximately 10:1. Negative controls received the same volume of fresh K-medium instead. After 94 h of incubation, the infected cells as well as the negative controls were harvested from each triplicate culture for (a) fixation and counting, (b) assessment of parasite infection rates and (c) PST cell content analysis.

Toxin analysis

Just before the parasite treatment experiment (0 h), samples (15 ml) of each cultured *Alexandrium* strain (in triplicate) were collected for PST analysis. After 94 h incubation, 10 ml of the parasite infected and negative control *Alexandrium* cultures were collected and the cells harvested by centrifugation (Centrifuge 5810R, Eppendorf, Hamburg, Germany) in 15 ml centrifuge tubes at 3,220 x *g* for 15 min. The supernatant was discarded and the cell pellet was transferred after resuspension with 1 ml of sterile-filtered seawater to a 2 ml reaction tube and centrifuged (Centrifuge 5415R, Eppendorf, Hamburg, Germany) at 15,682 x *g* for 10 min. After removing the supernatant, cell pellets were homogenized with Lysing Matrix D (MP Biomedicals, Santa Ana, California, USA) and 500 µl 0.03 M acetic acid in a Fast Prep™ FP 120 (Thermo-Fisher Scientific, Waltham, Massachusetts, USA) at speed 6.5 for 45 s. Afterwards, the extracts were centrifuged at 15,682 x *g* for 15 min. The supernatant was passed through a spin-filter (pore size 0.45 µm, Millipore Ultrafree, Eschborn, Germany) by

centrifugation at 835 x *g* for 30 s. The filtrate was transferred into 2 ml autosampler vials (Agilent Technologies, Santa Clara, USA) for toxin analysis.

The PSTs were analysed by reverse-phase ion-pair liquid chromatography with post-column fluorescence detection (LC-FD) as described in Krock *et al.* (2007). The analyses were carried out on an LC1100 series liquid chromatograph (Agilent Technologies, Santa Clara, USA) connected to a PCX 2500 post-column derivatization module (Pickering Laboratories, Mountain View, CA, USA) with an injection volume of 20 µl. The Agilent chromatography system comprised the following components: G1379A degasser, G1311A quaternary pump, G1329A autosampler, G13308 sample thermostat and G1321A fluorescence detector. Fluorescent toxin derivatives were detected with the dual monochromator fluorescence detector set at an excitation wavelength of 333 nm and an emission wavelength of 395 nm. The two mobile phases for chromatography consisted of Eluent A (6 mM octanesulphonic acid, 6 mM heptanesulphonic acid, 40 mM ammonium phosphate, adjusted to pH 7.0 with dilute phosphoric acid, and 0.75% tetrahydrofuran [THF]) and Eluent B (13 mM octanesulphonic acid, 50 mM phosphoric acid adjusted to pH 6.9 with ammonium hydroxide, 15% (v/v) of acetonitrile, 1.5% THF). The gradient, with a flow rate of 1 ml min^{-1}, was as follows: 0 min 100% A, 15 min 100 %A, 16 min 100 %B, 35 min 100 %B, 36 min 100% A, and 45 min 100% A.

Toxin separation was performed on a Luna C18 reversed-phase analytical column (Phenomenex, Aschaffenburg, Germany) equipped with a Phenomenex SecuriGuard pre-column. The eluate was oxidized with 10 mM periodic acid in 550 mM ammonium hydroxide at 50 °C and then acidified with 0.75 M nitric acid. Toxin identification and concentrations were determined according to toxin standard solutions from the CRMP Programme, Institute for Marine Biosciences, National Research Council Canada (Halifax, Nova Scotia, Canada).

Population genetic structure

The genetic structure among the geographical populations of *A. fundyense* was established with reference to a total of fifteen microsatellite loci (Alpermann *et al.* 2006; Nagai *et al.* 2004) (primers see Table S4.1.1). Amplifications were based upon the DNA templates of the 30 clonal *Alexandrium* isolates from the three geographical populations (Table 4.1.1). DNA extractions were performed with a DNeasy plant mini Kit (Qiagen, Hilden, Germany) and a NucleoSpin Plant II Kit (Macherey Nagel, Düren, Germany). DNA purification was performed according to the manufacturer's instructions (Qiagen). The PCR reactions were carried out with a Type-it Microsatellite PCR Kit (Qiagen, Hilden, Germany) in the PCR Mastercycler nexus (Eppendorf, Hamburg, Germany). The reaction mixture included 1 µl (10 ng) of template DNA, 10 µl 2 x Type-it Multiplex PCR Master Mix and 0.2 µl primers (forward and reverse). Nuclease free water was added to this mixture to an end-volume of 25 µl. The PCR cycle was as follows: initial denaturation at 95 °C for 5 min followed by 32 cycles of denaturation at 95 °C for 30 s, annealing at 60 °C for 90 s and elongation at 72 °C for 30 s. A final extension at 72 °C was performed for 30 min. Fragment lengths were analysed with ROX GS500 size standard (Applied Biosystems, Lincoln Centre Drive Foster City, CA, USA) on an ABI 3130lx sequencer (Applied Biosystems, CA, USA). Allele sizes were manually binned and genotypes were assessed in GENEMAPPER 3.7 (Applied Biosystems, Lincoln Centre Drive Foster City, CA, USA).

The Fixation index (F_{ST}), a measure of genetic differentiation among populations, was determined from the microsatellite sequence data. Pairwise F_{ST} values (Weir & Cockerham 1984) among populations were calculated in ARLEQUIN 3.5 (Excoffier & Lischer 2010). The F_{ST} values can range from 0 to 1, where 0 means complete sharing of genetic background and 1 signifies no sharing. A rough classification of F_{ST} values for microsatellite-based studies suggests strong population differentiation for a significant F_{ST} value more than 0.25 (Hartl *et al.* 1997).

A more robust Bayesian clustering analysis implementing a model-based method was performed with the software STRUCTURE 2.3.4 (Pritchard *et al.* 2000). Briefly, the software assumes a model comprising *K* populations (where *K* may be unknown), each of which is characterized by a set of allele frequencies at each locus. Individuals in the sample are assigned (probabilistically) to populations, or jointly to two or more populations if their genotypes indicate that they are admixed. Without being given any prior population information, *K* was let to range from 1 to 5. Analyses were performed with the independent allele frequencies option under the non-admixture model, i.e. under the assumption that there is no gene flow among populations, as well as with the correlated allele frequencies option under the admixture model, assuming certain genetic connectivity among populations. Twenty runs with 200,000 Markov chain Monte Carlo (MCMC) iterations after a burn-in period of 25,000 steps were carried out for each *K*. The results were uploaded onto STRUCTURE Harvester (Earl & vonHoldt 2012) and *K* was determined from the *ad hoc* statistic ΔK (Evanno *et al.* 2005), as well as mean estimates of posterior probability *L(K)* (Pritchard *et al.* 2000). Results from the 20 replicates of the most likely value for *K* were averaged with the software CLUMPP 1.1.2 (Jakobsson & Rosenberg 2007) and the output was visualized by DISTRUCT 1.1 (Rosenberg 2004).

Phylogenetic relationship

Sequencing of the D1/D2 region of the LSU rRNA was performed with an ABI Prism 3130xl Genetic Analyzer (Applied Biosystems, Darmstadt, Germany) after PCR amplification with the two primers D1R and D2C (Scholin *et al.* 1994). The PCR products were cloned into a vector, using the TOPO TA Cloning kit (Invitrogen, Carlsbad, CA, USA). Reference sequences of members of all five ribotype groups of the *Alexandrium tamarense* species complex, as well as of *A. affine* as an out-group, were obtained from the NCBI database. The obtained sequences were aligned with ClustalW in MEGA6 (Tamura *et al.* 2013). The maximum likelihood (ML) phylogenetic tree was finally generated in MEGA6, by accessing the Tamura-Nei model and with 1,000 bootstrap replications.

4.1.4 Results

Time-course study of infected cultures

Infections were observed in all *Alexandrium* host strains, where the hosts showed clear red auto-fluorescence, and the parasites showed distinctive green auto-fluorescence. Control cells were either untreated (negative control) or exposed to the standard host *Alexandrium* strain (Alex5) (positive control). In all cases there was a significant difference in cell concentrations between 0 h (immediately after the infection) and after 94 h (late growth stage) due to parasite infection. At 94 h, the growth rates of *Alexandrium* treated with the parasite *Amoebophrya* were significantly lower than those of their corresponding controls (ANOVA: $P < 0.05$) (Fig. 4.1.2). Although no significant difference in growth rate of *Alexandrium* was observed among the three populations (ANOVA: $P > 0.05$), the difference of growth rate between the treatment and the control in Gulf of Marine strains was three-fold higher than that of North Sea strains and one-and-a-half-fold higher than that of Alaska strains.

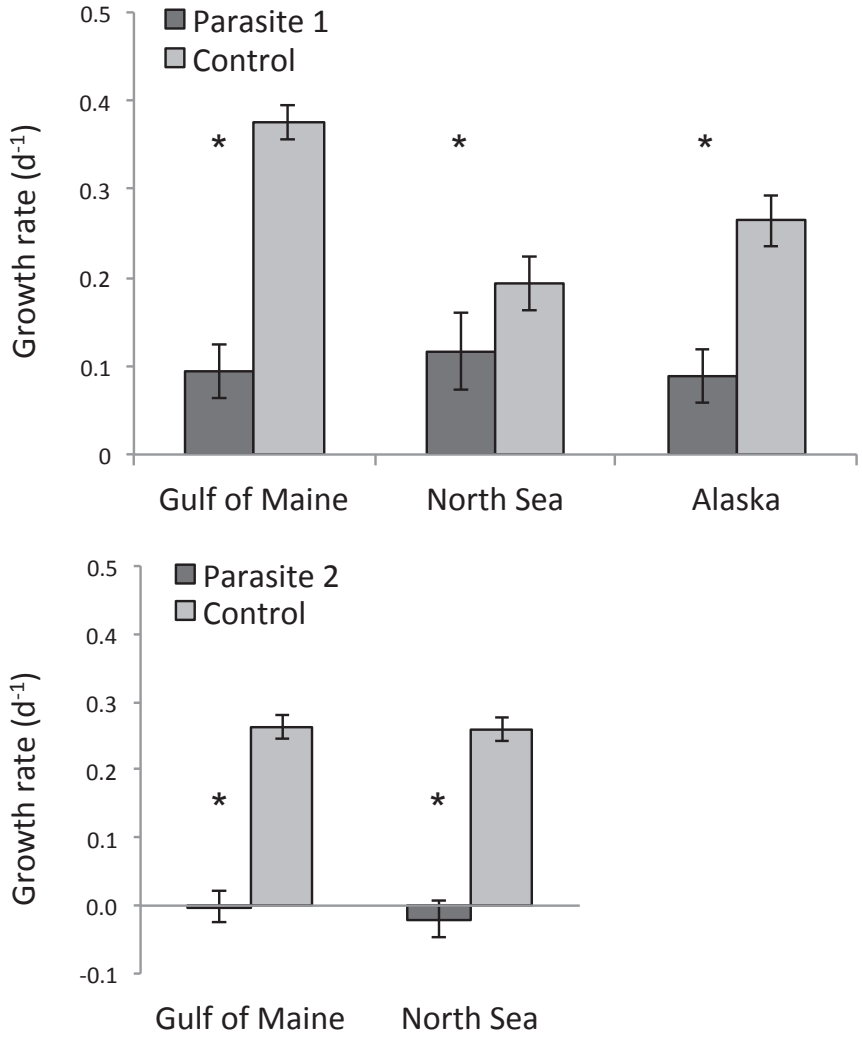


Fig. 4.1.2: Growth rates of *Alexandrium* from three very distant geographical locations (Alaska, the Gulf of Maine and the North Sea) exposed to two clones (Parasite 1 and 2) of the parasite *Amoebophrya*. Asterisks indicate a significant difference in the growth rate compared to the corresponding control in each geographical population (ANOVA: $P < 0.05$, $n = 30$, error bars indicate standard errors of the mean).

Lytic rate of Alexandrium

The lytic rate of *Alexandrium* due to parasite infection in Gulf of Maine strains was significantly higher than that of North Sea and Alaska strains under the incubations with the Parasite 1 (ANOVA: $P < 0.05$) (Fig. 4.1.3). By contrast, no significant difference between the two geographical populations was observed in the incubations with the Parasite 2 (ANOVA: $P > 0.05$).

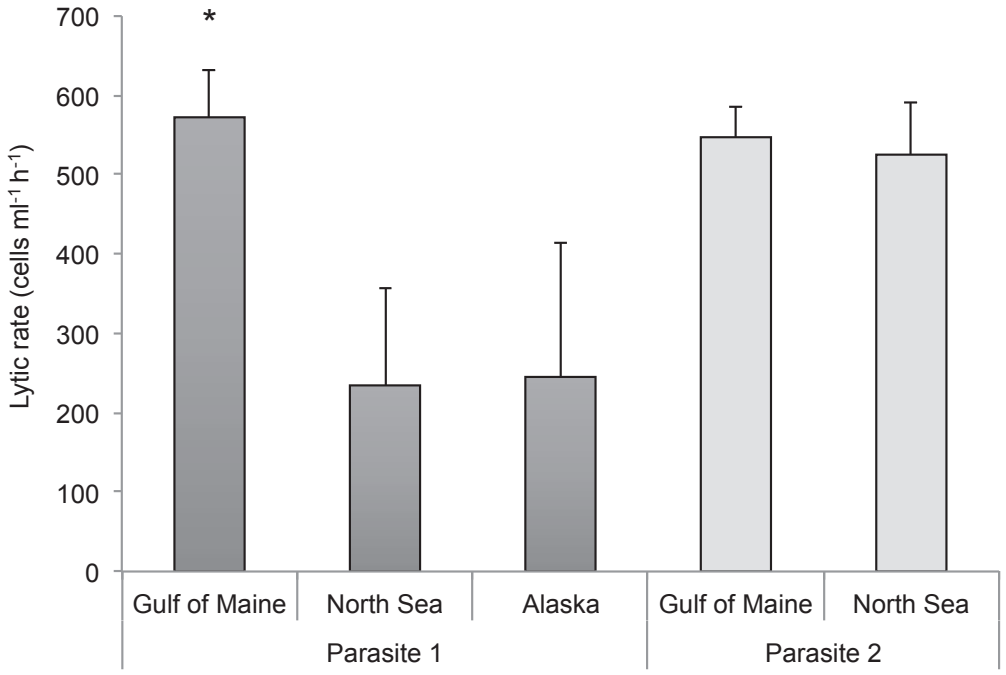


Fig. 4.1.3: Lytic rates of *Alexandrium* exposed to two clones (Parasite 1 and 2) of the parasite *Amoebophrya*. The asterisk indicates significant differences in the lytic rate under the incubations within the Parasite 1 (ANOVA: $P < 0.05$, $n = 30$, error bars indicate standard errors of the mean).

Infection percentage of Alexandrium

The distribution of infection percentages after 94 h of incubation showed some differences, but an ANOVA with *post hoc* test (Tukey's Honest Significant Differences test) only showed significant differences between the host strains from the North Sea with Parasite 1 (Fig. 4.1.4, See appendix for detailed results of the *post hoc* test). The North Sea strains with the Parasite 1 had an extremely low infection percentage (<5%), however, the infection percentage of the positive control was also very low (7%).

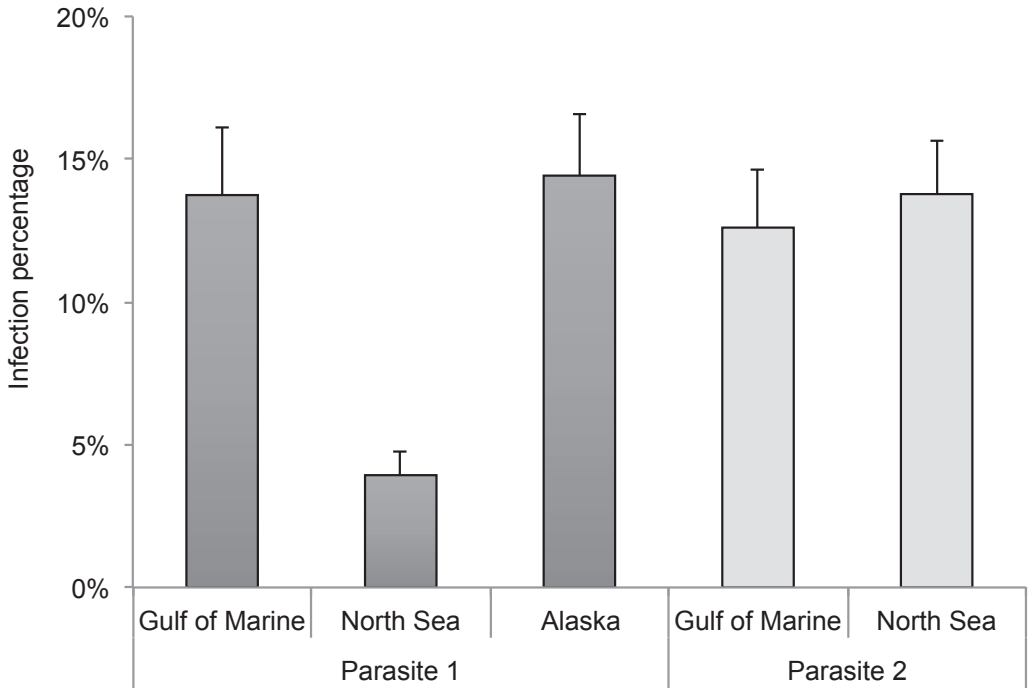


Fig. 4.1.4: Infection percentage of each *Alexandrium* strains from three geographical locations (Alaska, the Gulf of Maine and the North Sea) with two clones of parasite (Parasite 1 and 2) *Amoebophrya* (n = 30, error bars indicate standard errors of the mean).

The infection percentages were very different on the intra-population level (Fig. 4.1.5). There were two strains, GOM_11 and NS_4, that were entirely resistant to parasite infection and others that showed only a low infection percentage (<5%), e.g., GOM_14, NS_7, NS_8 or A_4. Conversely, some strains had a very high infection percentage that in the case of GOM_2 (36%) was even higher than that of the positive controls (30%).

Some strains of *Alexandrium* from the Gulf of Maine behaved identically when confronted with the two clones of parasites, for instance GOM_2 was the most highly infected and GOM_11 was the lowest infected isolate in both cases (Fig. 4.1.5). The majority of strains behaved at least similar in confrontation with the different parasites. The biggest difference in infection percentage occurred in the strain GOM_1 with a mean infection percentage of 20% (with Parasite 1) versus 3% (with Parasite 2). The comparison of the incubations of the North Sea strains with the two parasite clones is not valid, since infection percentages in the incubation with Parasite 1 were extremely low.

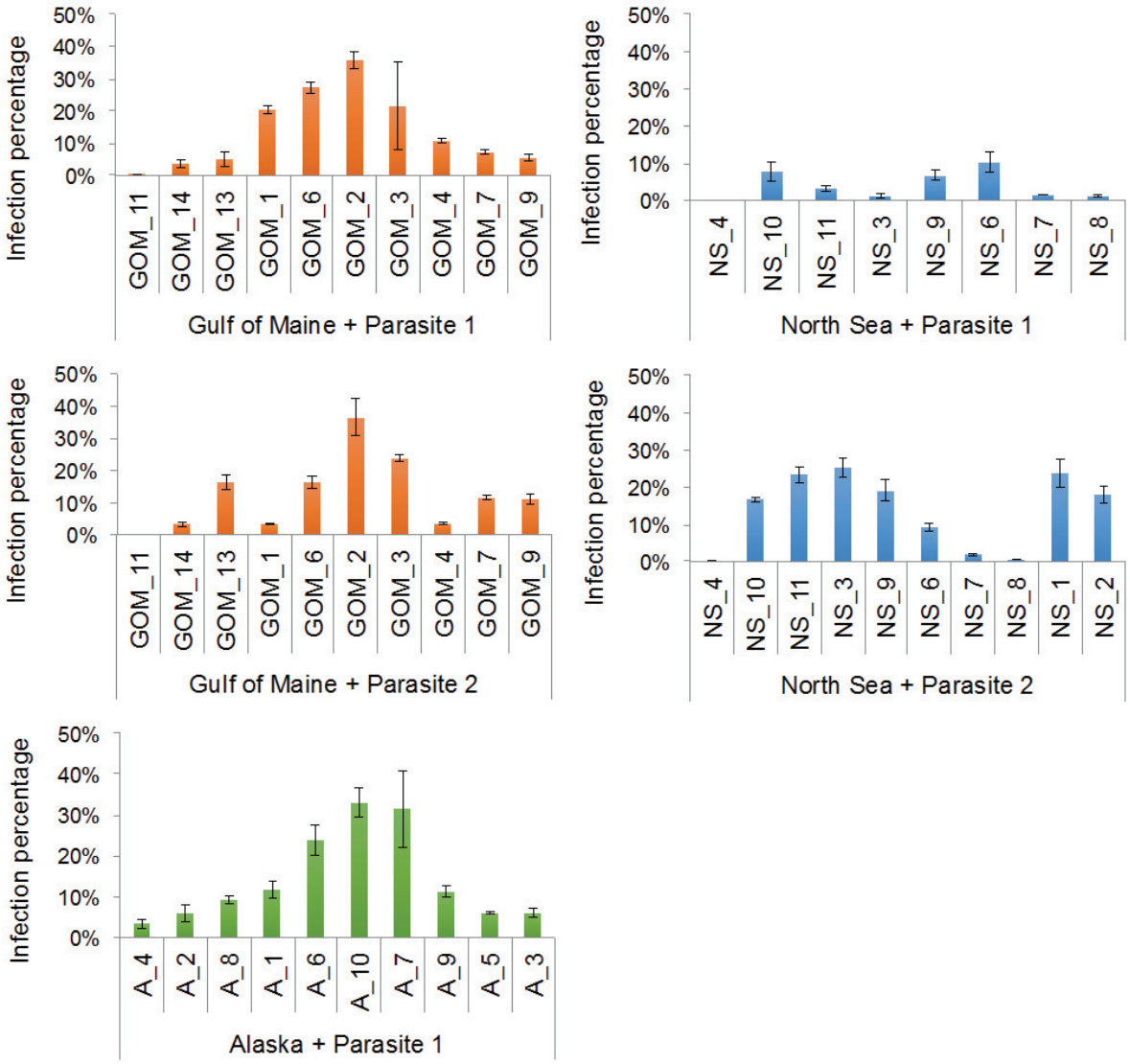


Fig. 4.1.5: Barplots of the infection percentages of *Alexandrium* strains from three geographical locations (Alaska, the Gulf of Maine and the North Sea) exposed to two clones (Parasite 1 and 2) of parasite *Amoebophrya* from with the highest in the middle and the lowest at the sides (n=3, error bars indicate standard errors of the mean).

Cellular toxin content of Alexandrium

Incubation of the Gulf of Maine strains with Parasite 2 caused a significant decline in *Alexandrium* toxin content per cell compared to the control (ANOVA: $P < 0.001$), although no significant differences were observed with Parasite 1 (ANOVA: $P > 0.05$) (Fig. 4.1.6). By contrast, after exposure to Parasite 1 and 2, no significant changes in cellular toxin

concentrations of North Sea strains could be verified (ANOVA: $P > 0.05$). Similarly, no significant change in cellular toxin content of the Alaska strains was observed when exposed to Parasite 1 (ANOVA: $P > 0.05$).

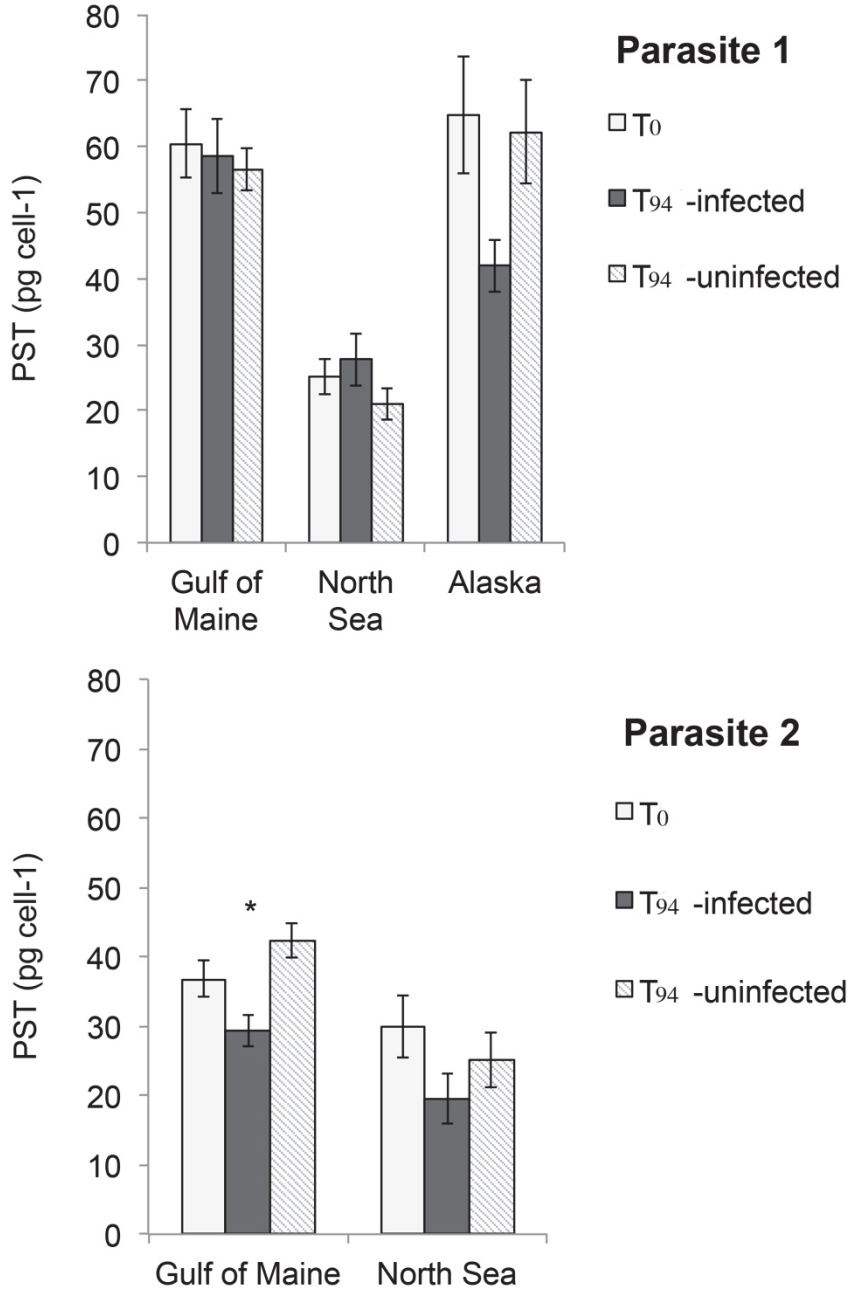


Fig. 4.1.6: Cellular PST content for each geographical population of *Alexandrium* at time 0 (T_0) compared to 94 h after the parasite infection (T_{96}-infected) and the corresponding control at 94 h (T_{94}-uninfected). The asterisk indicates significant differences of PST content inside each geographical population (ANOVA: $P < 0.05$, $n = 30$, error bars indicate standard errors of the mean).

The toxin composition is illustrated in Fig. 4.1.7. In the case of the Gulf of Maine strains the major saxitoxin analogues were C1/C2 toxins (31%), GTX1/4 (26%) and GTX2/3 (23%); neoSTX (10%) and STX (9%) were also present in considerable amounts, but GTX5 made up only a small part (1%). The mean toxin profile of the North Sea strains was considerably different, with C1/C2 toxins comprising nearly half of the toxin complement (47%) and a large portion was made up by GTX1/4 (30%). Conversely, GTX2/3 (3%) and STX (5%) were not as abundant as that in the Gulf of Maine strains; GTX5 (5%) and neoSTX (11%) were produced similar to the Gulf of Maine strains. More than 90% of toxins of Alaska strains were C1/C2 toxins and 5% of the toxins were GTX1/4. In contrast, GTX2/3, GTX5, STX and neoSTX were also produced, but only in small amounts (<1 pg cell^{-1}). The toxin concentrations and profiles show inconsistent on an intra-population level (for a detailed illustration see Figure Appendix 4.1.2).

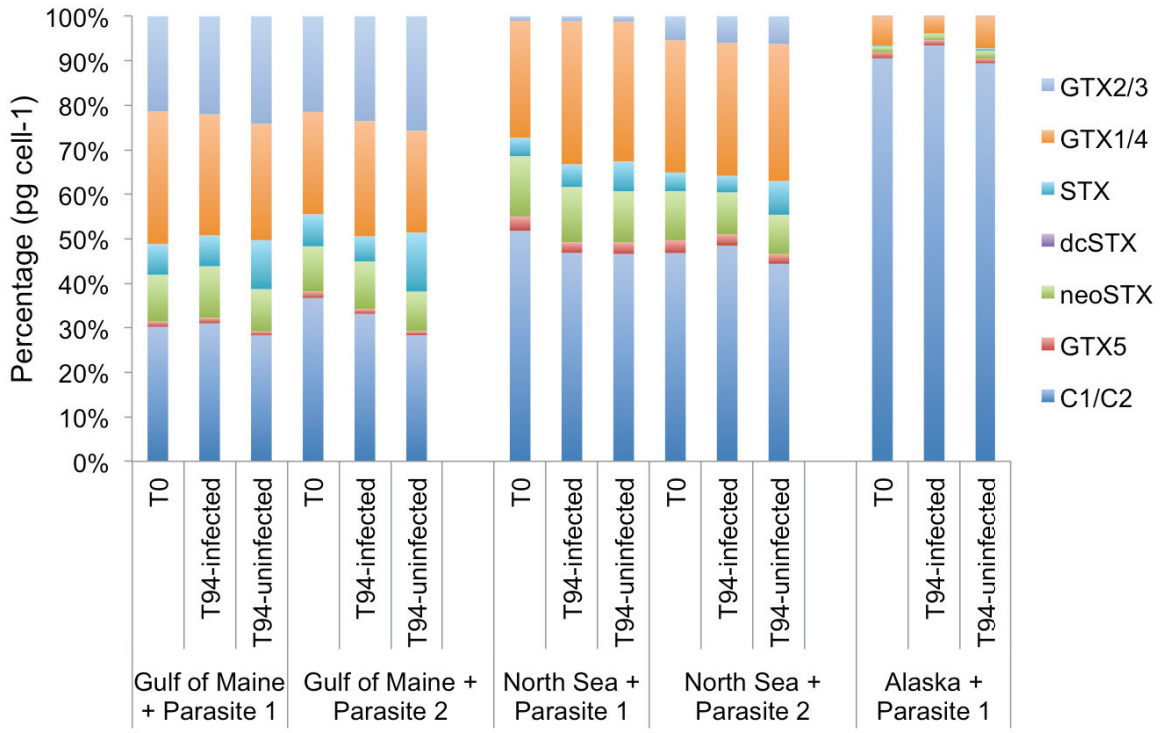


Fig. 4.1.7: Toxin composition variation among representative isolates of three geographical populations of *Alexandrium* at time 0 (T_0) compared to 94 h after the parasite infection (T_{94}-infected) and the corresponding control at 94 h (T_{94}-uninfected).

Regression analysis of toxin concentration

To investigate the potential effects of cellular toxin content (in pg $cell^{-1}$) on growth rate (d^{-1}), lytic rate (cells ml^{-1} h^{-1}) and infection rate, three separate simple linear regression models were fitted. Infection rate was the percentage of infected host cells after 94 h of incubation with the parasite. The independent variable was cellular toxin content, whereas growth rate, lytic rate and infection rate were the outcome variables of the respective models. *P*-values were calculated for the slope parameter, i.e. for the influence of toxin content on the outcome variable of the model. Normal distribution of the residuals was checked with normal-quantile-plots and variance homogeneity was checked with residual plots. Infection rates were transformed with the square root to obtain an approximate normal distribution. Only positive values of lytic rate were used and squared to obtain an approximate normal distribution. In neither case did the cellular toxin content influence the outcome significantly (to a naïve significance level of 0.05). The individual linear models are illustrated in Fig. 4.1.8. *P*-values were higher than 0.1 in all three models and R^2 values were always below 0.02.

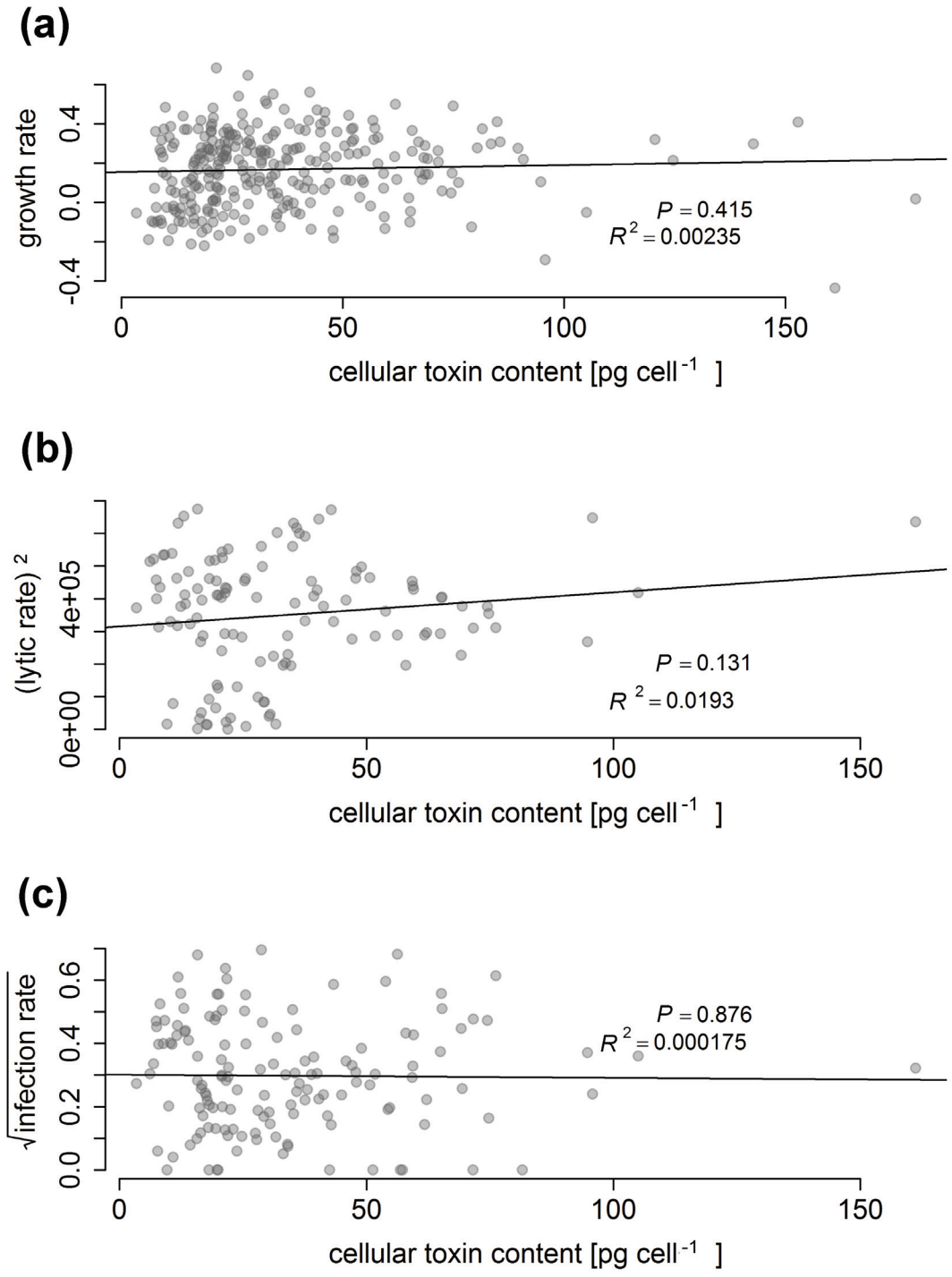


Fig. 4.1.8: Scatterplot of the total toxin content per cell after four days of infection. (a) Toxin content versus growth rate (n=288). (b) Toxin content versus lytic rate (only positive values; values squared to approximate normal distribution of residuals; n=124). (c) Toxin content and infection rate (values transformed with square root to approximate normal distribution of residuals; n=144).

On an intro-population level, the regression lines between toxin concentration and the three outcome variables (growth rate, lytic rate and infection rate) are illustrated for the three geographical populations separately (Fig. 4.1.9). There was a significant (to a Bonferroni corrected significance level of 0.05/9 = 0.005556) influence of toxin content on lytic rate in the Alaska strains ($R^2 = 0.561$, $P < 0.0001$). Furthermore, there were significant influences of cellular toxin content on the infection rate in the North Sea and Alaska strains. In the North Sea strains, the correlation was negative ($R^2 = 0.224$, $P < 0.001$), whereas in Alaska strains it was positive ($R^2 = 0.411$, $P < 0.001$). In all other cases, toxin content did not have a significant influence on the outcome.

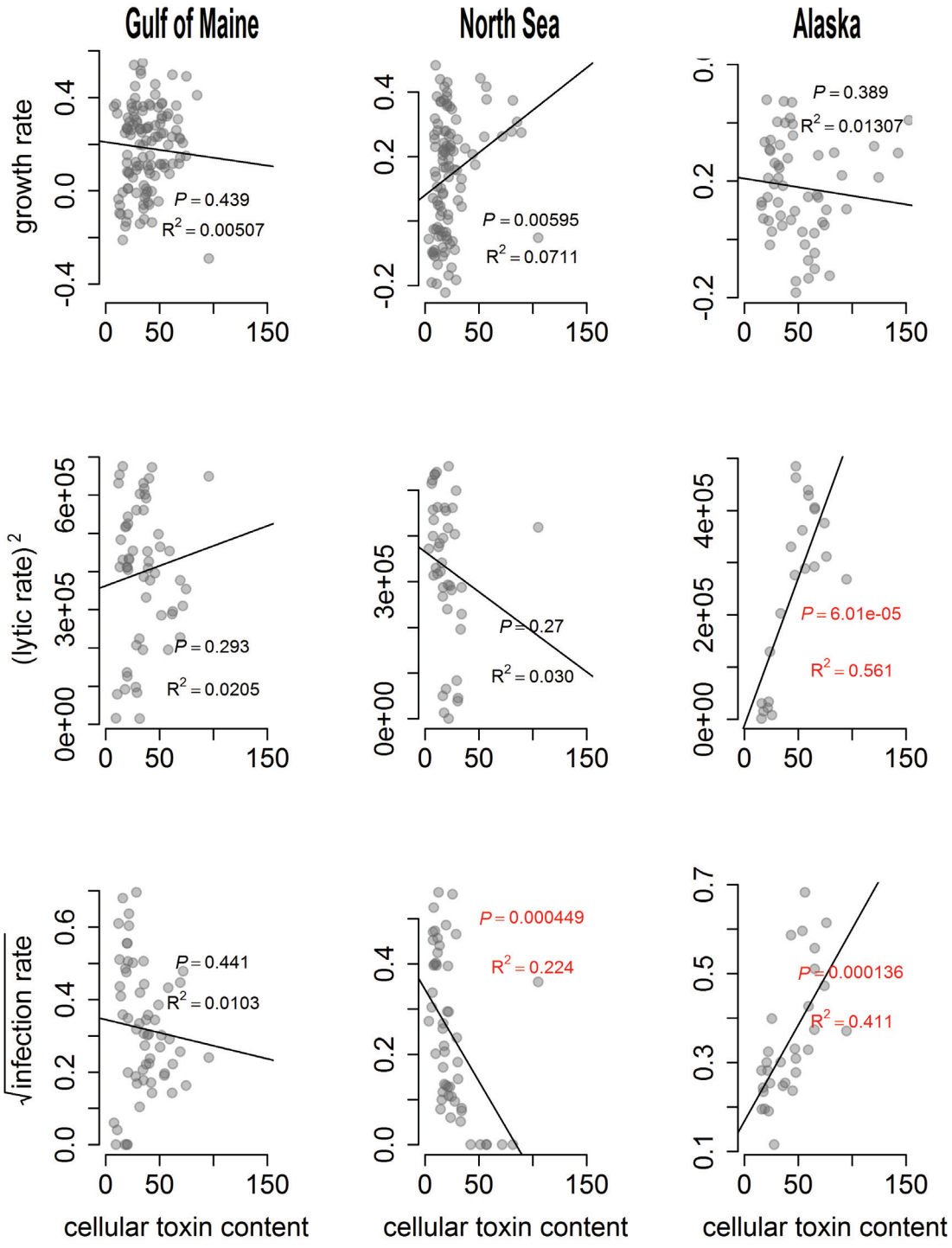


Fig. 4.1.9: Scatterplot of cellular toxin content versus growth rate, lytic rate and infection rate (from top to bottom) in three geographical populations (from left to right: the Gulf of Maine, the North Sea and Alaska) of *Alexandrium* infected by the parasite *Amoebophrya* (n=30). *P*-values smaller than the Bonferroni corrected significance level are marked in red.

Population genetic structure based on microsatellites

The genetic population differentiations estimated with fixation index (F_{ST}) for microsatellites showed that F_{ST} values among each pair of populations were all significant or highly significant. Overall, the three sampling regions displayed strong genetic differentiation among each other (F_{ST} ranged from 0.21 to 0.48). But the realized genetic pattern is unexpected with geographical distances of the three sampling populations. Geographically the nearest populations of North Sea and Gulf of Maine have the highest level of differentiation, whereas closest genetic relationship was found between the geographically farthest populations of North Sea and Alaska (Table 4.1.2).

Table 4.1.2 : Pairwise population differentiation measured as F_{ST} based on nine microsatellite loci. Pairwise F_{ST} values are below diagonal. Level of statistical significance are above diagonal represented by * ($p < 0.05$), ** ($p < 0.01$) or *** ($p < 0.001$).

	Alaska	Gulf of Maine	North Sea
Alaska	0	***	*
Gulf of Maine	0.38677	0	***
North Sea	0.21002	0.48224	0

The furthest genetic relationship between North Sea and Gulf of Maine was also reflected in the result of Bayesian clustering implemented in STRUCTURE. Under the assumption of correlated allele frequencies with admixture model, which hypothesizes that each strain may have mixed ancestry and has inherited some fraction of its genome from ancestors of different populations. The most probable K was 1 according to average log probability ($L(K)$), whereas the *ad hoc* statistic ΔK suggested $K = 2$. This is not a contradiction between both methods of determining best K, because the change in log probability (= ΔK) cannot account for the smallest and largest K (Evanno *et al.* 2005). Even the estimated membership coefficient (Q) of each strain from all sampling populations was largely shared between the putative two clusters (Fig. 4.1.10 a), implying an actual single genetic pool formed by the three geographical populations based on correlated allele frequencies with admixture model. Nevertheless, given the high level of differentiation (F_{ST}) among the three populations, the independent allele frequencies and non-admixture model is more plausible. Under this condition, average log probability ($L(K)$) suggested the most probable K was 3, whereas $K = 2$ was indicated by the highest statistic ΔK, yet $K = 3$ also had a high probability (Figure Appendix 4.1.3). In the scenario of either two or three clusters, all strains from the Gulf of Maine formed a separated cluster having no genetic share with the other two geographical populations (Fig. 4.1.10 b and c). In a two-cluster scenario, the North Sea group also formed an independent genetic cluster despite signs of very little introgression from the Gulf of Maine, while the remote group of Alaska showed unexpected pattern of mixture of the North Sea and Gulf of Maine (Fig. 4.1.10 b). The pattern of mixture for Alaska was still displayed in the three-cluster scenario, whereas all strains of the North Sea group had genetic share with Alaska (Fig. 4.1.10 c), implying historical genetic flow between these two farthest populations and between the western and eastern sides of North America rather than between the two sides of the northern Atlantic.

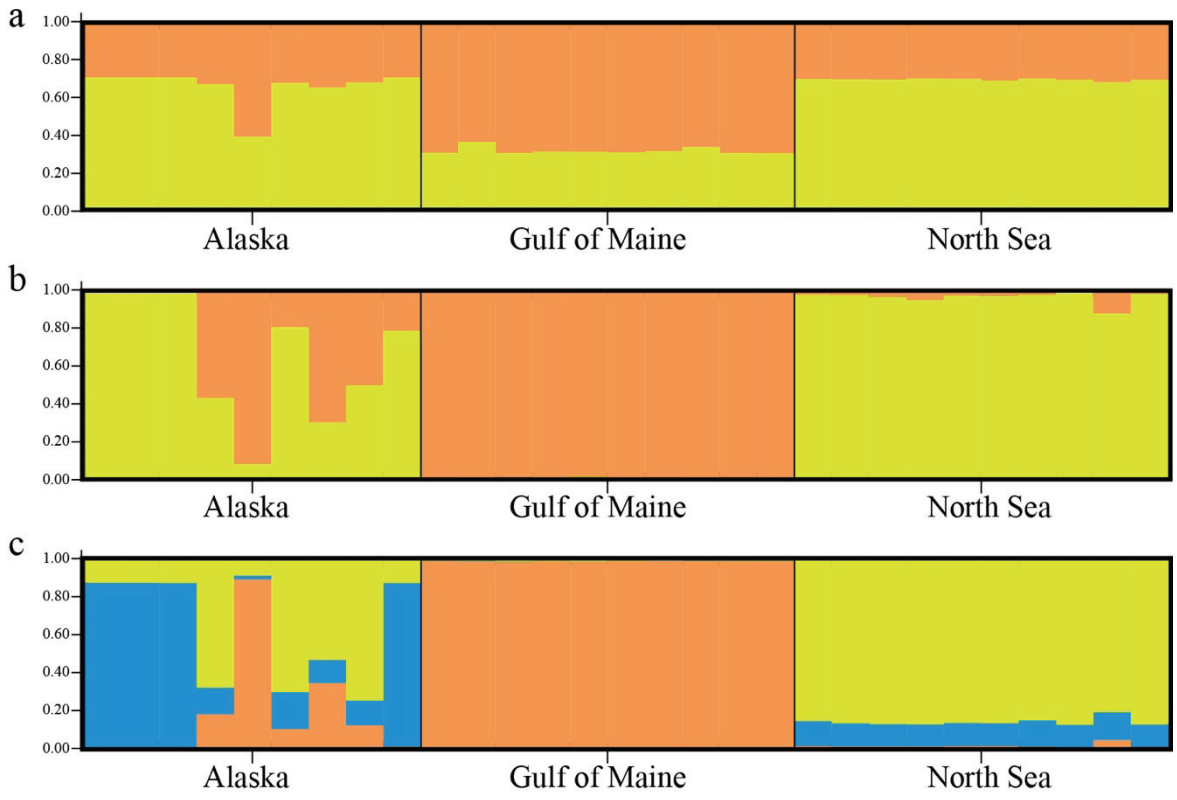


Fig. 4.1.10: Summary bar plot of estimated membership coefficient (Q) of 29 *Alexandrium fundyense* strains from 3 sampling populations (Alaska, the Gulf of Maine and the North Sea) from STRUCTURE analysis. A single vertical bar represents each individual. Under the assumption of correlated allele frequencies with admixture model: (a) $K = 2$ suggested by the *ad hoc* statistic ΔK. (b) $K = 2$ suggested by the *ad hoc* statistic ΔK . (c) $K = 3$ suggested by the average log probability ($L(K)$).

Phylogenetic relationship

The created maximum likelihood tree is presented in Fig. 4.1.11. All strains analysed fell into *Alexandrium fundyense* of the *Alexandrium* species complex.

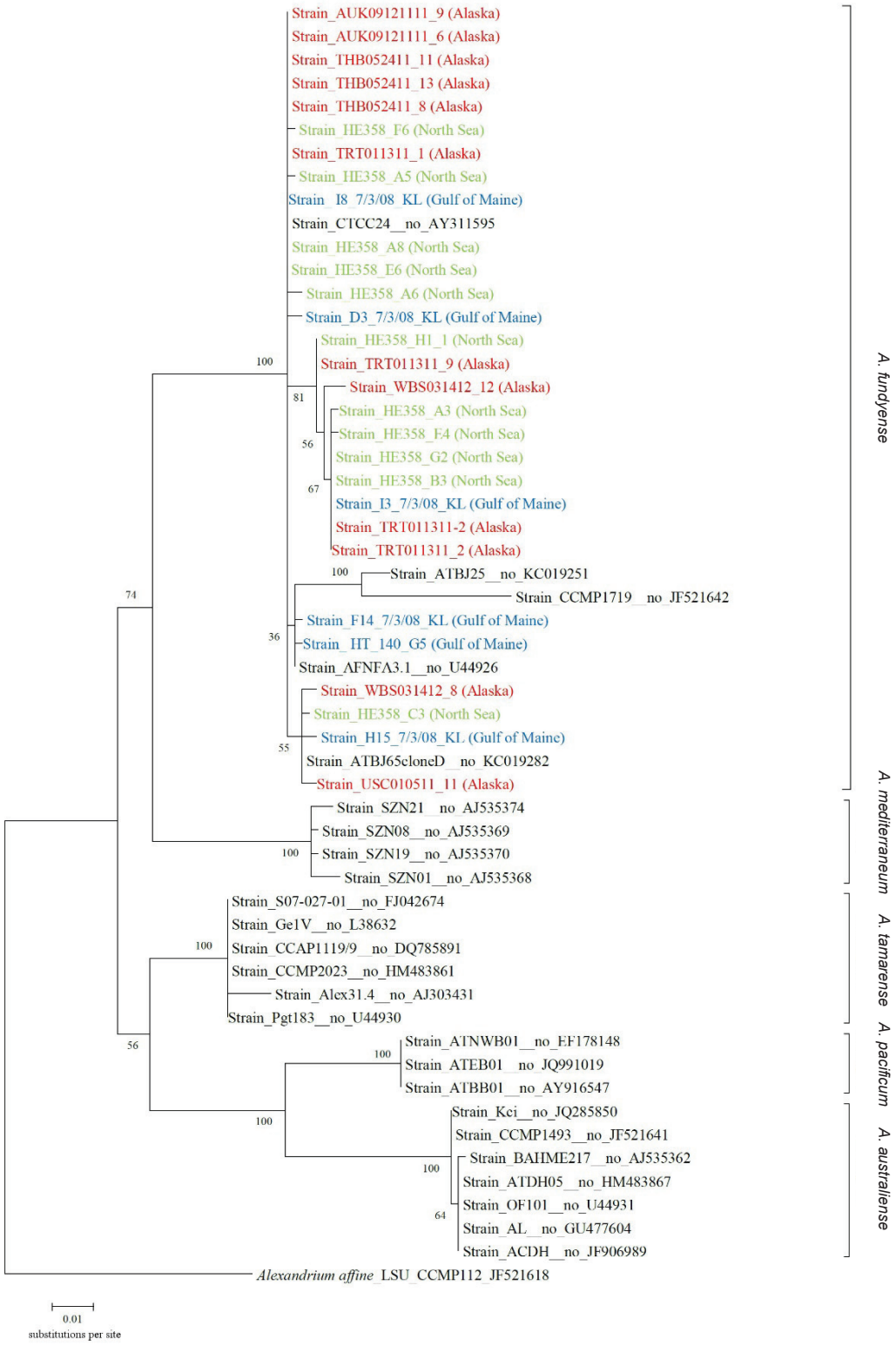


Fig. 4.1.11: Maximum likelihood phylogenetic tree of all members of the Alaska and North Sea strains and six members of the Gulf of Maine strains with reference strains for each of the other groups of the species complex and *A. affine* as an out-group. 1,000 bootstrap replications were done; the alignment contained 613 bp of 54 strains.

4.1.5 Discussion

The present study showed that parasite infection had a strong negative impact on the development of the host populations, but there was a lack of significant difference in parasite prevalence (the percentage of infected cells in the host population) among the three host populations from Alaska, North Sea and Gulf of Maine (Fig. 4.1.4). This result indicated that there was no adaptation of the host on an intra-specific level, which was unlike previous studies of *Alexandrium minutum* strains infected by the parasite *Parvilucifera sinerae,* where host strains showed different susceptibilities to parasite infection depending on the geographical origin of the Mediterranean and Atlantic strains (Figueroa *et al.* 2010a; Garcés *et al.* 2013). In contrast to the little difference in inter-population susceptibilities, significantly different susceptibilities to infection were observed within populations. In Gulf of Maine, strain GOM_2 was most susceptible to infection with both parasite clones, whereas the strain GOM_11 was least susceptible to infection with both parasite clones (Fig. 4.1.5). This observation could be explained that the susceptibility may be universal at least in this study with a relatively small number of strains/isolates and there may be some adaptations in the hosts on the intra-population level.

There was not only no apparent adaptation of the host, but also no specialization of the parasite on the intra-specific level. The host strains from the Gulf of Maine, coming from the same region as the parasites, were more susceptible with significantly higher lytic rates to infection by Parasite 1 than the North Sea and Alaska strains (Fig. 4.1.3). However, the lytic rates of Gulf of Maine strains were not significantly different to infection by Parasite 2. *Amoebophrya* are representative of strains in Syndiniales that vary from extremely species-specific to rather unspecific. Some strains of *Amoebophrya* only infect a unique host strain (Chambouvet *et al.* 2008; Coats & Park 2002), whereas a few strains infect a broader host range from the same or closely related genera (Coats *et al.* 1996; Kim 2006). The present study illustrates differences in the ability of *Amoebophrya* to infect strains from different geographical populations of a single host species.

The growth rates in the infection experiments with Parasite 2 were much lower than that with Parasite 1, where the growth rates of the negative controls were rather similar (Fig. 4.1.2). Furthermore, the lytic rates of *Alexandrium* treated with the Parasite 2 were higher than that with Parasite 1 (Fig. 4.1.3). This may suggest a higher capability of the Parasite 2 in controlling the host population. This is the first time where differences were demonstrated in the ability of multiple clones of parasite within a single species to infect dinoflagellates.

PSP toxins produced by toxic *Alexandrium* species adversely affect metazoan (copepod) grazers and competitors (Selander *et al.* 2006; Wohlrab *et al.* 2010; Yang *et al.* 2010), but do not apparently act as defense compounds against unicellular hetertrophs (Tillmann & Hansen 2009; Tillmann & John 2002). The parasite *Amoebophrya* acts as 'intracellular grazer' and is more similar to those of unicellular protistan grazers than metazoan copepods; it is thus likely that PSP toxins do not serve as defense compounds against protistan parasites in *Alexandrium*. This hypothesis is supported by the absence of change on PSP-toxin content, i.e., there is little effect of parasite infection on the production of PSP toxin in *Alexandrium*. Alternatively, synthesis of toxin by *Alexandrium* could have been inhibited following parasite infection or infected cells may become leaky and have lost synthesized toxin. The suggestion is consistent with the observation that the dinoflagellate *Akashiwo sanguinea* infected by an intra-nuclear species of *Amoebophrya* quickly lost photosynthetic periodicity and circadian rhythm as parasite utilized host resources (Park *et al.* 2002a).

Significantly, the correlation between toxins versus infection percentage in North Sea strains was clearly negative. By contrast, a significantly positive correlation between toxins versus infection percentage remained in Alaska strains. Bai *et al.* (2007) described a positive correlation between karlotoxin concentration in *Karlodinium veneficum* and infection by *Amoebophrya*, although they concluded that the hosts' toxin content may be coupled to other factors, such as an increased cell size, and the correlation did not necessarily reflect a direct connection. A recent study by Kim and Park (2016) examined *Amoebophrya* infecting the

dinoflagellate *Alexandrium fundyense* showed different results that in mean toxin content for infected cultures was significantly higher than that for uninfected cultures, but a requirement of much large number of replicates is also mentioned in their study. Since little is known about the impact of parasitism on *Alexandrium* toxin production and repeated exposures of *Alexandrium* PSP-toxins during parasite infection remain to be explored.

All strains included in this study belong to the monophyletic group of *A. fundyense*, and none of the three geographical populations formed monophyly, which confirms the single specific status of all sampled strains in this study. The population genetic analyses revealed strong differentiation among the sampling regions, suggesting that the geographical distribution of *A. fundyense* reflects the genetic structure distribution. However, the level of genetic relationship among the three populations is not as anticipated as to follow a pattern of geographical distances. The geographically nearest two populations, the Gulf of Maine and North Sea, for which there seems to be no barrier to the dispersal of cells from both populations, have paradoxically the strongest differentiation (measured as F_{ST}) and almost complete absence of gene flow. By contrast, the remote population of Alaska, which is either separated from the North Sea by distance and pathway of cold water, or separated from the Gulf of Maine by the continent, show largely genetic exchange with these two populations. Based on current result a bold hypothesis is that the population of the Gulf of Maine serves as a genetic source, whereas the other two populations act as genetic sink. To test this hypothesis a larger data set (including more strains and more loci) is clearly needed in order to 1) depict a more refined genetic structure since the limited number of strains from each population used in this study may convey an amplified genetic differentiation due to the high variation of microsatellite markers; 2) conduct more comprehensive analyses in revealing directionality of gene flow and ancestral reconstruction.

Caveat

The significantly low parasite prevalence of the North Sea strains infected by Parasite 1 was due to the low starting concentration of the parasite. Dinospores of the parasites were

harvested after 5 days (1 day later than the other treatments), because of the time shortages in the experiment execution. Therefore the infection percentage of North Sea strains and their positive controls were both low. This at least shows the fitness of parasite dinospores decreased quickly after release from the host.

4.1.6 Conclusions

This study demonstrates a strong negative effect of parasitism on the population of the bloom-forming dinoflagellate. Although a selection by parasite is not strong enough to make host populations adapt, the parasite *Amoebophrya* has great potential in control of *Alexandrium* as a biological factor. It is likely there is no effect of the parasite on host's PSP toxin production, and PSP toxins do not serve as defense compounds towards the parasite in *Alexandrium*. Since the present study investigates the effect of two clones of *Amoebophrya* on three host geographical populations, using more samples/strains will pave the way forward for future research.

4.2 Eukaryotic-aerobic life is possible without a mitochondrial genome: The parasitic dinoflagellate *Amoebophrya* sp.

4.2.1 Abstract

Dinoflagellates are a group of microbial eukaryotes that display a number of unusual characteristics in cellular structure, life histories, and genomics. Dinoflagellate genomes are exceptionally large and the DNA is packaged in a unique way using a viral protein and only low levels of histones. Organellar genomes are typically small, fragmented, and contain few genes because most organellar genes have moved to the nucleus. In this study, the relatively small nuclear genome (less than 100Mb) of the basal dinoflagellate *Amoebophrya ceratii*, which is a parasite of other dinoflagellates, was sequenced to further investigate character evolution in dinoflagellates, and parasitic dinoflagellates in particular. Several features were observed in the genome that are consistent with its dependence on a host cell, such as an inability to generate certain amino acids. Notably, tryptophan synthesis is physically interlocked at the genetic level with the shikimate pathway, constituting a novel type of metabolic regulation. *Amoebophrya* has lost its plastid genome and nearly all genes related to the former endosymbiotic organelle. Although mitochondria are observed in *Amoebophrya* at all life stages in electron microscopy, no trace of a mitochondrial genome was found. Genes that have been retained in the minimal mitochondrial genomes of other dinoflagellates were found to have been transferred to the nucleus in *Amoebophrya,* making this the first aerobic eukaryote to exhibit complete loss of the mitochondrial genome.

4.2.2 Introduction

The eukaryote tree of life has been resolved into a small number of major branches (Bachvaroff *et al.* 2014; Baldauf 2008; Keeling *et al.* 2005), one of which comprises the Alveolata; a diverse group of organisms including apicomplexans, ciliates, and dinoflagellates. The genomes of core dinoflagellates are 10-100 times larger than the human genome, and exhibit several unusual features whose evolutionary origins are unclear (LaJeunesse *et al.* 2005). Dinoflagellate chromosomes are permanently condensed in a liquid-crystalline state throughout the cell cycle, but they have only low amounts of histones and additionally use a viral derived protein, and dinoflagellate genes are all expressed with a short leader sequences that is added by trans-splicing (Zhang *et al.* 2007).

The ancestor of dinoflagellates and apicomplexans (and perhaps all alveolate) was photosynthetic (Janouškovec *et al.* 2010), but currently only the basal apicomplexan relatives *Chromera* and *Vitrella*, and approximately half of the Dinoflagellates maintain photosynthesis (Gómez 2012). Furthermore, dinoflagellates and apicomplexans, collectively known as myzozoans, have highly derived mitochondrial genomes that encode only three protein-coding genes and two rRNA genes on one DNA molecule (Hikosaka *et al.* 2010; Jackson *et al.* 2012; Mungpakdee *et al.* 2014; Waller & Jackson 2009). This suite of genes appeared to mark a minimal set for reduced mitochondrial genomes in aerobic species (Flegontov & Lukes 2012), however, recent examination of the respiratory chain in the synthetic *Chromera velia* showed that oxidative phosphorylation complexes I and III were lost, leaving only two protein coding genes (cox I and III) together with fragments of the rRNA genes (Flegontov *et al.* 2015).

Several species of dinoflagellates can produce potent toxins and are able to form harmful algal blooms (HABs) that have enormous impact on ecosystem functions (Anderson *et al.* 2012a; Lewitus *et al.* 2012). The genus *Alexandrium* contains the most prominent HAB-forming dinoflagellates, some of which can form HABs that persist for extended time periods under favorable abiotic and biotic conditions (Anderson *et al.* 2012a; John *et al.* 2015). *Alexandrium* species produce the potent neurotoxins, saxitoxin and its derivates, which are associated with Paralytic Shellfish Poisoning (PSP) (Anderson *et al.* 2012b; Anderson *et al.* 1994). *Alexandrium* blooms therefore have the potential to cause serious human disease and pose economic problems for fisheries.

The dynamics of HABs can be strongly affected by parasites (Chambouvet *et al.* 2008; Mazzillo *et al.* 2011; Salomon & Stolte 2010). In several instances, dinoflagellates are parasitized by other members of the same lineage, although morphological features and their basal phylogenetic position place them outside the core dinoflagellate group, and they have been classified separately (Syndinea and Perkinsozoa, and a few taxa of uncertain position like *Oxyrrhis* and *Psammosa*) (Bachvaroff *et al.* 2014; Saldarriaga *et al.* 2003). The genome of the basal syndinian *Hematodinium* revealed that it likely has secondarily lost plastid organelle (Gornik *et al.* 2015). The Amoebophryidae (Syndinea) is an exclusively endoparasitic family that fall within the marine alveolate Group II, MALVII, clade, a large and diverse group made up mostly of environmental sequences. Amoebophryidae comprises only a single genus, *Amoebophrya* (Guillou *et al.* 2008) with seven described species that exhibit high genetic diversity and play a major part in the small picoplankton assembly in most if not in all field samples studied so far (Alves-de-Souza *et al.* 2012). *Amoebophrya* species were observed to infect populations of the bloom-forming dinoflagellate *Alexandrium* (Chambouvet *et al.* 2011a; Montagnes *et al.* 2008). A high proportion of bloom populations can be infected, and this infection is claimed to affect HAB formation and persistence (Chambouvet *et al.* 2008).

The life cycle of *Amoebophrya* was described more than 40 years ago and was recently examined in more detail using electron microscopy (Cachon & Cachon 1970; Miller *et al.* 2012). The infectious free-living stage, the dinospores, possess two flagella (Fig. 4.2.1). The main function of this non-dividing unicellular form is to search for an amenable host cell. The dinospore subsequently attaches and enters the host cytoplasm, losing its flagella in the process, and becomes enclosed in a parasitophorous membrane. In most cases, the parasite crosses the host nuclear envelope, loosing its parasitophorous membrane in the process. Then, the growing parasite starts to digest its host, increases in size, and eventually forms a so-called beehive structure as a result of several consecutive mitotic divisions. The host cell wall then breaks down and releases the parasite as a short-lived vermiform stage that divides into hundreds of infective dinospores (Coats & Bockstahler 1994; Coats & Park 2002). The maturation of the parasite within the host takes 2–3 days, during which time phases of differential gene expression are distinguishable (Lu *et al.* 2014).

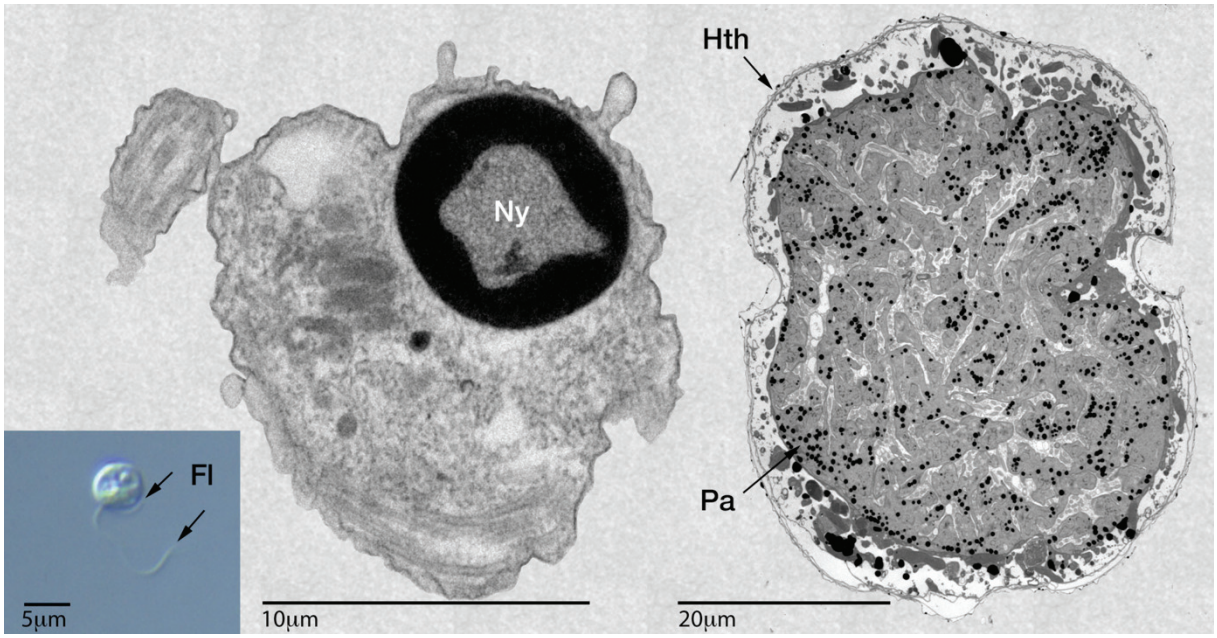


Fig. 4.2.1: Free-living stage and endocellular development of the parasite *Amoebophrya* infecting the toxic dinoflagellate host *Alexandrium fundyense*. Transmission electron micrographs of free-living stage (left) and infected host cell (right, late parasitic development at the beehive stage) The parasite (Pa) fills the complete host cell of which only the Theca remains (Hth). Inset: Light microscopy image of a free-swimming bi-flagellate cell. Fl: Flagellum, Ny: Nucleus, Hth: Host theca, Pa: Parasite.

Here, I describe the complete sequence and analysis of the genome of a strain of *Amoebophrya* that infects the toxic microalga *Alexandrium fundyense* (Lu *et al.* 2014; Lu *et al.* 2016). I find some features that are shared in common with dinoflagellates, but also several unique characteristics that could reflect its basal position or parasitic lifestyle, altogether making this genome an important aid in helping us reconstruct evolutionary transitions in early dinoflagellate evolution.

4.2.3 Materials and Methods

Library preparation and sequencing (Bentley *et al.* 2008) was performed using the Illumina NGS platform and methodology. DNA was extracted from an *Amoebophrya* culture and used to construct paired-end (PE) and mate-pair (MP) libraries for sequencing. Approximately 5 µg of DNA was used for PE library preparation according to the

manufacturer's instructions (Illumina PE sample preparation kit). The library was sequenced in a single lane of a HiSeq2000 device in 100 bp PE mode. Reads were extracted in FastQ format using CASAVA v1.8.2 (supported by Illumina). Sequencing produced 127,251,149 read pairs.

Approximately 3 µg of DNA was used for MP library preparation using a Roche/Illumina hybrid protocol. First, DNA was fragmented using a hydroshear to obtain fragments of around 3 kb in length. Circularization was performed according to the Roche Paired End Library Preparation Method Manual (20 kb and 8 kb Span) using Titanium linkers. Circularized fragments were nebulized and used for library preparation with a TruSeq DNA sample prep kit v2 (Illumina). The library was sequenced in a single lane of a HiSeq2000 device in 100 bp PE Rapid mode. Reads were extracted in FastQ format using bcl2fastq v1.8.3 (supported by Illumina). Sequencing produced 180,661,377 read pairs.

Read data were quality filtered prior to assembly. Reads corresponding to Illumina adapter sequences were discarded (~0.3% of the total). ClC assembly cell (CLCbio) was used to trim reads at a quality cut-off of Q20 and discard reads when remaining read length was <40 bp. Reads were discarded when the respective paired read was discarded for quality/length reasons. For PE data, 85,139,449 read pairs passed the quality criteria. For MP data, 88,499,311 read pairs passed the quality criteria.

Only PE data were used in the first assembly step (using CLC assembly cell). Contigs were blasted against an EST reference/database (Lu *et al.* 2014). Contigs with coverage <30x and >300x which were smaller than 1 kb were discarded. The resulting 9,449 contigs were bridged with MP data using SSPACE2 (Boetzer *et al.* 2011). The bridging procedure produced 4,630 scaffolds.

Sequences from co-cultured bacteria were identified via blastn. After an initial screen of the largest contigs against the complete NCBI nucleotide database (from 30. August 2013) genus-specific databases were set up for each identified contaminant and these were used to re-screen the entire assembly. A further measure of contamination was the coverage combined with the GC content of the contigs (GC content lower than 48 % and a coverage of lower than 30). This way, 87.8 Mb were defined as *Amoebophrya*-specific sequences and 16

Mb were removed as contaminants. The completeness of the assembly was checked using cegma (Parra *et al.* 2007) using default settings.

Augustus (Stanke *et al.* 2004) was employed for gene prediction. A training set was constructed that contained 16 genes previously identified with blast. After adjusting Augustus parameters, the previously-generated EST sequences were used to detect *bona fide* splice sites. In total, 19,952 protein-coding genes were identified. The self-training prediction tool genemarkES (Borodovsky & Lomsadze 2011) was used to evaluate the Augustus predictions. The number of genes predicted with this tool was similar to that predicted by Augustus, and the overlap between the predicted gene sets was extensive.

Predicted ORF sequences from gene models were annotated with the Trinotate pipeline, using blast against NCBI nucleotide data with a cut-off evalue = 10^{-11} [http://trinotate.github.io/]. Annotations were extended by KEGG mapping and were finally manually inspected.

4.2.4 Results and Discussion

Amoebophrya genome

Being an intercellular parasite, the *Amoebophrya* culture was not axenic. After assembly, therefore, I removed contaminating sequences from bacteria or host genomes. Ultimately, 2,352 *Amoebophrya* scaffolds totaling 87.7 Mb remained. The genome coverage is around 110 and the GC content averages to 55.9 %. Gene predictions defined 19,925 protein-coding genes and 39 tRNAs (according to tRNAscan) (Lowe & Eddy 1997). Genome details are shown in Table 4.2.1.

Table 4.2.1: Features of the Amoebophrya genome

Genome		Genes		Introns	
size (Mb)	87.7	predicted CDS	19,925	predicted number	49,619
scaffolds (number)	2,352	average protein length (aa)	653	median size (bases)	184
GC content (%)	54.6%	predicted tRNAs	39	intronless CDS	5,464
		CDS with transcript data	3,714		
		CDS with domain annotation	8,768		

SL-RNA, introns, transcription factors and gene expression

Most, if not all, transcripts of dinoflagellates are trans-spliced to a 22 bp splice leader sequence (SL) (Zhang *et al.* 2007) and appear to be regulated largely at the post-transcriptional level (Slamovits & Keeling 2008a). Individual mRNAs might be processed even from large precursors by trans splicing and polyadenylation. The presence of such SLs associated with coding gene loci indicates the potential for mRNAs to be reintegrated into the genome as intron-less genes after reverse transcription (Slamovits & Keeling 2008a).

Here, the *Amoebophrya* genome was examined for SLs and traces of reintegration events. None of the predicted gene models were associated with a full-length SL motif; however, five gene models were identified that had truncated motifs (SI Appendix, Fig. S.4.2.1; Table S.4.2.1A). This low frequency of SL motifs at gene loci suggests that mature RNA reintegration events are rare or that the traces of integrations are frequently cleaned from the genome by mechanisms such as those observed in some fungi (Ramakrishnan *et al.* 2011). Fifty-three orphan full-length SL motifs distributed across 50 scaffolds were found (SI Appendix, Table S.4.2.1B), and another 713 truncated SL motifs with identities of 73–100%. (SI Appendix, Table S4.2.1C). This distribution differs from the tandem repeat organization found in *Trypasomona,* which also has spliced leaders (Vanhamme & Pays 1995). In the transcriptome dataset, 70 transcripts with single SL motifs were observed (SI Appendix, Fig. S.4.2.1*C*). Only one contig contained a second truncated SL repeat (60% identity to the consensus sequence), and no third or fourth SL repeats were identified.

The predicted gene set contained 49,619 introns in 15,016 genes, and 71.6% of genes had at least one intron. This is substantially fewer genes as compared to the first published dinoflagellate genomes *Symbidinium minimum* and *S. kawagutii* (41,925 and 36,850 genes, respectively) but a rather large number of genes for a parasite (Lin *et al.* 2015; Shoguchi *et al.* 2013). Median intron size was 184 bases (SI Appendix, Fig. S4.2.2A). Besides the canonical U2 donor site, a potential U12 dependent donor motif was also found in the confirmed set (SI Appendix, Fig. S.4.2.3). RNAseq read mapping against the *Amoebophrya* genes revealed no expression enhancement of genes without introns or of mRNAs where I found SL signatures. Our data thus suggest that gene reintegration events are rare in *Amoebophry*a and such events maybe more common in the more complex genomes of core dinoflagellates. Gene expression in trypanosomes seems to be mostly constitutive and thus gene regulation should occur post-transcriptionally (Vanhamme & Pays 1995). Similar mechanisms for dinoflagellates have been proposed, as gene expression was only weakly induced in short term treatments (Jaeckisch *et al.* 2011; Lin 2011) (Morey *et al.* 2011). However, in biotic interactions, such as grazing (Wohlrab *et al.* 2010) and the infection cycle of *Amoebophrya*, gene expression changes were strongly induced (Lu *et al.* 2014; Lu *et al.* 2016). Such inductions are supposedly induced via transcription factors, but so far no comprehensive survey has been done in dinoflagellates (Roy & Morse 2013). I searched the Amoebophrya genome for such factors, and found 60 common transcription factors and 46 proteins with domains corresponding to specific transcriptional regulatory factors. These factors were found also in

other dinoflagellates and alveolates. Thus, while transcription factors are not abundant in dinoflagellates, they are not absent and likely do play a role in adaptation to changing conditions, as is common in eukaryotic gene expression.

Amplified domains

Using the OrthoMCL algorithm (Li *et al.* 2003), all the predicted proteins were clustered into 4,879 families through comparison with complete protein sets from *Perkinsus marinus*, *Phaeodactylum tricornutum*, *Paramecium tetraurelia*, *Naegleria gruberi*, *Reticulomyxa filosa*, *Arabidopsis thaliana*, and *Dictyostelium discoideum*. Of these, 499 proteins clustered to 12 different families associated with transposon domains (SI Appendix, Table S4.2.3), indicating high transposon activity in this genome. Other prominent categories included proteins involved in carbohydrate binding or degradation (157 proteins) and proteins involved in detoxification (44 proteins). However, most of the amplified gene families exhibited no recognizable domains and their functions remain unknown (SI Appendix, Table S4.2.3). The amplification of genes for proteins with domains involved in breakdown of macromolecules is likely associated with the parasitic life style of *Amoebophrya*.

Metabolism of the parasite

Pathogens and parasites frequently use the resources of the host to obtain the basic organic building blocks required for their own metabolism. This close "partnership" often leads to losses of certain synthesis pathways in the parasite. Such losses preclude the organism from sustaining activity for anything other than short periods of time without close contact with the host. In this study, I scrutinized the metabolic abilities of *Amoebophrya* to determine any potential host dependencies. However, the *Amoebophrya* genome contained the genes for most of the common metabolic pathways. For example, in contrast with Plasmodium species, which cannot synthesize purines (Cowman & Crabb 2002), the *Amoebophrya* genome encodes complete purine and pyrimidine synthesis pathway components (SI Appendix, Fig. S4.2.3; S4.2.4). In addition, the *Amoebophrya* genome encoded more than one enzyme for several steps in these pathways. This redundancy was also observed for the fatty acid degradation and amino acid biosynthesis pathways. The composition of the latter two pathways more closely resembled that of autotrophs than heterotrophs, indicating a requirement for enhanced metabolic activity within these pathways (SI Appendix, Fig. S4.2.3).

Amino acid synthesis and shikimate pathway

The *Amoebophrya* genome contained a full set of genes needed to synthesize all amino acids with the exception of histidine and two of the three aromatic amino acids. Enzymes that would enable the production of L-histidine from phosphoribosylpyrophosphate were not found in the *Amoebophrya* genome. This is similar to the *Simbiodinium kawagutii* genome (Lin *et al.* 2015). The components of the tyrosine and phenylalanine synthesis pathways were also absent. However, components for the synthesis of tryptophan via the shikimate pathway (Fig. 4.2.2) were encoded in the genome. The shikimate pathway produces precursors for aromatic amino acids and secondary metabolites to which approximately 20% of photosynthetically fixed carbon is directed in vascular plants (Herrmann 1995). Generally, the complete shikimate pathway consists of seven components, five of which (units aroB, A, K, D, and E; following the notation system in *E.coli*) were fused in the last common ancestor (LCA) of all eukaryotes (Richards *et al.* 2006). In *Amoebophrya*, aroC was fused to the N-terminus of this five-domain array to form a single large protein (g6770; Fig. 4.2.2A). This conformation has not been observed for any other organism present in the Pfam database. Interestingly, the seventh component, aroG, was found to be fused to the multifunctional tryptophan synthesis gene (g13589; Fig. 4.2.2B), so that the shikimate and tryptophan pathways are uniquely physically linked in this species. Normally, the shikimate pathway also provides material for the synthesis of tyrosine and phenylalanine. Since these pathways were not present in *Amoebophrya*, the shikimate pathway appears to be required for the synthesis of tryptophan only. This direct “hard-wired” linkage of the shikimate and tryptophan pathways may constitute a simple mechanism to ensure the concerted expression and output of the two pathways.

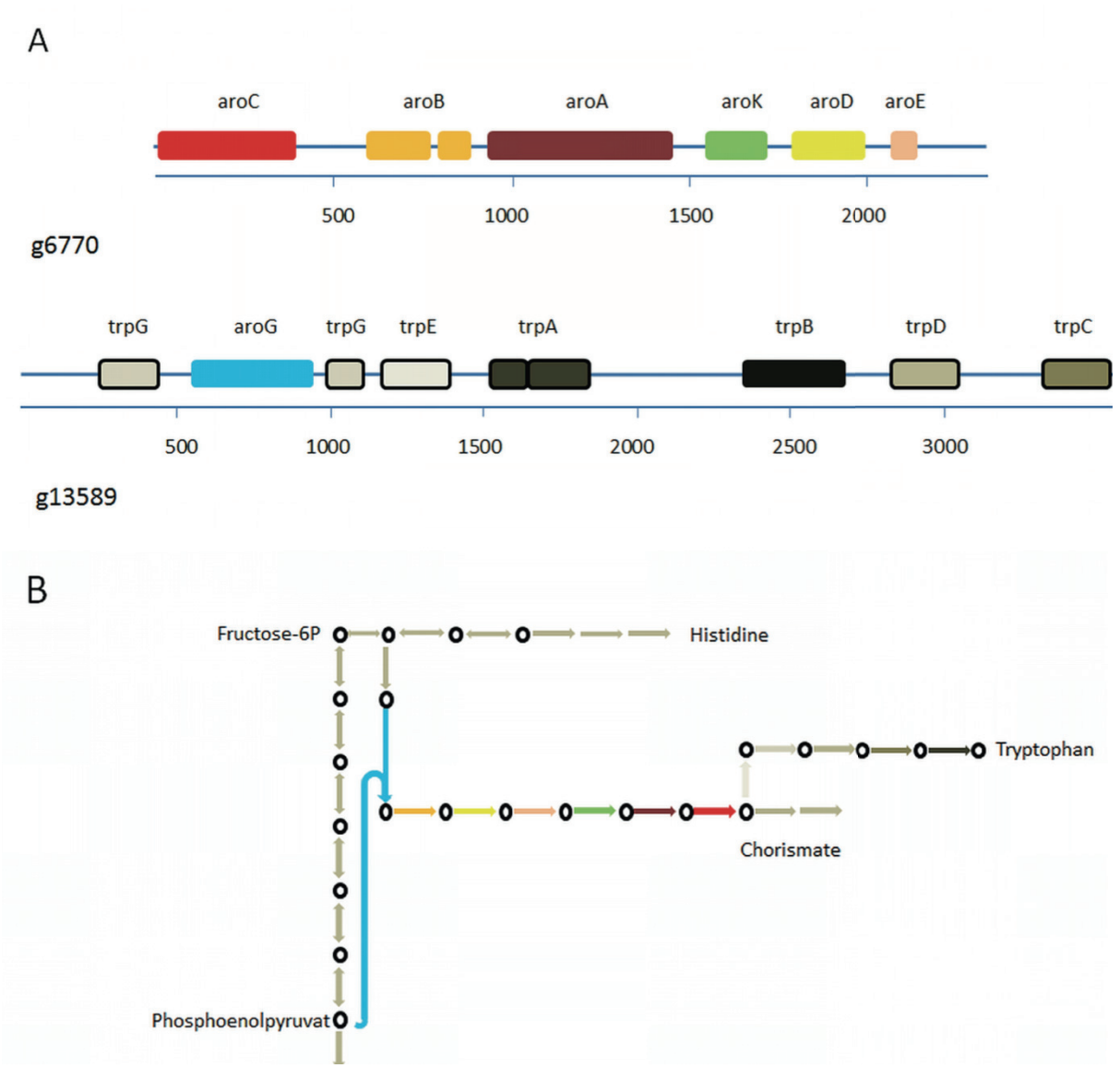


Fig. 4.2.2: Shikimate (g6770) and tryptophan (g13589) synthesis pathway multidomain genes. A: Individual domains of the shikimate pathway are illustrated with colored boxes and domains of the tryptophan pathway are represented with differently shaded gray boxes. B: Schematic view of the biosynthetic pathway for tryptophan in *Amoebophrya*. Circles represent intermediates that can be synthesized in *Amoebophrya* and arrows indicate the respective enzymatic activities. Arrows without circles indicate missing pathway components in *Amoebophrya*. The colors for the shikimate enzymatic activities are as in A. For simplicity, all tryptophan pathway steps are depicted in gray.

Fatty Acid (FAS) and Polyketide Synthases (PKS)

In photosynthetic organisms, fatty acid synthesis in the plastid is carried out by the ancestral, cyanobacterium-derived multienzyme complex type II FAS (Kohli *et al.* 2016). The ancestral eukaryote also possessed a single modular multidomain enzyme complex type I FAS

that functions in the cytosol. Apicomplexa, having a reduced plastid, can contain both type I and type II FAS or only one of the two. Similar to *Hematodinium* (Gornik *et al.* 2015), *Amoebophrya* has only the type I FAS complex, and they do not have plastid type II FAS (Fig. S4.2.4). However, both *Hematodinium* and *Amoebophrya* also have a PKS type I gene which is evolutionarily is closely related to type I FAS, but primarily involved in secondary metabolisms such as toxin production in many dinoflagellates (Kohli *et al.* 2016).

Mitochondrial genome and function

Apicomplexans, *Perkinsus marinus*, the syndinian *Hematodinium*, and core dinoflagellate mitochondrial genomes encode only three protein-coding genes (Jackson *et al.* 2012), generally found as repeats either in tandem, on short DNA fragments or as circular fragments. It is hypothesized that the genome was reduced and reorganized in the common ancestor of the Myzozoa (Apicomplexa and Dinoflagellates). In core dinoflagellates, these genes also require extensive RNA editing to reconstitute an open reading frame (Waller & Jackson 2009). The assembled *Amoebophrya* genome and unassembled reads were both searched for mitochondrion-encoded genes using mitochondrial sequences from *Plasmodium falciparum*, *Alexandrium tamarense*, *Oxyrrhis marina*, *Tetrahymena thermophile*, *Hematodinium*, and *Perkinsus marinus*, but the short contigs typical for Myzozoa mitochondrial genes could not be found. In electron microscopy two mitochondria are generally observed, and even if each possessed only a single genome, the coverage of these sequences should exceed that of the nuclear genome and they would be readily identified. Further examination showed that, although the expected short contigs were absent, that two genes encoding parts of coxI reside on different, longer scaffolds (scaffold46 and scaffold1091) with the same coverage as other nuclear scaffolds and with genes in the neighborhood with clear nuclear origin. The two coxI fragments together constitute the C-terminal part of coxI, but the N-terminal part was absent from our assembly. I found spliced transcripts of each of these two fragments in the transcript data and confirmed those via PCR from genomic DNA and cDNA. Thus, both fragments are transcribed and spliced individually to form short ORFs covering positions 323–370 (g15932) and 390–444 (g833) of the Pfam domain PF00115. There was no indication in the transcriptome data that the two mRNAs would be united via trans-splicing. In addition, no transcript was observed that would constitute the missing N-terminal section of coxI. It is therefore not yet clear whether the N-terminal part of coxI is no longer needed or whether it has been replaced by a different gene with similar functions. Another possibility is that this part of the gene may be present,

but that it was missed in the genome assembly and transcriptome data. The transfer of other cox gene fragments from the mitochondrial genome to the nucleus has been previously observed in other organisms (Adams & Palmer 2003), but to date cox1 has not been seen to be nucleus-encoded. In other cases, for example the movement of cox2 to the nucleus in apicomplexans, dinoflagellates, and green algae, it has been hypothesized that the transfer was made possible by the fragmentation of the gene. In this case the full length gene possesses too many hydrophobic trans-membrane domains to be readily targeted to the mitochondrion, the fragments individually can be targeted (Gawryluk & Gray 2010; Pérez-Martınez *et al.* 2001; Waller & Keeling 2006).

The other expected mitochondrial genes are cox3 and cytb, and fragments of the rRNAs. Cox3 has also been lost in Chromera, so I looked more intensively for cytb and the mitochondrial rRNA genes, but found no match in the genome assembly. Attempts to amplify parts of cytb via PCR on gDNA or cDNA using degenerate primers derived from conserved sites failed (Lin *et al.* 2009). The whole predicted proteome of *Amoebophrya* was then examined for the presence or absence of the five complexes constituting the respiratory chain normally found in mitochondria. Complexes 1 and 3 were not found, but all essential components of complexes 2, 4, and 5 are present (Fig. 4.2.3, SI Appendix, Table S4.2.5). Complex I (NADH dehydrogenase) is absent from all apicomplexans and dinoflagellates and is compensated via an alternative NAD(P)H dehydrogenase (also found in *Amoebophrya*, g180) (Danne *et al.* 2013). But the absence of complex III was only otherwise observed in *Chromera velia* (Flegontov *et al.* 2015). Electrons could enter the Q pool over the alternative NAD(P)H-DH and the succinate dehydrogenase. As the next step the electrons are transferred to cytochrome C and from there to complex IV. The dinospores can survive approximately four days without a host while searching for a new host. The energy needed during this time (e.g. for swimming with flagella) is probably synthesized by this rudimentary oxidative phosphorylation chain (see model Fig. 4.2.3). Interestingly, none of the identified components of the respiratory chain is predicted to encode a mitochondrial targeting transit peptide, according to SignalP (Emanuelsson *et al.* 2007). This might be due to incomplete protein coding gene prediction, or because leaders in this organism are difficult to predict, but could also point to a mechanism by which sequences encoding targeting signals could be attached via trans-splicing. Overall, the genomic data indicate that *Amoebophrya* no longer retains a mitochondrial genome, making it the first described aerobic species to to have completely lost it. As only an incomplete coxI gene was found, it is possible that the

respiratory chain is not fully functional and that the generation of ATP is solely achieved by substrate phosphorylation. Conversely, a complete complex V is present, which indicates that a functional proton gradient is available for ATP generation. In either case, the genome loss appears to be the result of both the loss of some of the few remaining genes and the transfer of others, or functionally critical fragments of them, to the nucleus.

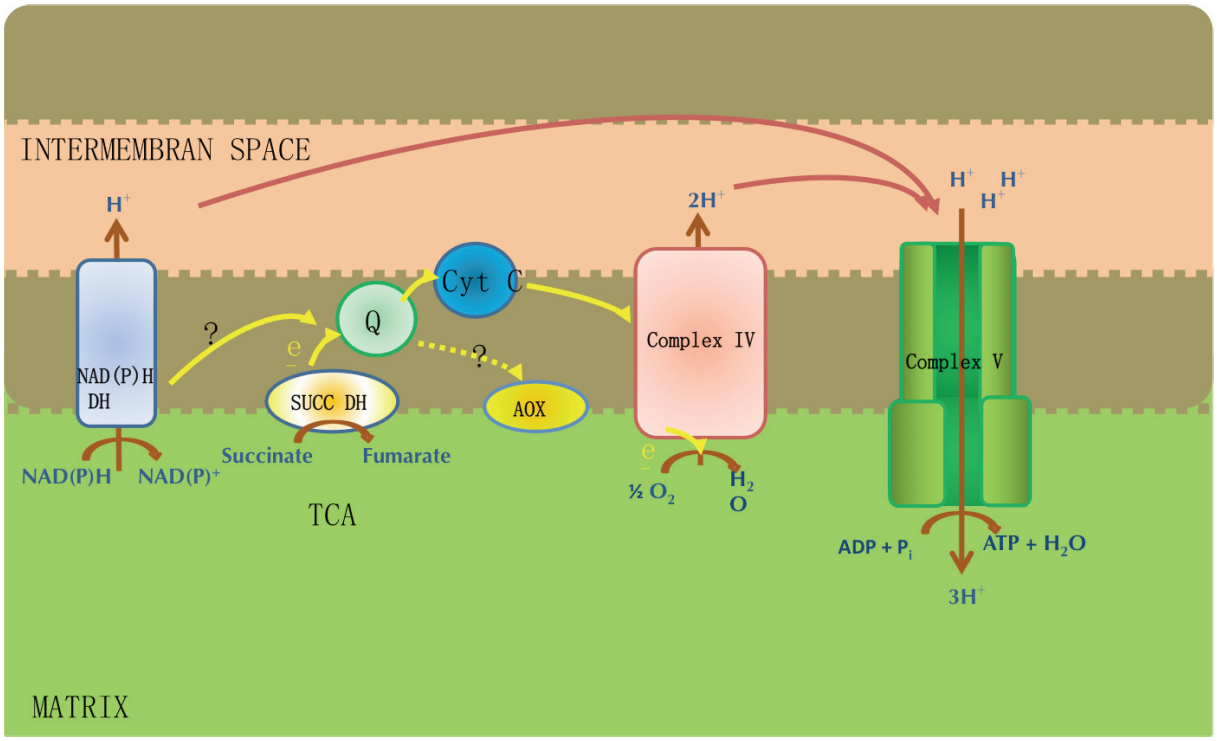


Fig. 4.2.3: Model of mitochondrial function in *Amoebophrya* based on the genome gene content.

Multigene phylogeny and relationships among early-diverging dinoflagellate lineages

Accurate character evolution requires a detailed, completely resolved phylogeny of the evolutionary lineage in question. I therefore used a concatenated set of 100 conserved proteins to construct a maximum likelihood tree of dinoflagellate and their relatives (Fig. 4.4.4). In this tree, *Amoebophrya* branches after *O. marina* and *P. marinus*, closely to *Hematodinium*, as expected based on previous analyses including a recent tree based on concatenated ribosomal proteins (Bachvaroff *et al.* 2014).

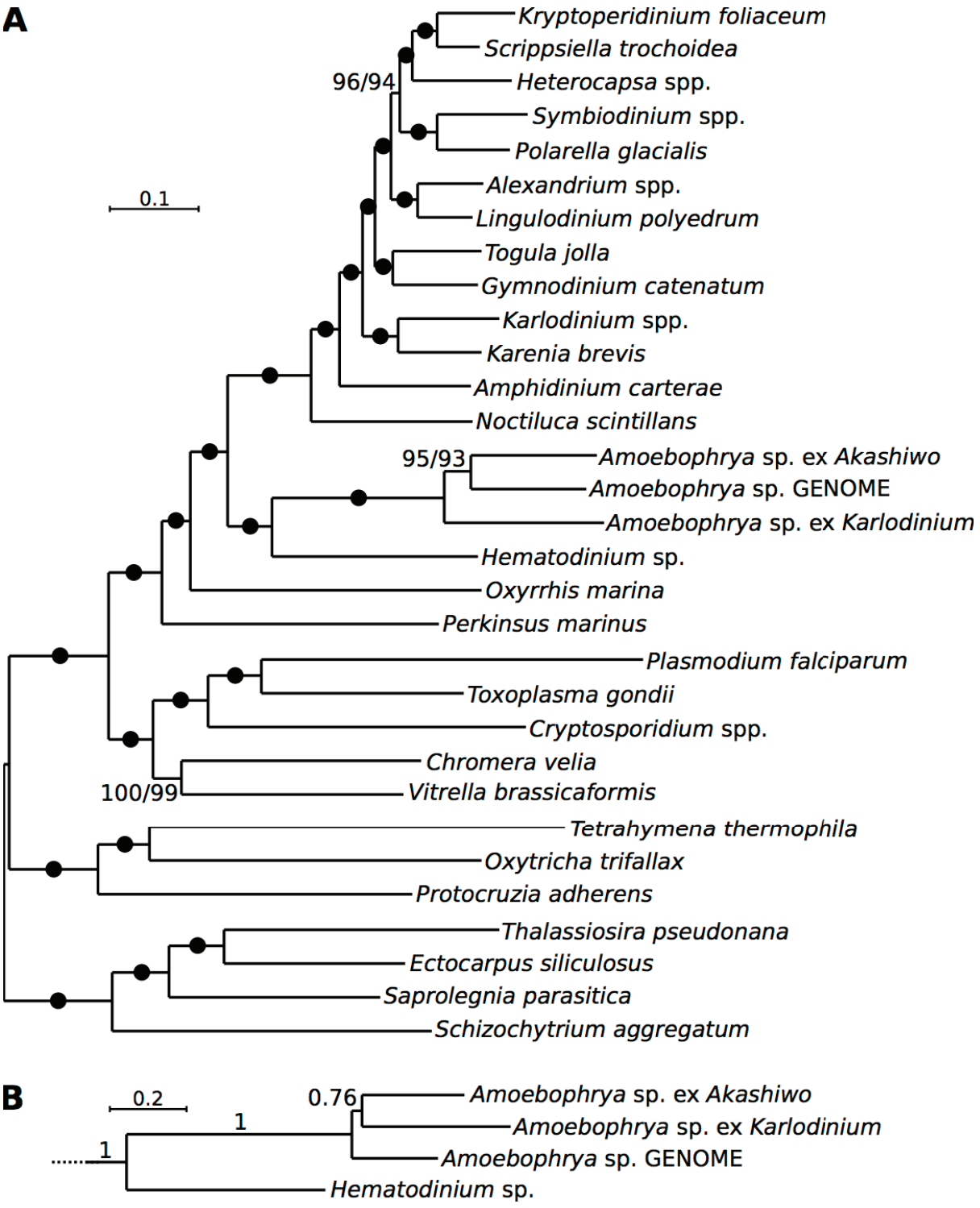


Fig 4.2.4: Multiprotein phylogeny of *Amoebophrya*, dinoflagellates, and their relatives. (A) The best Maximum Likelihood tree (IQ-Tree) under LG+G4+I+F model with ultrafast / nonparametric bootstrap supports at branches (black circles denote 100/100 support). (B) The relationships among *Amoebophrya* species are different in a Phylobayes GTR+CAT+G4 inference; the rest of the tree is identical to (A) and fully supported at all branches.

Genes of potential chloroplast origin

The ancestor of all myzozoa is inferred to have harbored a plastid, and today most species are either still photosynthetic or known to retain a reduced, non-photosynthetic plastid. Neither *Hematodinium* (Gornik *et al.* 2015) nor *Amoebophrya* has been shown to contain a relict plastid organelle by microscopy (Miller *et al.* 2012). In *Oxyrrhis marina*, eight genes were found that had a likely plastid origin (Slamovits & Keeling 2008a), indicating that this species once possessed a plastid, although it has never been observed directly. *Hematodinium* has more recently been shown to have completely lost its plastid organelle, but has retained few gene of plastid origin (Gornik *et al.* 2015). The *Amoebophrya* genome was searched for the presence of genes of plastid origin (as defined from the analysis of *O. marina* and *Hematodinium*). Homologues of three genes also identified in *Oxyrrhis* (DXR, KARI, and RPI) were found in *Amoebophrya*. In *Hematidinium*, the heme biosynthesis pathway is made up of some genes of plastid origin and other genes of cytosolic origin, and in *Amoebophrya* I found cytosolic homologues of porphobilinogen synthase (HemB), porphobilinogen deaminase (HemC), and uroporphyrinogen III decarboxylase (HemE), but also found plastid-derived homologues of uroporphyrinogen III synthase (HemD) and coproporphyrinogen oxidase (HemF), according to phylogenetic reconstruction (SI appendix Fig. S4.2.5 A-E). In none of these is there evidence for plastid targeting via a recognizable signal peptide and transit peptide. As suggested by Gornik *et al.*(2015), this reallocation of plastidial genes to the cytosol argues for a secondary loss of the plastid organelle in *Hematodinium* and for *Amoebophrya*.

Plastid-derived genes were also sought by classification of all predicted proteins using BLAST and calculating phylogenetic origin. Potential “photosynthetic” proteins were further scrutinized with refined alignments and phylogenies were reconstructed using several matrices, but yielded no convincing candidates for a gene derived from the plastid, indicating a nearly complete loss of all plastid genes.

4.2.5 Conclusions

Parasite genomes are often thought to be characterized by a reduction of genome size, loss of duplicated genes, and loss of function as the parasite becomes more dependent on the host. The *Amoebophrya* genome has virtually none of these characteristics. It is missing some pathways relative to other genomes (e.g., some aerobatic amino acids), but not more than one might expect from comparing two heterotrophic lineages. It harbors extended gene families that encode degrading enzymes, and has other unique characteristics, such as the physical linkage between functionally related shikimate and tryptophan synthesis pathways, which may constitute a hitherto unknown mechanism of pathway regulation. The lack of 'loss' in this parasite is interesting because it reflects the challenges I face in inferring ecological roles from the genome of heterotrophic eukaryotes: with even this complete genome it would be impossible to infer that *Amoebophrya* was a parasite in the absence of other biological information about its nature.

The *Amoebophrya* genome does exhibit some intriguing and unique evolutionary features as well. Most noteworthy is the complete loss of the mitochondrial genome, which has not been observed in an aerobic organism previously. Ancestrally, the mitochondrial genome of the myzozoan lineage was reduced to three protein coding genes (cox1, cox3, and cytb), which encode three proteins that appear recalcitrant to relocation to the nucleus. In some species this compliment has been reduced by the loss of an electron transport chain complex that makes the gene non-essential. *Amoebophrya* has taken this one step further: by loosing electron transport chain complexes. The other two, cox1 and cob, seem recalcitrant to relocation or loss. *Amoebophrya* has managed to loose its entire genome due to a mixture of factors. Electron transport chain complexes I and III are lost making cox3 and cytb non-essential, whereas cox1 has been retained, but moved to the nucleus and fragmented, much like cox2. The mitochondria in *Amoebophrya* likely remain functional, as indicated by the presence of a similar oxidative phosphorylation chain as is found in *Chromera velia,* but because of this combination of reduction and novel relocation to the nucleus, the entire genome and its maintenance and expression (including the rRNA gene fragments) have been eliminated.

4.2.6 Supporting information

A

B

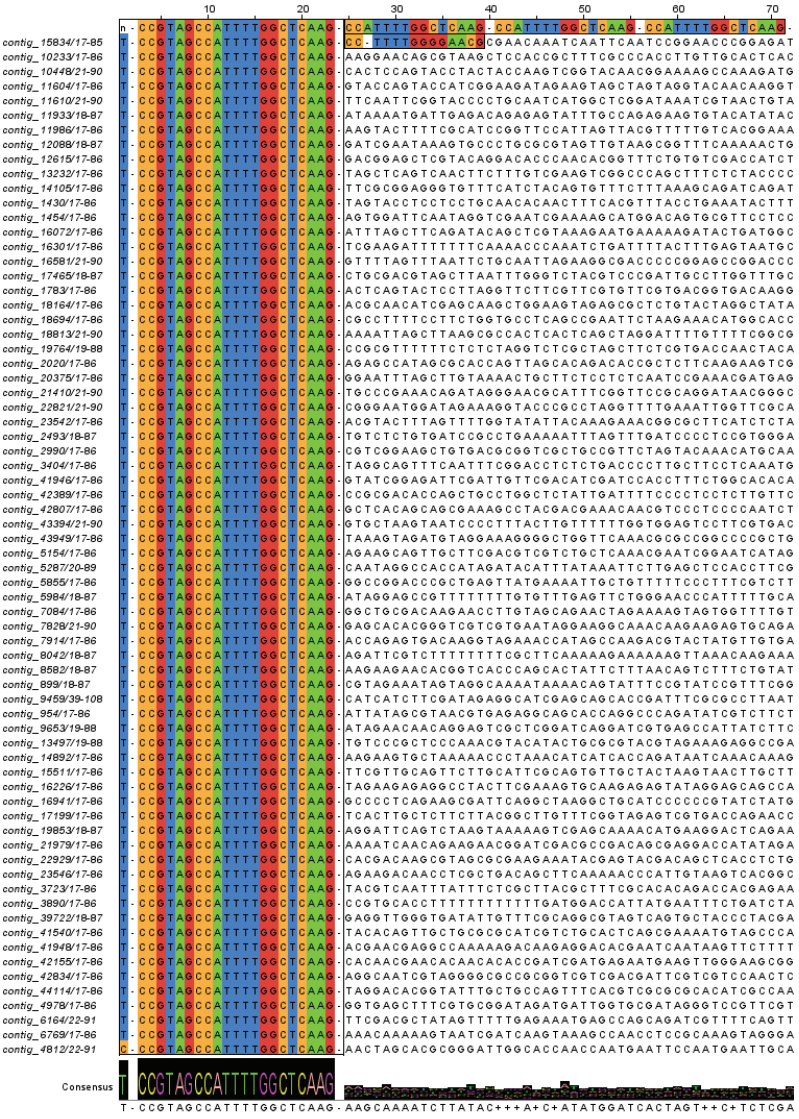


Fig. S4.2.1: Alignment of 5' end of transcripts of *Amoebophrya* showing spliced leader (SL) and relict SL-repeats in (*A*) *Amoebophrya* gene models and (*B*) transcriptome dataset. The first sequence shows a reference SL consensus sequence. The generated consensus sequence and sequence logo are aligned under the transcripts.

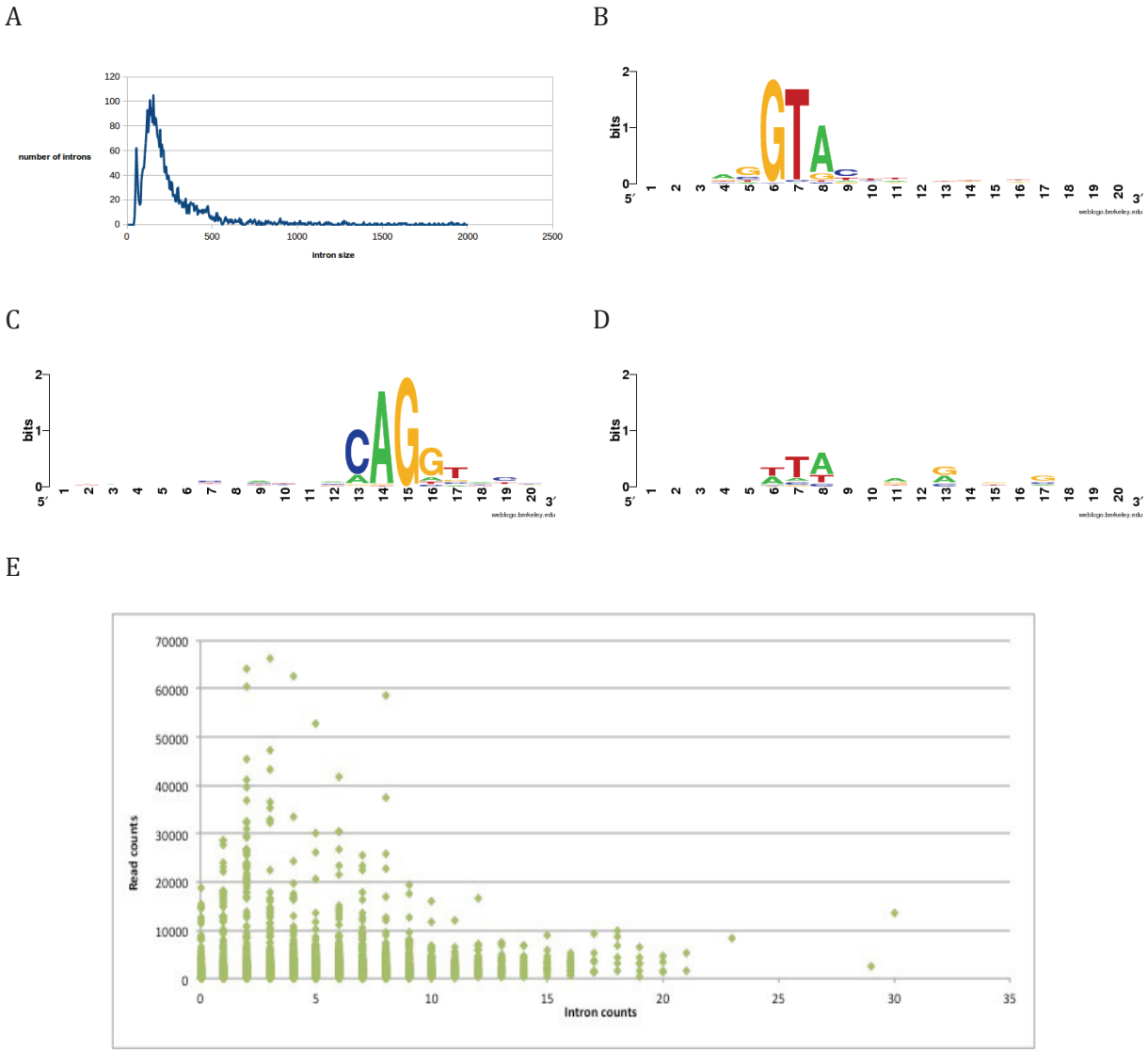


Fig. S4.2.2: (A) The distribution of intron sizes in *Amoebophrya* genes. Only introns confirmed by transcriptome data were taken into account to calculate donor-acceptor distances irrespective of gene models The length of introns varies around 100 bases. A TopHat (Version 2.0.1) (Trapnell *et al.* 2009) mapping of reads on the assembled scaffold data to resolve junction sites independently confirms the gene-model based estimate of 49.000 introns in total. I defined consensus donor and acceptor sites of introns confirmed by transcripts: (B) U2 donor; (C) acceptor; (D) likely U12 donor site. Exon intron boundaries were centered on the splice site (donors: position 5/6; acceptor: position 15/16). Interestingly, some introns seem to be U12 dependent introns. (E) Expression level dependency on intron number. RNAseq read counts were plotted against intron numbers per gene.

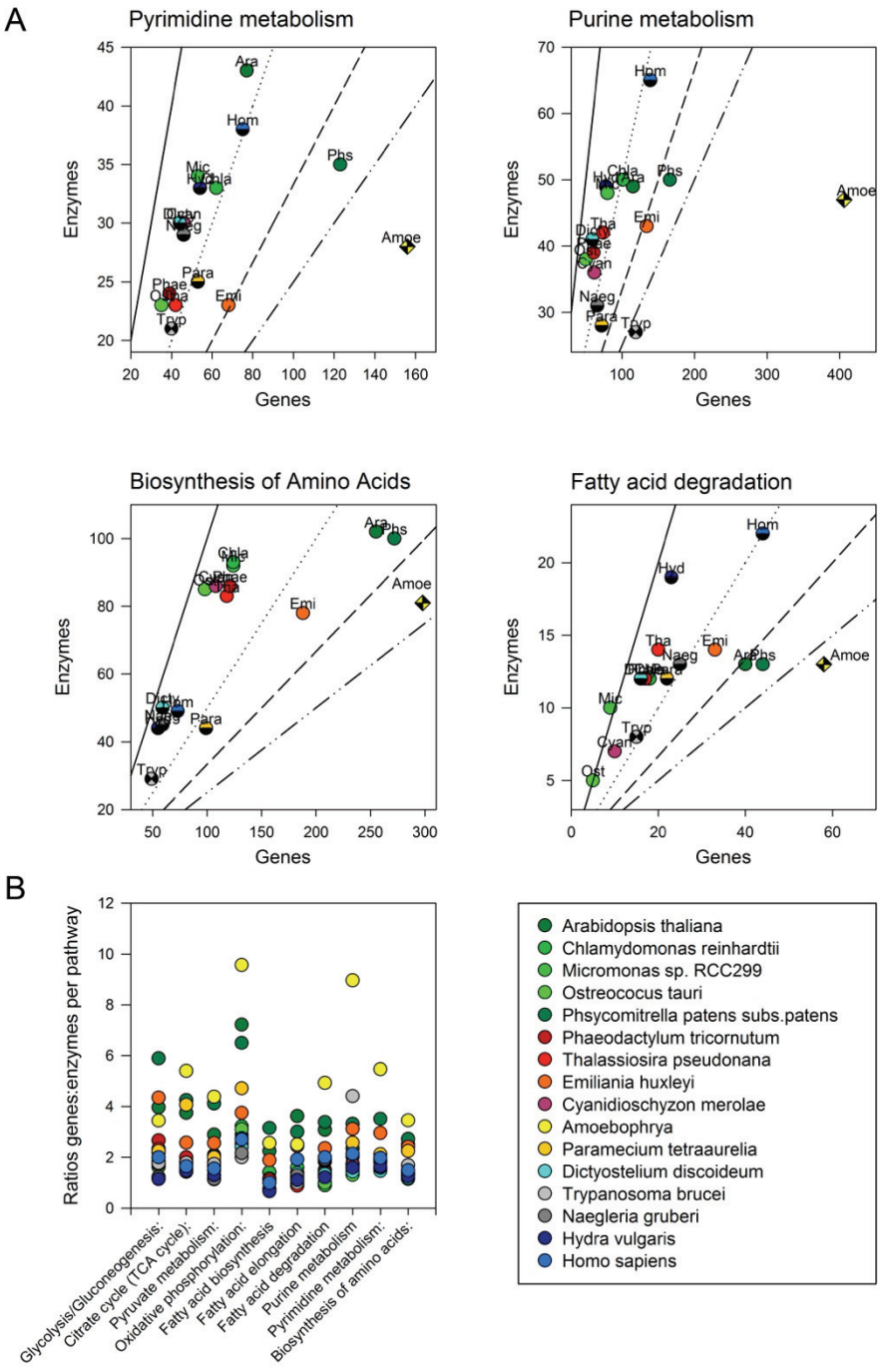


Fig. S4.2.3: Expansion of gene numbers per metabolic pathway in *Amoebophrya*. Numbers were obtained from the Kyoto Encyclopedia of Genes and Genomes (KEGG database; http://www.genome.jp/kegg/pathway.html) and compared with the EC annotations of *Amoebophrya* (combined from the Trinotate annotation (http://trinotate.github.io/) and KAAS annotation pipeline (Moriya *et al.* 2007). (A) Numbers of genes per pathway enzymes in different organisms. For each species the number of genes in a certain pathway were plotted against the number of enzymatic activities in this pathway. (B) Ratios of genes: enzymes of key metabolic pathways of different species

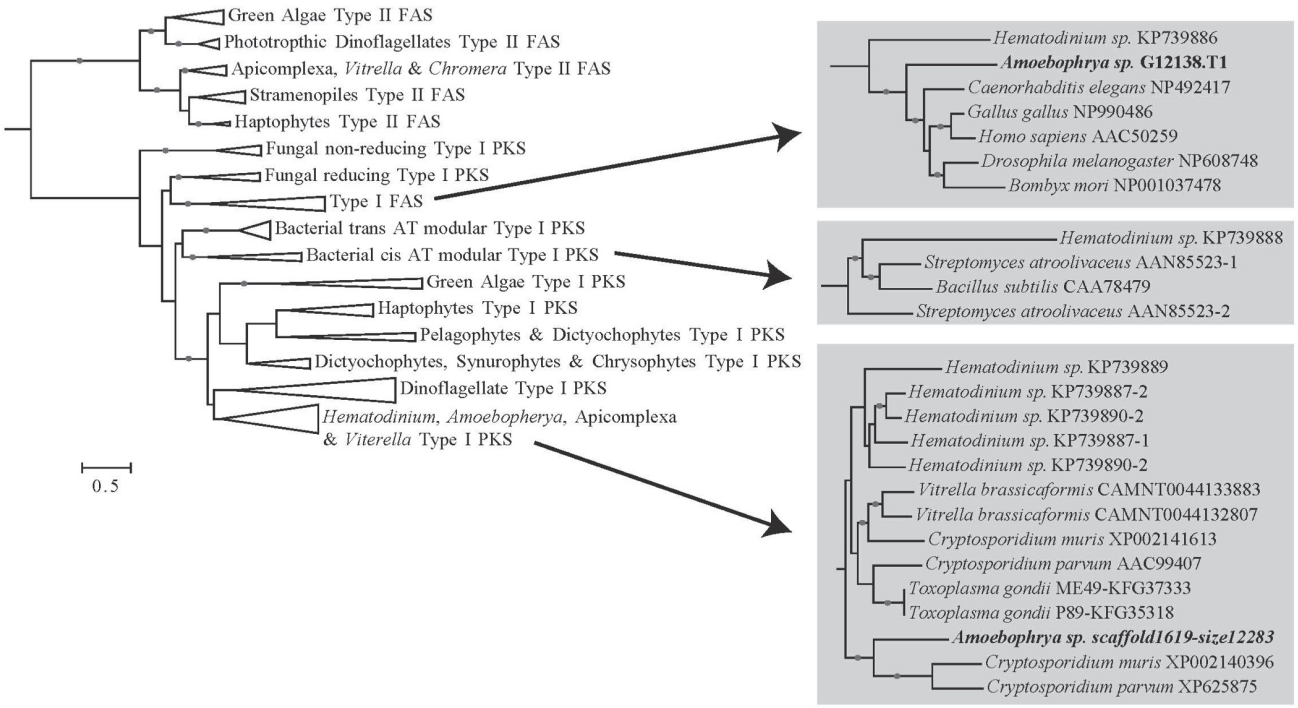
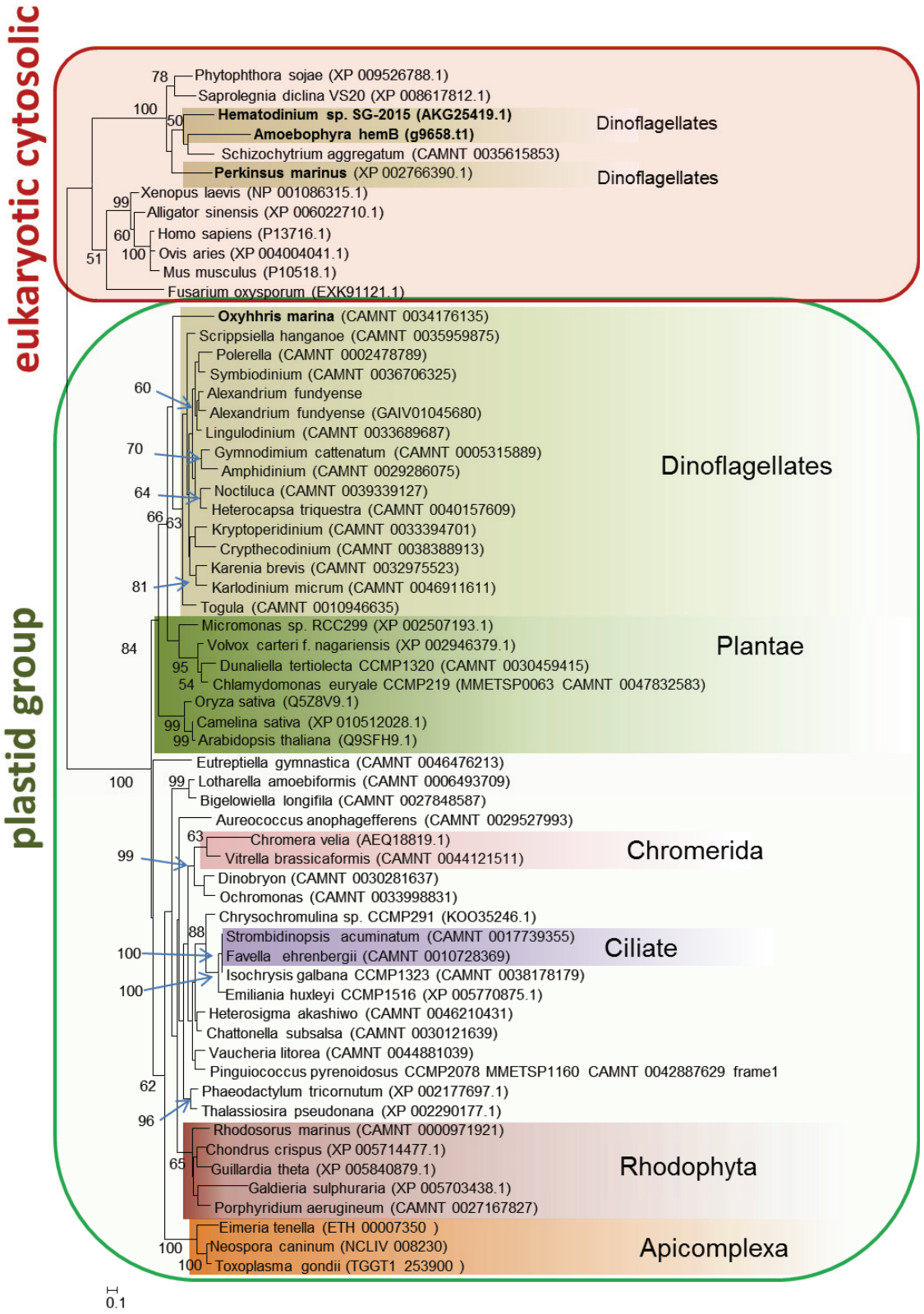

Green Algae Type II FAS
Phototropthic Dinoflagellates Type II FAS
Apicomplexa, Vitrella & Chromera Type II FAS
Stramenopiles Type II FAS
Haptophytes Type II FAS
Fungal non-reducing Type I PKS
Fungal reducing Type I PKS
Type I FAS
Bacterial trans AT modular Type I PKS
Bacterial cis AT modular Type I PKS
Green Algae Type I PKS
Haptophytes Type I PKS
Pelagophytes & Dictyochophytes Type I PKS
Dictyochophytes, Synurophytes & Chrysophytes Type I PKS
Dinoflagellate Type I PKS
Hematodinium, Amoebopherya, Apicomplexa & Viterella Type I PKS
0.5
Hematodinium sp. KP739886
Amoebophrya sp. G12138.T1
Caenorhabditis elegans NP492417
Gallus gallus NP990486
Homo sapiens AAC50259
Drosophila melanogaster NP608748
Bombyx mori NP001037478
Hematodinium sp. KP739888
Streptomyces atroolivaceus AAN85523-1
Bacillus subtilis CAA78479
Streptomyces atroolivaceus AAN85523-2
Hematodinium sp. KP739889
Hematodinium sp. KP739887-2
Hematodinium sp. KP739890-2
Hematodinium sp. KP739887-1
Hematodinium sp. KP739890-2
Vitrella brassicaformis CAMNT0044133883
Vitrella brassicaformis CAMNT0044132807
Cryptosporidium muris XP002141613
Cryptosporidium parvum AAC99407
Toxoplasma gondii ME49-KFG37333
Toxoplasma gondii P89-KFG35318
Amoebophrya sp. scaffold1619-size12283
Cryptosporidium muris XP002140396
Cryptosporidium parvum XP625875

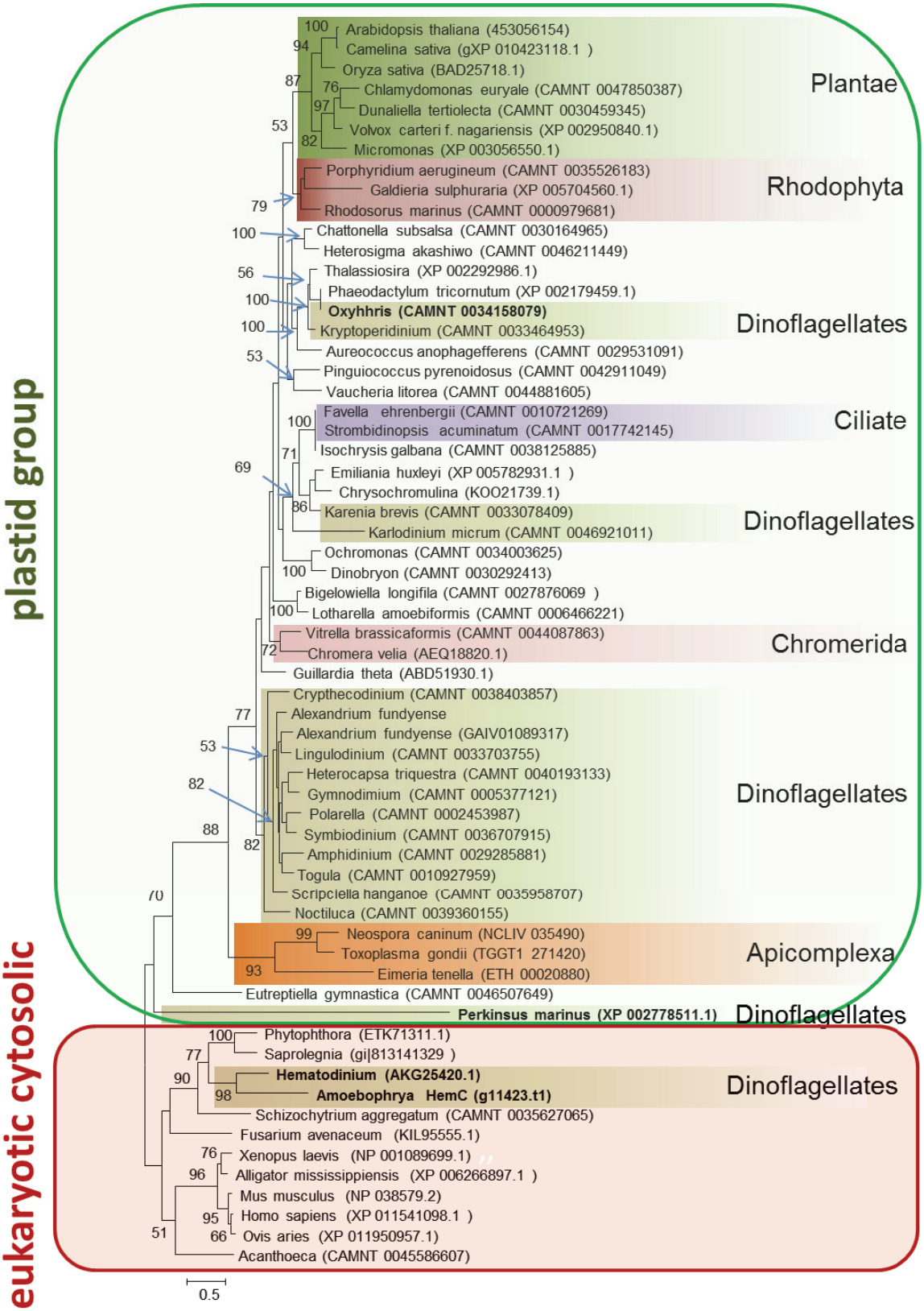
Fig. S4.2.4: Phylogenetic analysis of 26 type II 3-ketoacyl ACP synthase II and 74 type I ketosynthase domains from prokaryotic and eukaryotic polyketide synthase and fatty acid synthases, showing the position of each major group, inferred in RAxML using GAMMA model of rate heterogeneity and 1000 bootstraps (584 characters). Solid circles indicate bootstrap values above 75. To align sequences, MAFFT (Katoh *et al.* 2002) and ClustalW (Thompson *et al.* 1994) were used and alignments were trimmed manually to ensure they spanned the same coding region of each enzyme. Maximum likelihood phylogenetic analysis was carried out using RAxML with 1000 bootstraps using the GAMMA and LG model of rate heterogeneity (Stamatakis 2006). Phylogenetic trees were visualised using Geneious (Kearse *et al.* 2012) and MEGA:Version6 (Tamura *et al.* 2013).

The genome of *Amoebophrya* was searched for genes encoding Type I & II fatty acid synthases (FAS) and type I polyketide synthases (PKS). Searches were conducted using HMMER (Finn *et al.* 2011) and procedures used have been described in detail in (Kohli *et al.* 2016). No type II FAS domains were found in the genome of *Amoebophrya*. However, two type I ketosynthase (KS) domains were encoded in the genome. Contig G12138.T1 encoded the KS, acyltransferase (AT), dehydratase (DH), enoylreductase (ER) and ketoreductase (KR) domains. Another contig scaffold.619.12283 encoded thioesterase (TE), acyl carrier protein (ACP), AT, KS and ACP domains. Phylogenetic analysis was performed on the KS domains to predict the origin and function of these domains. KS domain encoded in the contig G12138.T1 clustered within the type I FAS clade. Together with the presence of other domains in the sequence of KS-AT-DH-ER-KR and the position of this KS domain in the phylogenetic tree indicate that this cluster of genes might be involved in the synthesis of fatty acids in *Amoebophrya*. Another, KS domain from contig scaffold.619.12283 clustered with type I PKS domains from other protists at it might be involved in the production of natural products.

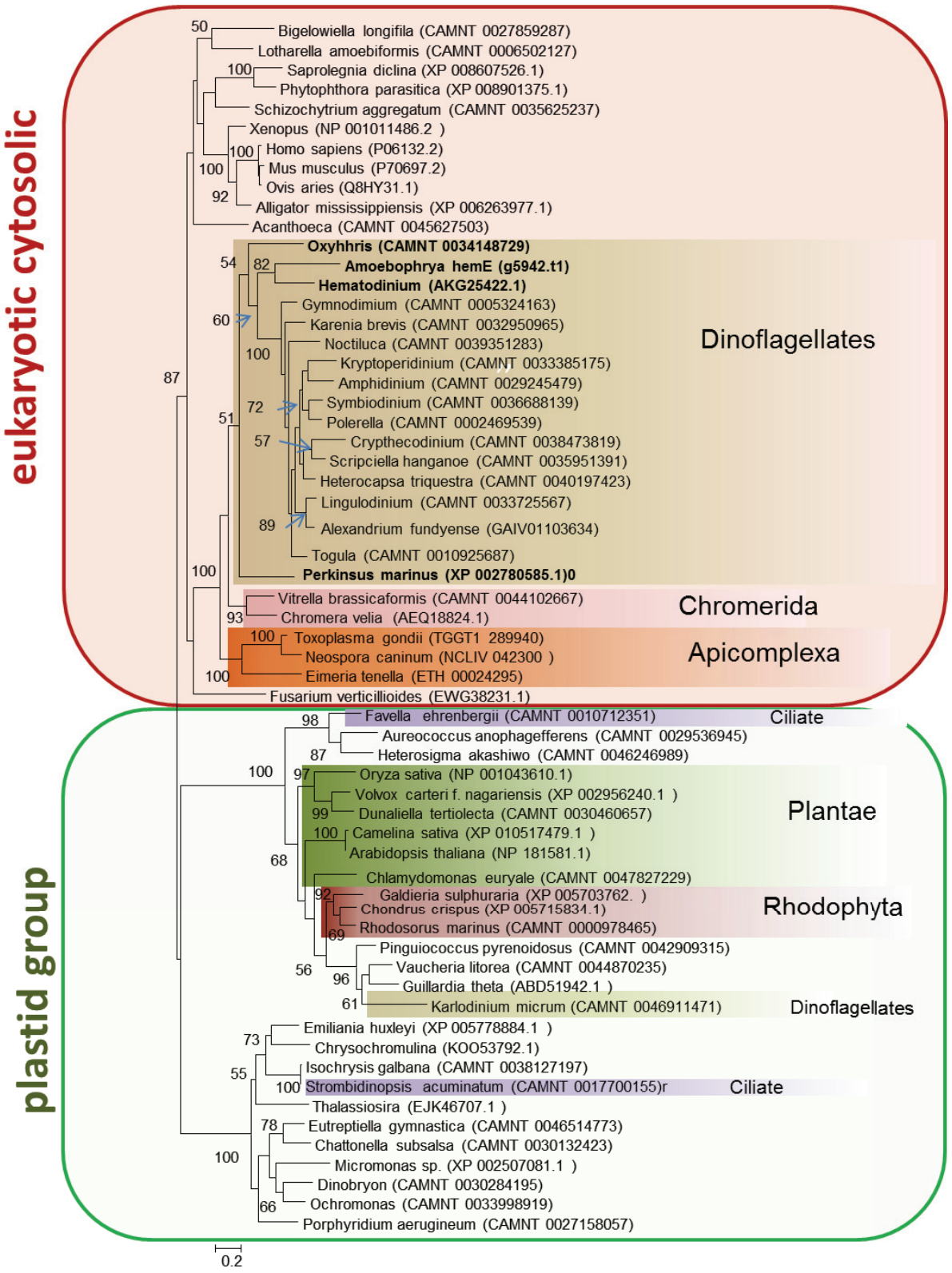
A) HemB


eukaryotic cytosolic
plastid group
Phytophthora sojae (XP 009526788.1)
Saprolegnia diclina VS20 (XP 008617812.1)
Hematodinium sp. SG-2015 (AKG25419.1)
Amoebophyra hemB (g9658.t1)
Dinoflagellates
Schizochytrium aggregatum (CAMNT 0035615853)
Perkinsus marinus (XP 002766390.1)
Dinoflagellates
Xenopus laevis (NP 001086315.1)
Alligator sinensis (XP 006022710.1)
Homo sapiens (P13716.1)
Ovis aries (XP 004004041.1)
Mus musculus (P10518.1)
Fusarium oxysporum (EXK91121.1)
Oxyhhris marina (CAMNT 0034176135)
Scrippsiella hanganoe (CAMNT 0035959875)
Polerella (CAMNT 0002478789)
Symbiodinium (CAMNT 0036706325)
Alexandrium fundyense
Alexandrium fundyense (GAIV01045680)
Lingulodinium (CAMNT 0033689687)
Gymnodimium cattenatum (CAMNT 0005315889)
Amphidinium (CAMNT 0029286075)
Dinoflagellates
Noctiluca (CAMNT 0039339127)
Heterocapsa triquestra (CAMNT 0040157609)
Kryptoperidinium (CAMNT 0033394701)
Crypthecodinium (CAMNT 0038388913)
Karenia brevis (CAMNT 0032975523)
Karlodinium micrum (CAMNT 0046911611)
Togula (CAMNT 0010946635)
Micromonas sp. RCC299 (XP 002507193.1)
Volvox carteri f. nagariensis (XP 002946379.1)
Plantae
Dunaliella tertiolecta CCMP1320 (CAMNT 0030459415)
Chlamydomonas euryale CCMP219 (MMETSP0063 CAMNT 0047832583)
Oryza sativa (Q5Z8V9.1)
Camelina sativa (XP 010512028.1)
Arabidopsis thaliana (Q9SFH9.1)
Eutreptiella gymnastica (CAMNT 0046476213)
Lotharella amoebiformis (CAMNT 0006493709)
Bigelowiella longifila (CAMNT 0027848587)
Aureococcus anophagefferens (CAMNT 0029527993)
Chromera velia (AEQ18819.1)
Vitrella brassicaformis (CAMNT 0044121511)
Chromerida
Dinobryon (CAMNT 0030281637)
Ochromonas (CAMNT 0033998831)
Chrysochromulina sp. CCMP291 (KOO35246.1)
Strombidinopsis acuminatum (CAMNT 0017739355)
Favella ehrenbergii (CAMNT 0010728369)
Ciliate
Isochrysis galbana CCMP1323 (CAMNT 0038178179)
Emiliania huxleyi CCMP1516 (XP 005770875.1)
Heterosigma akashiwo (CAMNT 0046210431)
Chattonella subsalsa (CAMNT 0030121639)
Vaucheria litorea (CAMNT 0044881039)
Pinguiococcus pyrenoidosus CCMP2078 MMETSP1160 CAMNT 0042887629 frame1
Phaeodactylum tricornutum (XP 002177697.1)
Thalassiosira pseudonana (XP 002290177.1)
Rhodosorus marinus (CAMNT 0000971921)
Chondrus crispus (XP 005714477.1)
Guillardia theta (XP 005840879.1)
Rhodophyta
Galdieria sulphuraria (XP 005703438.1)
Porphyridium aerugineum (CAMNT 0027167827)
Eimeria tenella (ETH 00007350)
Neospora caninum (NCLIV 008230)
Apicomplexa
Toxoplasma gondii (TGGT1 253900)
0.1

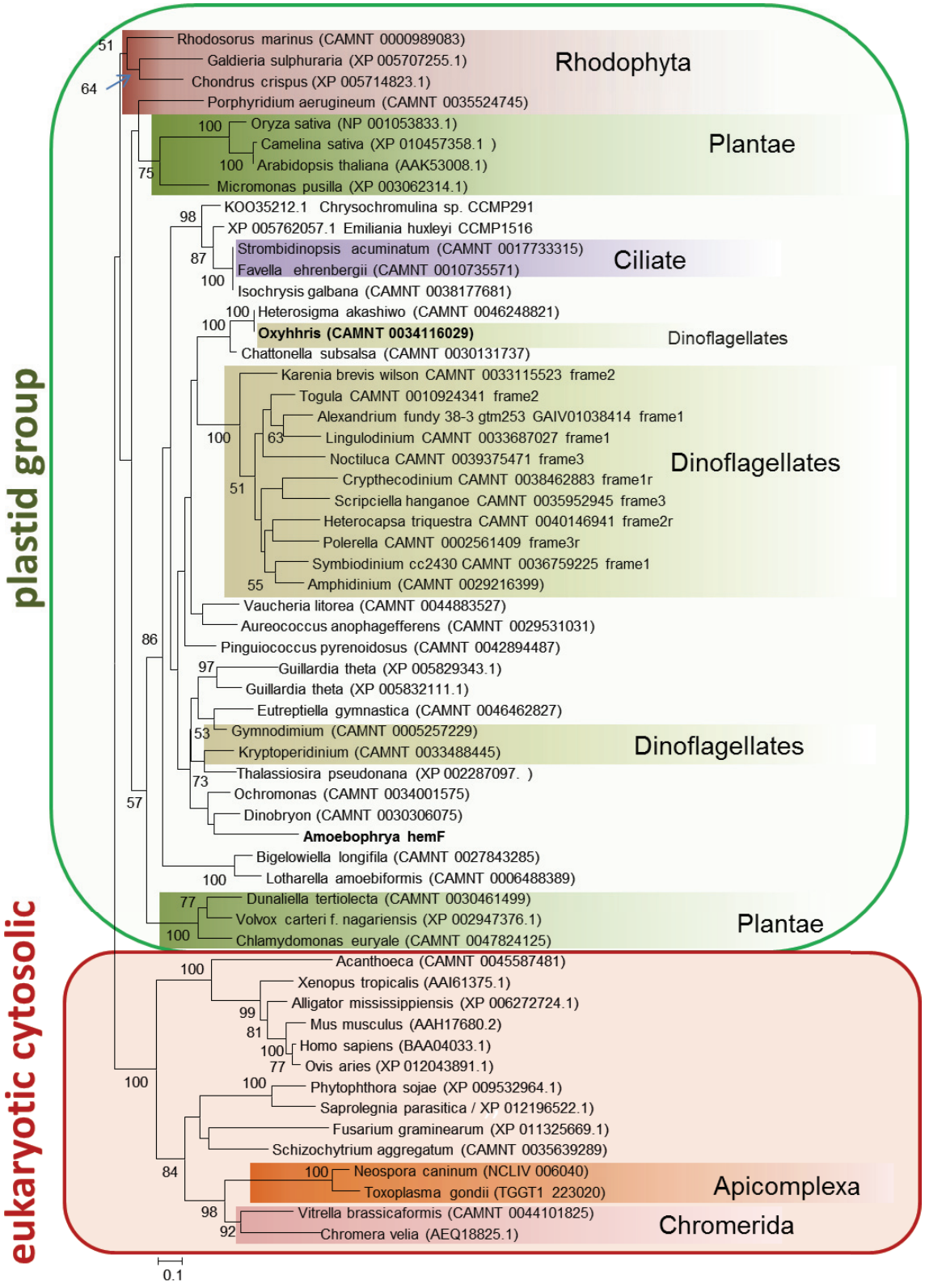
B) HemC


plastid group
eukaryotic cytosolic
Arabidopsis thaliana (453056154)
Camelina sativa (gXP 010423118.1)
Oryza sativa (BAD25718.1)
Chlamydomonas euryale (CAMNT 0047850387)
Dunaliella tertiolecta (CAMNT 0030459345)
Volvox carteri f. nagariensis (XP 002950840.1)
Micromonas (XP 003056550.1)
Plantae
Porphyridium aerugineum (CAMNT 0035526183)
Galdieria sulphuraria (XP 005704560.1)
Rhodosorus marinus (CAMNT 0000979681)
Rhodophyta
Chattonella subsalsa (CAMNT 0030164965)
Heterosigma akashiwo (CAMNT 0046211449)
Thalassiosira (XP 002292986.1)
Phaeodactylum tricornutum (XP 002179459.1)
Oxyhhris (CAMNT 0034158079)
Kryptoperidinium (CAMNT 0033464953)
Dinoflagellates
Aureococcus anophagefferens (CAMNT 0029531091)
Pinguiococcus pyrenoidosus (CAMNT 0042911049)
Vaucheria litorea (CAMNT 0044881605)
Favella ehrenbergii (CAMNT 0010721269)
Strombidinopsis acuminatum (CAMNT 0017742145)
Ciliate
Isochrysis galbana (CAMNT 0038125885)
Emiliania huxleyi (XP 005782931.1)
Chrysochromulina (KOO21739.1)
Karenia brevis (CAMNT 0033078409)
Karlodinium micrum (CAMNT 0046921011)
Dinoflagellates
Ochromonas (CAMNT 0034003625)
Dinobryon (CAMNT 0030292413)
Bigelowiella longifila (CAMNT 0027876069)
Lotharella amoebiformis (CAMNT 0006466221)
Vitrella brassicaformis (CAMNT 0044087863)
Chromera velia (AEQ18820.1)
Chromerida
Guillardia theta (ABD51930.1)
Crypthecodinium (CAMNT 0038403857)
Alexandrium fundyense
Alexandrium fundyense (GAIV01089317)
Lingulodinium (CAMNT 0033703755)
Heterocapsa triquestra (CAMNT 0040193133)
Gymnodimium (CAMNT 0005377121)
Polarella (CAMNT 0002453987)
Symbiodinium (CAMNT 0036707915)
Amphidinium (CAMNT 0029285881)
Togula (CAMNT 0010927959)
Scripciella hanganoe (CAMNT 0035958707)
Noctiluca (CAMNT 0039360155)
Dinoflagellates
Neospora caninum (NCLIV 035490)
Toxoplasma gondii (TGGT1 271420)
Eimeria tenella (ETH 00020880)
Apicomplexa
Eutreptiella gymnastica (CAMNT 0046507649)
Perkinsus marinus (XP 002778511.1)
Dinoflagellates
Phytophthora (ETK71311.1)
Saprolegnia (gi|813141329)
Hematodinium (AKG25420.1)
Amoebophrya HemC (g11423.t1)
Dinoflagellates
Schizochytrium aggregatum (CAMNT 0035627065)
Fusarium avenaceum (KIL95555.1)
Xenopus laevis (NP 001089699.1)
Alligator mississippiensis (XP 006266897.1)
Mus musculus (NP 038579.2)
Homo sapiens (XP 011541098.1)
Ovis aries (XP 011950957.1)
Acanthoeca (CAMNT 0045586607)
0.5

C) HemE


eukaryotic cytosolic
plastid group
Bigelowiella longifila (CAMNT 0027859287)
Lotharella amoebiformis (CAMNT 0006502127)
Saprolegnia diclina (XP 008607526.1)
Phytophthora parasitica (XP 008901375.1)
Schizochytrium aggregatum (CAMNT 0035625237)
Xenopus (NP 001011486.2)
Homo sapiens (P06132.2)
Mus musculus (P70697.2)
Ovis aries (Q8HY31.1)
Alligator mississippiensis (XP 006263977.1)
Acanthoeca (CAMNT 0045627503)
Oxyhhris (CAMNT 0034148729)
Amoebophrya hemE (g5942.t1)
Hematodinium (AKG25422.1)
Gymnodimium (CAMNT 0005324163)
Karenia brevis (CAMNT 0032950965)
Noctiluca (CAMNT 0039351283)
Kryptoperidinium (CAMNT 0033385175)
Amphidinium (CAMNT 0029245479)
Symbiodinium (CAMNT 0036688139)
Polerella (CAMNT 0002469539)
Crypthecodinium (CAMNT 0038473819)
Scripciella hanganoe (CAMNT 0035951391)
Heterocapsa triquestra (CAMNT 0040197423)
Lingulodinium (CAMNT 0033725567)
Alexandrium fundyense (GAIV01103634)
Togula (CAMNT 0010925687)
Perkinsus marinus (XP 002780585.1)0
Dinoflagellates
Vitrella brassicaformis (CAMNT 0044102667)
Chromera velia (AEQ18824.1)
Chromerida
Toxoplasma gondii (TGGT1 289940)
Neospora caninum (NCLIV 042300)
Eimeria tenella (ETH 00024295)
Apicomplexa
Fusarium verticillioides (EWG38231.1)
Favella ehrenbergii (CAMNT 0010712351)
Ciliate
Aureococcus anophagefferens (CAMNT 0029536945)
Heterosigma akashiwo (CAMNT 0046246989)
Oryza sativa (NP 001043610.1)
Volvox carteri f. nagariensis (XP 002956240.1)
Dunaliella tertiolecta (CAMNT 0030460657)
Camelina sativa (XP 010517479.1)
Arabidopsis thaliana (NP 181581.1)
Chlamydomonas euryale (CAMNT 0047827229)
Plantae
Galdieria sulphuraria (XP 005703762.)
Chondrus crispus (XP 005715831.1)
Rhodosorus marinus (CAMNT 0000978465)
Rhodophyta
Pinguiococcus pyrenoidosus (CAMNT 0042909315)
Vaucheria litorea (CAMNT 0044870235)
Guillardia theta (ABD51942.1)
Karlodinium micrum (CAMNT 0046911471)
Dinoflagellates
Emiliania huxleyi (XP 005778884.1)
Chrysochromulina (KOO53792.1)
Isochrysis galbana (CAMNT 0038127197)
Strombidinopsis acuminatum (CAMNT 0017700155)r
Ciliate
Thalassiosira (EJK46707.1)
Eutreptiella gymnastica (CAMNT 0046514773)
Chattonella subsalsa (CAMNT 0030132423)
Micromonas sp. (XP 002507081.1)
Dinobryon (CAMNT 0030284195)
Ochromonas (CAMNT 0033998919)
Porphyridium aerugineum (CAMNT 0027158057)
0.2

D) HemF


plastid group
eukaryotic cytosolic
Rhodophyta
Rhodosorus marinus (CAMNT 0000989083)
Galdieria sulphuraria (XP 005707255.1)
Chondrus crispus (XP 005714823.1)
Porphyridium aerugineum (CAMNT 0035524745)
Plantae
Oryza sativa (NP 001053833.1)
Camelina sativa (XP 010457358.1)
Arabidopsis thaliana (AAK53008.1)
Micromonas pusilla (XP 003062314.1)
KOO35212.1 Chrysochromulina sp. CCMP291
XP 005762057.1 Emiliania huxleyi CCMP1516
Ciliate
Strombidinopsis acuminatum (CAMNT 0017733315)
Favella ehrenbergii (CAMNT 0010735571)
Isochrysis galbana (CAMNT 0038177681)
Heterosigma akashiwo (CAMNT 0046248821)
Dinoflagellates
Oxyhhris (CAMNT 0034116029)
Chattonella subsalsa (CAMNT 0030131737)
Dinoflagellates
Karenia brevis wilson CAMNT 0033115523 frame2
Togula CAMNT 0010924341 frame2
Alexandrium fundy 38-3 gtm253 GAIV01038414 frame1
Lingulodinium CAMNT 0033687027 frame1
Noctiluca CAMNT 0039375471 frame3
Crypthecodinium CAMNT 0038462883 frame1r
Scripciella hanganoe CAMNT 0035952945 frame3
Heterocapsa triquestra CAMNT 0040146941 frame2r
Polerella CAMNT 0002561409 frame3r
Symbiodinium cc2430 CAMNT 0036759225 frame1
Amphidinium (CAMNT 0029216399)
Vaucheria litorea (CAMNT 0044883527)
Aureococcus anophagefferens (CAMNT 0029531031)
Pinguiococcus pyrenoidosus (CAMNT 0042894487)
Guillardia theta (XP 005829343.1)
Guillardia theta (XP 005832111.1)
Eutreptiella gymnastica (CAMNT 0046462827)
Dinoflagellates
Gymnodimium (CAMNT 0005257229)
Kryptoperidinium (CAMNT 0033488445)
Thalassiosira pseudonana (XP 002287097.)
Ochromonas (CAMNT 0034001575)
Dinobryon (CAMNT 0030306075)
Amoebophrya hemF
Bigelowiella longifila (CAMNT 0027843285)
Lotharella amoebiformis (CAMNT 0006488389)
Plantae
Dunaliella tertiolecta (CAMNT 0030461499)
Volvox carteri f. nagariensis (XP 002947376.1)
Chlamydomonas euryale (CAMNT 0047824125)
Acanthoeca (CAMNT 0045587481)
Xenopus tropicalis (AAI61375.1)
Alligator mississippiensis (XP 006272724.1)
Mus musculus (AAH17680.2)
Homo sapiens (BAA04033.1)
Ovis aries (XP 012043891.1)
Phytophthora sojae (XP 009532964.1)
Saprolegnia parasitica / XP 012196522.1)
Fusarium graminearum (XP 011325669.1)
Schizochytrium aggregatum (CAMNT 0035639289)
Neospora caninum (NCLIV 006040)
Toxoplasma gondii (TGGT1 223020)
Apicomplexa
Vitrella brassicaformis (CAMNT 0044101825)
Chromera velia (AEQ18825.1)
Chromerida
0.1

E) HemD

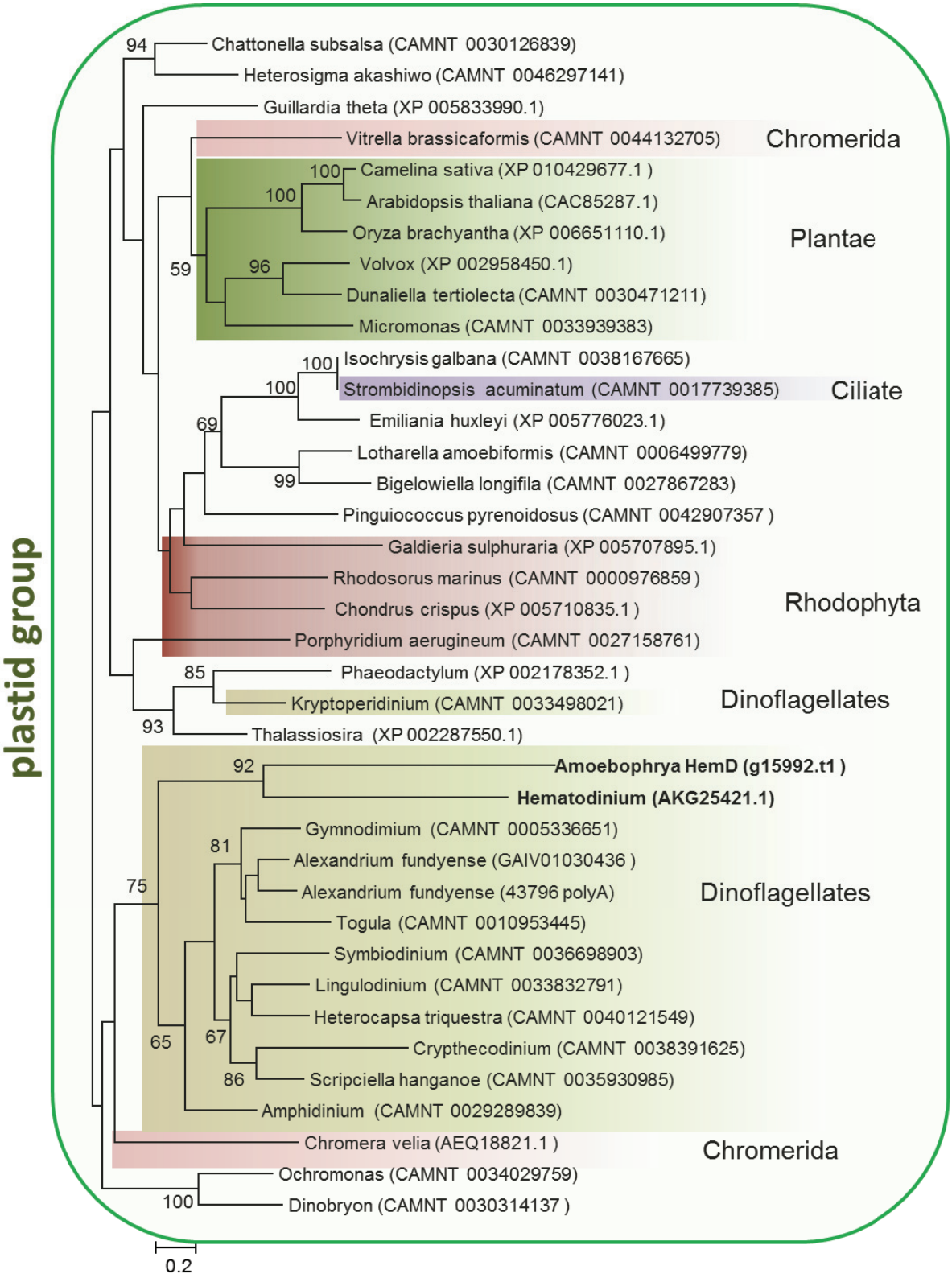

plastid group
Chattonella subsalsa (CAMNT 0030126839)
Heterosigma akashiwo (CAMNT 0046297141)
Guillardia theta (XP 005833990.1)
Vitrella brassicaformis (CAMNT 0044132705)
Chromerida
Camelina sativa (XP 010429677.1)
Arabidopsis thaliana (CAC85287.1)
Oryza brachyantha (XP 006651110.1)
Plantae
Volvox (XP 002958450.1)
Dunaliella tertiolecta (CAMNT 0030471211)
Micromonas (CAMNT 0033939383)
Isochrysis galbana (CAMNT 0038167665)
Strombidinopsis acuminatum (CAMNT 0017739385)
Ciliate
Emiliania huxleyi (XP 005776023.1)
Lotharella amoebiformis (CAMNT 0006499779)
Bigelowiella longifila (CAMNT 0027867283)
Pinguiococcus pyrenoidosus (CAMNT 0042907357)
Galdieria sulphuraria (XP 005707895.1)
Rhodosorus marinus (CAMNT 0000976859)
Chondrus crispus (XP 005710835.1)
Rhodophyta
Porphyridium aerugineum (CAMNT 0027158761)
Phaeodactylum (XP 002178352.1)
Kryptoperidinium (CAMNT 0033498021)
Dinoflagellates
Thalassiosira (XP 002287550.1)
Amoebophrya HemD (g15992.t1)
Hematodinium (AKG25421.1)
Gymnodimium (CAMNT 0005336651)
Alexandrium fundyense (GAIV01030436)
Alexandrium fundyense (43796 polyA)
Dinoflagellates
Togula (CAMNT 0010953445)
Symbiodinium (CAMNT 0036698903)
Lingulodinium (CAMNT 0033832791)
Heterocapsa triquestra (CAMNT 0040121549)
Crypthecodinium (CAMNT 0038391625)
Scripciella hanganoe (CAMNT 0035930985)
Amphidinium (CAMNT 0029289839)
Chromera velia (AEQ18821.1)
Chromerida
Ochromonas (CAMNT 0034029759)
Dinobryon (CAMNT 0030314137)
0.2

Fig. S4.2.5: Phylogenetic analysis of 5 Hem genes (HemB, HemC, HemE, HemD, HemF). **A)** Porphobilinogen synthase (HemB); **B)** porphobilinogen deaminase (HemC); **C)** uroporphyrinogen III decarboxylase (HemE) are encoded by a cytosolic orthologue. **D)** coproporphyrinogen oxidase (HemF; **E)** are plastid) uroporphyrinogen III synthase (HemD) derived. To align sequences, MAFFT (Katoh *et al.* 2002) and ClustalW (Thompson *et al.* 1994) were used and alignments were trimmed manually to ensure they spanned the same coding region of each enzyme. Maximum likelihood phylogenetic analysis was carried out using RAxML Version 8 with 1000 bootstraps using the GAMMA and WAG model of rate heterogeneity (Stamatakis 2006). Phylogenetic trees were visualised using MEGA:Version6 (Tamura *et al.* 2013). The genome of *Amoebophrya* was searched for genes encoding Hem synthesis. Searches were conducted using HMMER (Finn *et al.* 2011) and BlastP using the *Hematodinium* sequences or other derived from GenBank as template.

Table S4.2.1: Splice leaders (SL) sequence detected in the genome of *Amoebophrya.* (A) Predicted genes with 5' adjacent SL sequences. (B) Distribution of full length SL on scaffolds of *Amoebophrya*. There are three scaffolds (*Italic, Underline*) which contain a second SL-sequence. (C) Distribution of truncated SL in scaffolds of *Amoebophrya* genome.

A

Sequence ID	Hit length	Hit start	Hit end	Hit strand	Function
g18355.t1	16.00	4738	4753	Plus	Unknown
g6020.t1	15.00	25	11	Minus	Uncharacterized protein containing DHHC-type Zn finger
g13239.t1	13.00	292	304	Plus	Unknown
g11183.t1	13.00	348	336	Minus	Pentatricopeptide repeat-containing protein At1g62930
g3407.t1	13.00	3292	3304	Plus	Unknown

B

Sequence ID	Hit length	Hit start	Hit end	Hit strand
scaffold1189	21	32704	32724	Minus
scaffold1205	21	43000	43020	Minus
scaffold1270	21	12201	12221	Plus
scaffold1289	21	18168	18188	Plus
scaffold1308	21	40822	40842	Plus
scaffold132	21	41109	41129	Plus
scaffold1320	21	19617	19637	Minus
scaffold1358	21	1354	1374	Plus
scaffold1359	21	11485	11505	Plus

Sequence ID	Hit length	Hit start	Hit end	Hit strand
scaffold1385	21	14549	14569	Plus
scaffold1398	21	23563	23583	Plus
scaffold1458	21	309	329	Minus
scaffold1489	21	25315	25335	Plus
scaffold17	21	6576	6596	Minus
scaffold1737	21	4962	4982	Plus
scaffold180	21	268396	268416	Plus
scaffold2066	21	3630	3650	Plus
scaffold2171	21	7973	7993	Plus
scaffold2275	21	4514	4534	Minus
scaffold2289	21	8172	8192	Minus
scaffold233	21	81864	81884	Plus
scaffold2671	21	36	56	Minus
scaffold283	21	211262	211282	Plus
scaffold286	21	34005	34025	Minus
scaffold286	21	25201	25221	Minus
scaffold297	21	106066	106086	Plus
scaffold3000	21	7663	7683	Plus
scaffold307	21	85828	85848	Plus
scaffold309	21	71658	71678	Plus
scaffold309	21	74700	74720	Plus
scaffold329	21	225792	225812	Minus
scaffold3442	21	973	993	Minus
scaffold37	21	271411	271431	Minus
scaffold381	21	4775	4795	Minus
scaffold385	21	8435	8455	Minus
scaffold3942	21	282	302	Minus
scaffold4061	21	1083	1103	Plus
scaffold48	21	17146	17166	Plus
scaffold48	21	215298	215318	Plus
scaffold544	21	63960	63980	Plus
scaffold549	21	2921	2941	Minus
scaffold574	21	85014	85034	Plus
scaffold600	21	7053	7073	Plus
scaffold61	21	114241	114261	Plus
scaffold656	21	1686	1706	Plus
scaffold68	21	131099	131119	Minus
scaffold811	21	61758	61778	Minus
scaffold820	21	12792	12812	Minus
scaffold912	21	56882	56902	Plus
scaffold94	21	33762	33782	Minus
scaffold956	21	41864	41884	Minus
scaffold958	21	1772	1792	Minus

C

Sequence ID	Sequence length	Number of truncated SL*
scaffold61	406243	56
scaffold544	230577	53
scaffold180	286137	47
scaffold329	318215	42
scaffold48	276268	41
scaffold68	248039	40
scaffold37	324791	37
scaffold17	333242	35
scaffold286	196863	35
scaffold283	211282	28
scaffold1358	40547	26
scaffold309	124822	24
scaffold233	133269	23
scaffold94	168647	22
scaffold811	107907	20
scaffold297	150960	18
scaffold381	76577	18
scaffold132	124252	16
scaffold1189	72863	15
scaffold307	88209	13
scaffold1385	64388	12
scaffold574	85034	12
scaffold820	86658	11
scaffold956	84797	11
scaffold1489	38893	10
scaffold958	45976	9
scaffold1205	59972	8
scaffold385	40098	8
scaffold1289	45660	7
scaffold1308	47888	7
scaffold600	54455	7
scaffold1737	13558	6
scaffold549	37577	6
scaffold1398	42967	5
scaffold912	59273	5
scaffold1458	27774	4
scaffold656	38666	4
scaffold1359	45323	3
scaffold2066	8694	3
scaffold2171	8014	3

*too many truncated SL therefore no positions were provided

Table S4.2.2: (A) Illumina RNAseq reads of *Amoebophrya* sp. ex *Alexandrium fundyense* at 6 and 96 hours during the infection. RNA was extracted following the manufacturer's protocol (Sigma-Aldrich, Steinheim, Germany). RNA quality and quantity check were performed using a NanoDrop ND-100 spectrometer (PeqLab, Erlangen, Germany) and a RNA Nano Chip Assay by the 2100 Bioanalyzer (Agilent Technologies, Böblingen, Germany). Library preparation was done using Illumina's TruSeq RNA sample prep kit (Illumina, San Diego, CA, USA). Libraries were sequenced with a HiSeq2000/2500 (high-throughput mode) in single-read/50 cycle mode (Bentley *et al.* 2008). After quality checking the reads were aligned to the *Amoebophrya* predicted genes (20,969 genes) by using CLC Genomics Workbench (CLC Bio, Aarhus, Denmark) to obtain the read counts. (B) KOG enrichments of *Amoebophrya* predicted genes with intronless (0 intron), low-intron (1-3 introns) and high-intron (>3 introns) (***Bold Italic***: p-value < 0.01). Significant enrichments of the transcripts were tested by calculating the *P* value from a hypergeometric distribution at the background level of the annotation of eukaryotic orthologous groups (KOGs) (Subramanian *et al.* 2005). KOGs were considered significantly enriched for a given intron number when test statistics gave a *P* value of < 0.01.

A

Time point	Name	Sequencing approach	No.of reads	Read length
6 hours	Sample_6h_A	single	34.132.086	50
	Sample_6h_B	single	27.920.490	50
	Sample_6h_C	single	29.078.452	50
96 hours	Sample_96h_A	single	33.502.911	50
	Sample_96h_B	single	36.356.519	50
	Sample_96h_C	single	37.047.919	50

B

intron counts	0			0-3			>3		
	Count	*p*-value	%	Count	*p*-value	%	Count	*p*-value	%
Cell wall/membrane/envelope biogenesis	16.33	0.079633653	3.12	29.00	0.093435737	2.66	19.33	0.055213818	2.08
Cell motility	1.50	0.389494754	0.29	2.00	0.315209359	0.18	0.00	0.255190159	0.00
Posttranslational modification, protein turnover, chaperones	53.00	0.064681456	10.13	115.50	0.041555137	10.59	88.00	0.041136207	9.46
Signal transduction mechanisms	***40.83***	***1.93014E-07***	***7.81***	150.00	0.041466774	13.75	***168.33***	***4.60454E-06***	***18.10***
Intracellular trafficking, secretion, and vesicular transport	2.50	0.293493683	0.48	***9.00***	***0.008696886***	***0.82***	0.00	0.00661944	0.00
Defense mechanisms	***28.50***	***4.25763E-06***	***5.45***	33.00	0.021217713	3.02	***0.00***	***5.77826E-13***	***0.00***
Cytoskeleton	1.00	0.362025869	0.19	4.50	0.220721174	0.41	1.50	0.166697724	0.16
RNA processing and modification	0.00	0.31608484	0.00	1.00	0.228131368	0.09	4.00	0.056509636	0.43
Chromatin structure and dynamics	1.50	0.328614056	0.29	2.50	0.178715282	0.23	4.50	0.202707915	0.48
Translation, ribosomal structure and biogenesis	68.50	0.05350139	13.10	***157.00***	***0.001891763***	***14.39***	***91.50***	***0.000367389***	***9.84***
Transcription	16.00	0.109138109	3.06	31.33	0.058576815	2.87	35.50	0.046144398	3.82
Replication, recombination and repair	44.00	0.040346179	8.41	***89.33***	***0.005604752***	***8.19***	***111.00***	***0.000583883***	***11.94***
Energy production and conversion	31.00	0.0655722	5.93	51.00	0.033679218	4.67	53.33	0.057782765	5.73
Cell cycle control, cell division, chromosome partitioning	***0.83***	***0.009829731***	***0.16***	8.00	0.174292284	0.73	11.83	0.0432854	1.27
Amino acid transport and metabolism	***42.00***	***0.001087038***	***8.03***	61.00	0.063035812	5.59	***33.57***	***0.000568139***	***3.61***
Nucleotide transport and metabolism	18.00	0.03486392	3.44	21.33	0.031284004	1.96	23.83	0.10524688	2.56
Carbohydrate transport and	31.50	0.020896894	6.02	44.33	0.045920928	4.06	39.95	0.06673197	4.30

metabolism intron counts	0 Count	*p*-value	%	0-3 Count	*p*-value	%	>3 Count	*p*-value	%
Lipid transport and metabolism	14.17	0.047791046	2.71	48.20	0.016068817	4.42	30.00	0.06497314	3.23
Inorganic ion transport and metabolism	20.50	0.062218657	3.92	***37.00***	***0.003067602***	***3.39***	***59.75***	***0.000561138***	***6.42***
Secondary metabolites biosynthesis, transport and catabolism	13.17	0.042586633	2.52	21.83	0.089222102	2.00	***8.50***	***0.00536165***	***0.91***
General function prediction only	60.17	0.060949467	11.50	124.15	0.047574857	11.38	110.75	0.049142591	11.91
Function unknown	4.00	0.073973436	0.76	18.00	0.079998993	1.65	13.00	0.13907704	1.40

Table S4.2.3: Enriched gene families. I clustered all predicted proteins using OrthoMCL (http://www.orthomcl.org/orthomcl/). The families were then examined for functions according to their domain structure. Furthermore, I combined the gene families according to the functions of their members if the majority of members of a specific family contained the same indicative domain. Families without any definable function were combined to one group, which comprises the highest number of families and proteins. Among the most prominent functions found in families with at least 10 members in *Amoebophrya* are many transposon derived sequences followed by proteins involved in carbohydrate or protein binding. The prominent occurrence of carbohydrate degradation in this list indicates an important role of these functions in the parasites life.

Function	Number of proteins	Number of families
Calcium binding	10	1
Cation channel	11	1
Nucleotide-diphospho-sugar transferase	12	1
Galactosyltransferase	13	1
ABC transporter	16	1
Peptidase	16	1
Cell surface	21	1
Phosphoesterase	24	1
Cell cycle or growth phase-related regulation	32	1
Kinase	32	3
Protease or protease inhibitor	32	3
Cytoskeletton	36	1
Detoxification and degradation	44	2
Carbohydrate degradation	50	3
Protein-protein or protein-carbohydrate interactions	107	6
Transposon	499	12
No definition	1,424	76

Table S4.2.4: Signalling. I searched the predicted gene complement for signatures of known components of these functions using the iprscan algorithm.

Name	ID	Domain	Number
GPCRs	PF10192	Rhodopsin like GPCR transmembrane domain	2
	PF05462	slime mold cyclic AMP receptor	1
Cyclic nucleotide signaling	PF00211	Adenylyl cyclase class-3/4/guanylyl cyclase	12
	PF00233	3'5'-cyclic nucleotide phosphodiesterase, catalytic domain	10
	PF00027	Cyclic nucleotide-binding domain	82
PIP signaling	PF00454	Phosphatidylinositol 3-/4-kinase, catalytic	8
	PF01504	Phosphatidylinositol-4-phosphate 5-kinase, core	9
	PF00613	Phosphoinositide 3-kinase, accessory (PIK) domain	2
	PF00387	Phospholipase C, phosphatidylinositol-specific, Y domain	0
	PF00388	Phospholipase C, phosphatidylinositol-specific , X domain	0
	PF13180	PDZ domain	2
Calcium signaling	PF00168	C2 calcium-dependent membrane targeting	44
	PF00036	EF-hand	10
	PF00122	ATPase, P-type, ATPase-associated domain	27
	PF00612	IQ motif, EF-hand binding site	64
	PF01699	Sodium/calcium exchanger membrane region	25
Heterotrimeric G proteins	PF00503	Guanine nucleotide binding protein (G-protein), alpha subunit	0
	PF00615	Regulator of G protein signalling	1
small G proteins	PF00071	Ras	36
	PF00616	Ras GTPase-activating protein	0
	PF00617	Guanine-nucleotide dissociation stimulator CDC25	0
	PF00618	Ras-like guanine nucleotide exchange factor, N-terminal	0
	PF00025	ARF/SAR superfamily	13
	PF01412	Arf GTPase activating protein	9
	PF01369	SEC7-like	6
	PF00621	RhoGEF domain	0
	PF00620	Rho GTPase-activating protein domain	0
	PF02263	Guanylate-binding protein, N-terminal	5
	PF01926	GTP-binding domain, HSR1-related	21

Name	ID	Domain	Number
Phosphate signaling	PF00069	Serine/threonine-protein kinase-like domain	333
	PF07714	Serine-threonine/tyrosine-protein kinase	7
	PF00149	Metallo-dependent phosphatase	52
	PF03372	Endonuclease/exonuclease/phosphatase	40
	PF00782	Dual specificity phosphatase, catalytic domain	25
	PF00328	Histidine phosphatase superfamily, clade-2	7
	PF00244	14-3-3 domain	5
	PF00498	Forkhead-associated (FHA) domain	17
Histidine kinase	PF00072	Signal transduction response regulator, receiver domain	4
	PF00512	Signal transduction histidine kinase, subgroup 1, dimerisation/phosphoacceptor domain	2
Sensors	PF00989	PAS fold	0
	PF08376	Nitrate/nitrite sensing protein	0
	PF04940	BLUF	0

4.3 Genomic insights into processes driving the infection of *Alexandrium fundyense* by the parasite *Amoebophrya* sp.

4.3.1 Abstract

The regulatory circuits during infection of dinoflagellates by their parasites are largely unknown on the molecular level. Here I provide molecular insights into these infection dynamics. *Alexandrium fundyense* is one of the most prominent harmful algal bloom dinoflagellate. Its pathogen, the dinoflagellate parasite *Amoebophrya* sp., has been observed to infect and control the blooms of this species. I generated a dataset of transcripts from three time points (0, 6 and 96 hours) during the infection of this parasite-host system. Assembly of all transcript data from the parasite (>900.000 reads/313MBp with 454/Roche next-generation sequencing [NGS]) yielded 14,455 contigs, to which I mapped the raw transcript reads of each time point of the infection cycle. I show that particular surface lectins are expressed at the beginning of the infection cycle which likely mediate the attachment to the host cell. In a later phase, signal transduction-related genes together with transmembrane transport and cytoskeleton proteins point to a high integration of processes involved in host recognition, adhesion, and invasion. At the final maturation stage, cell division- and proliferation-related genes were highly expressed, reflecting the fast cell growth and nuclear division of the parasite. The molecular insights in dinoflagellate parasite interaction point to general mechanisms also known from other eukaryotic parasites, especially from the Alveolata. These similarities indicate the presence of fundamental processes of parasite infection that have remained stable throughout evolution within different phyla.

*Formerly described as *A. tamarense* in the publication.

4.3.2 Introduction

Dinoflagellates, ciliates and apicomplexa belong to the Alveolata superphylum (Bachvaroff *et al.* 2014; Baldauf 2008; Keeling *et al.* 2005). The dinoflagellates are among the most important primary producers in the marine ecosystem. Some of these species can form harmful algal blooms (HABs) that can be noxious to animals and aquatic ecosystems and thus profoundly impact marine environments (Anderson et al. 2012b). One of the most prominent microalgae recorded to form such HABs is the toxigenic dinoflagellate *Alexandrium* sp., associated with paralytic shellfish poisoning (PSP) (Anderson *et al.* 2012b; Anderson *et al.* 1994). *Alexandrium* life-cycle transitions include sexual and clonal reproduction and both play important roles in the dynamics and recurrence of blooms (Anderson 1998; Wyatt & Jenkinson 1997). Sexual reproduction results in the formation of resting cysts that can remain viable in sediments for years, whereas clonal reproduction is responsible for proliferation that may lead to HABs (Anderson 1998; Anderson *et al.* 2012a; Wyatt & Jenkinson 1997). The formation of HABs and their ability to sustain themselves over time are strongly dependent on favorable abiotic conditions, such as solar radiation, nutrient concentration, salinity, and water mass stability (Shilo 1967; Smayda 2002), but also on biotic parameters, for instance the avoidance of competition and grazing (Tillmann & John 2002). Recent studies demonstrated the active participation of parasitic pathogens in the control of toxic bloom formation and development in both field observations (Chambouvet *et al.* 2008; Mazzillo *et al.* 2011; Salomon & Stolte 2010) and model predictions (Montagnes *et al.* 2008).

The parasitoid *Amoebophrya* sp. has been observed to infect populations of the bloom-forming dinoflagellate *Alexandrium minutum* (Chambouvet *et al.* 2011a; Montagnes *et al.* 2008), with a prevalence up to 40% infected cells observed in the Penzé estuary (Brittany, France) (Chambouvet *et al.* 2008). Specific characteristics make the effects of parasitoid on HABs different from those of predators or parasites: first, a parasitoid infects only one host during its lifetime, whereas a predator kills many prey (Kuris 1974). Second, infection by a parasitoid suppresses further host division and the host inevitably is killed to complete the parasitoid's life cycle, whereas a parasite has effects on the host fitness influencing only indirectly its viability (Kuris 1974; Park *et al.* 2004). The Amoebophryidae belong to Syndiniales (Alveolata) and only one genus, *Amoebophrya*, is known, but this genus has a high genetic diversity (Alves-de-Souza *et al.* 2012). Corresponding environmental sequences

belonging to this group were clustered into the widespread eukaryotic marine alveolate group II (MALV ll) (Guillou *et al.* 2008).

The life cycle of *Amoebophrya* elucidated by Cachon in 1964 (Cachon 1964) was recently confirmed using electron microscopy (Cachon & Cachon 1970; Miller *et al.* 2012). The life cycle begins with small infective biflagellate cells termed dinospores (Cachon 1964; Coats & Park 2002; Park *et al.* 2002b). After finding and recognizing the host and adhering to the host's surface, the dinospores exhibit electron-dense bodies within a microtubular basket. *Amoebophrya* seems to use this structure to enter the host cytoplasm, resembling the rhoptries employed by apicomplexa parasites (Carruthers & Sibley 1997; Dubremetz *et al.* 1998). The dinospores lose their flagella and penetrate first into the host cytoplasm protected by a parasitophorous membrane. In some cases, the parasitoid crosses the nuclear envelope into the host nucleus, losing this protecting membrane (Miller *et al.* 2012). The parasitoid maturation, after being initiated inside the nucleus, takes 2 to 3 days. During this time, the trophont of the parasitoid increases in size, followed by consequent cellular divisions involved during sporogenesis to ultimately form a typical intracellular and multicellular stage called the beehive structure. A motile vermiform stage of brief duration, which is composed of several rows of biflagellate cells with a concerted swimming behavior, is then released from the host after the intracellular maturation of the parasitoid. Soon after this release, this structure dissociates into hundreds of free-living infective dinospores (Coats & Bockstahler 1994; Coats & Park 2002).

Amoebophrya infection has been studied on morphological and physiological levels, and yet the molecular processes of infection are poorly understood. Gene expression data are scarce. At present only Bachvaroff *et al.* (2009) has performed a survey of a host-parasite system at the gene expression level, publishing 898 expressed sequence tags (ESTs) from the host *Karlodinium veneficum* and the parasitoid *Amoebophrya* sp. However, the transcriptomic changes during the host-parasitoid interaction still remain enigmatic.

The objective of this study was to obtain transcriptomic insights into the life cycle of the parasitoid *Amoebophrya* sp. During infection of the toxic dinoflagellate host *Alexandrium* I profiled the transcriptome of *Amoebophrya* sp. at three different life-stages: pure dinospores (0 h), initial infection/penetration stage (6 h) and maturation stage (96 h). By analyzing the

ESTs obtained from different life stages, I could identify processes and genes that may be relevant to these three different life stages. These data foster our understanding of complex host-parasitoid interactions and deliver a mechanistic understanding of the genetic basis enabling *Amoebophrya* sp. to dominate over toxic *Alexandrium* blooms.

4.3.3 Materials and Methods

Cultures

The *Alexandrium fundyense* (formerly described as *A. tamarense* (John *et al.* 2014) strain (Alex5) used in this study was isolated from the North Sea coast of Scotland (Tillmann *et al.* 2009). The strain was grown in K-medium (Keller *et al.* 1987), prepared from 0.2 µm sterile-filtered (VacuCap, Pall Life Sciences, Dreieich, Germany) North Sea water.

The *Amoebophrya* sp. strain (AT5.2) used was isolated from host *Alexandrium* cells sampled from the Gulf of Maine, USA (Chambouvet *et al.* 2011b). *Amoebophrya* sp. infecting *Alexandrium* was maintained on the above-described *Alexandrium* strain from the North Sea (3-4 days per generation). All cultures were grown at 15 °C, with cool-white fluorescent lamps providing 150 µmol photons m^{-2} s^{-1} on a light:dark cycle of 14h:10h.

Fixation and counting methods

Host samples (10 ml) were fixed with Lugol's solution (10 g potassium iodide; 5 g iodine in 100 ml distilled water) with a final concentration of 2% (Tillmann *et al.* 2009) and three 1-ml aliquots were counted after sedimentation in chambers under an inverted microscope (Zeiss Axiovert 200M). The total number of cells counted was always >400 per sample. The growth rate (µ) of *Alexandrium* was calculated with the following formula (Guillard 1973; Tillmann *et al.* 2009):

$$\mu = \frac{1}{t_2 - t_1} ln \frac{N_2}{N_1}$$

Where µ is the growth rate (day^{-1}), N_1 and N_2 are the abundances of *Alexandrium* at t_1 and t_2, respectively, and *t* is the sampling day.

The persistence of green autofluorescence indicated the survival of the dinospores. Samples (10 ml) were fixed with formaldehyde (10% $CaCO_3$ buffered formaldehyde; 2% final concentration), and three 1-ml aliquots were counted in duplicate after sedimentation in chambers using a microscope (Zeiss Axiovert 200M) (Coats & Park 2002). Parasitoid prevalence (the percentage of infected host cells) was assessed by detecting well maturing parasitoid using the natural autofluorescence of the parasitoid (Coats & Bockstahler 1994). More than 400 cells from each aliquot were screened using an epifluorescence microscope (Carl Zeiss AG, Göttingen, Germany).

Infection experiments

The infection experiments were set up to cover one complete life cycle of the parasitoid and included three harvesting time points (0 h, 6 h and 96 h; see Fig. S4.3.1 in the supplemental material). Infection of the host culture was established following the methods of Coats and Park (2002). Infective parasitoid dinospores for the experiment were harvested from an infected host culture by gravity filtration through a 10 µm pore size mesh. Harvested cultures of three 1-ml were checked under an inverted microscope (Zeiss Axiovert 200M) to make sure that no host cells remained. One part of this dinospore culture (15 ml) was used for RNA extraction (time point 0h). The remaining dinospore culture was immediately used to inoculate the triplicate exponential phase cultures of *Alexandrium* with a parasitoid:host ratio of 10:1. Triplicate cultures of 400 ml *Alexandrium* with a concentration of approximately 3,000 cells ml^{-1} were initially prepared in three flasks (500 ml Erlenmeyer flask) for the treatment. Three vials for each triplicate culture with hosts only (~400 ml *Alexandrium* at a concentration of 3,000 cells ml^{-1}) served as controls. Two incubation times were chosen after adding the dinospores: 6 and 96 hours. At each time point samples were taken from the same of each triplicate parasitoid-treated culture for (a) fixation and counting, (b) parasitoid prevalence assessment, (c) RNA extraction and sequencing, and (d) PSP toxin analysis.

RNA extraction and sequencing

Samples (100 ml) were taken for RNA extraction. Cells for RNA extraction were filtered through a 10 µm pore size mesh in order to remove the free living dinospores, suspended from the filter using fresh K-medium, and harvested by centrifugation at 4 °C for 10 min. The supernatant was decanted, and the resulting pellet was immediately re-suspended

in 1 mL of 60 °C hot TriReagent (Sigma-Aldrich, Steinheim, Germany) and transferred to a 2 mL cryovial containing acid washed glass beads. Cells were lysed using a Bio101 FastPrep instrument (Thermo Savant Illkirch, France) at maximum speed for 45 s. Afterwards, 200 mL of pure chloroform was added and the sample was vortexed for 15 s. The mixture was incubated for 10 min at room temperature and then centrifuged at 4 °C for 15 min with 13,000 g. The upper aqueous phase was transferred to a new vial, filled with an equal volume of 100% isopropanol, vortexed and incubated for 2 h at -20 °C to precipitate the RNA. The RNA pellet was collected by 20 min centrifugation at 4 °C with 13,000 g. The pellet was washed twice, first with 70% ethyl alcohol (EtOH) followed by 10 min centrifugation at 4 °C with 13,000 g, then with 96% EtOH followed by 5 min centrifugation at maximum speed, air dried and dissolved with 30 µL RNase free water (Qiagen, Hilden, Germany). The RNA-sample was further cleaned with the RNeasy Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol for RNA clean up including on-column DNA-digestion. An RNA quality check was performed using a NanoDrop ND-100 spectrometer (PeqLab, Erlangen, Germany) and a RNA Nano Chip assay by the use of a 2100 Bioanalyzer (Agilent Technologies, Böblingen, Germany).

Subsamples of the triplicate RNAs from each treatment were pooled and were used for cDNA library preparation. The construction of the cDNA library was done by Vertis Biotechnology AG (Freising-Weihenstephan, Germany). In brief, poly(A)$^+$ RNA was prepared from the total RNA and first-strand cDNA synthesis was primed with random hexamers. The 454 sequencing adaptors were ligated to the 5' and 3' ends of the cDNA. The cDNA was amplified with 19 PCR cycles using a proof reading polymerase. The amplified cDNA was normalized by one cycle of denaturation and reannealing. The cDNA was passed over a hydroxylapatite column to separate the reassociated cDNA from the single-stranded cDNA (ss-cDNA). The ss-cDNA obtained was then amplified with 9 PCR-cycles. cDNAs with a size range between 450 to 650 bp were cut out and eluted from an agarose gel and converted to a 454 Roche titanium sequencing library according to the protocols of the manufacturer (Roche). The sequencing run was performed on a 454 GS FLX system (Roche) by the Max Planck-Genome-centre Cologne, Germany (http://mpgc.mpipz.mpg.de/home/).

Sequence reads were quality trimmed and assembled using a CLC Genomics Workbench (CLC Bio, Aarhus, Denmark) with default settings. In order to identify the cDNA

sequences derived from the parasitoid genome. A further test of contamination of the transcript contigs to be analyzed was done using all available host *Alexandrium* ESTs from the NCBI nucleotide database. The read mapper from CLC Genomics Workbench was used to align all single reads from each of the three time points to the reference *Amoebophrya* contigs. The number of hits were extracted as read counts, and were used to classify the occurrence of *Amoebophrya* contigs at each time points.

Functional annotation

Amoebophrya cDNA contigs (larger than 200bp) were annotated by BLAST search (blastx) (Altschul *et al.* 1997) against the nonredundant protein sequences database of the National Center for Biotechnology information (NCBI; http://www.ncbi.nlm.nih.gov) and the Universal Protein Resource (Uniprot; http://www.uniprot.org). Searches were conducted with Blast2GO (BioBam Bioinformatics S.L.,Valencia, Spain) with an e-value cut-off of $1e^{-6}$. For further functional annotation, the *Amoebophrya* genes were translated into amino acid sequences with the Virtual Ribosome (http://www.cbs.dtu.dk/services/VirtualRibosome) and the batch web CD-search tool (http://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi) was used to assign eukaryotic orthologous groups (KOGs) with an e-value cut-off of $1e^{-7}$. The translated sequences were additionally screened for the occurrence of Pfam domains and families (http://pfam.sanger.ac.uk). Homology searches and potential coding sequences searches were performed following the Trinotate annotation suite guidelines (Grabherr *et al.* 2011; Haas *et al.* 2013). Briefly, putative orthologs were predicted from reciprocal best BLAST hits, peptide sequences were predicted with use of the Trinity transdecoder, and protein families, signal peptides and transmembrane domains were determined using Pfam (Punta *et al.* 2012), signalP (Petersen *et al.* 2011) and tmHMM (Krogh *et al.* 2001), respectively. Comparison to currently currated annotation databases were derived from eggNOG/GO (Powell *et al.* 2012), Gene Ontology (Ashburner *et al.* 2000) and KEGG (Kanehisa *et al.* 2012) pathways databases. Significant enrichments of the contigs were tested by calculating *P* value from a hypergeometric distribution at the background level of all KOGs and Pfam families (Subramanian *et al.* 2005). KOGs and Pfam families were considered significantly enriched for a given experimental time point when test statistics replied a *P* value of < 0.05.

4.3.4 Results

Time-course study of infected cultures

To assess the host-parasitoid interaction, host abundances and infection percentages were measured over a 96-h period (Fig. 4.3.1). As a control, I used the noninfected culture of *Alexandrium*. The growth rate of the infected culture ranged from $\mu = 0.32$ at 24 h to $\mu = 0.17$ at 72 h and was significantly lower than that of the control ($\mu = 0.46$ at 24 h, $\mu = 0.38$ at 72 h) (one-way analysis of variance [ANOVA], $F = 19.96$, $P < 0.05$) (Fig. 4.3.1A, see also Table S4.3.1 in the supplemental material).

The percentage of parasitoid infection increased slightly during the first 48 h. Afterwards, a drastic increase appeared after 72 h, when the infection percentage ascended rapidly from 2% to 36%. The infection coverage remained high and increased during the following 24 h, finally reaching 39% after 96 h (Fig. 4.3.1B).

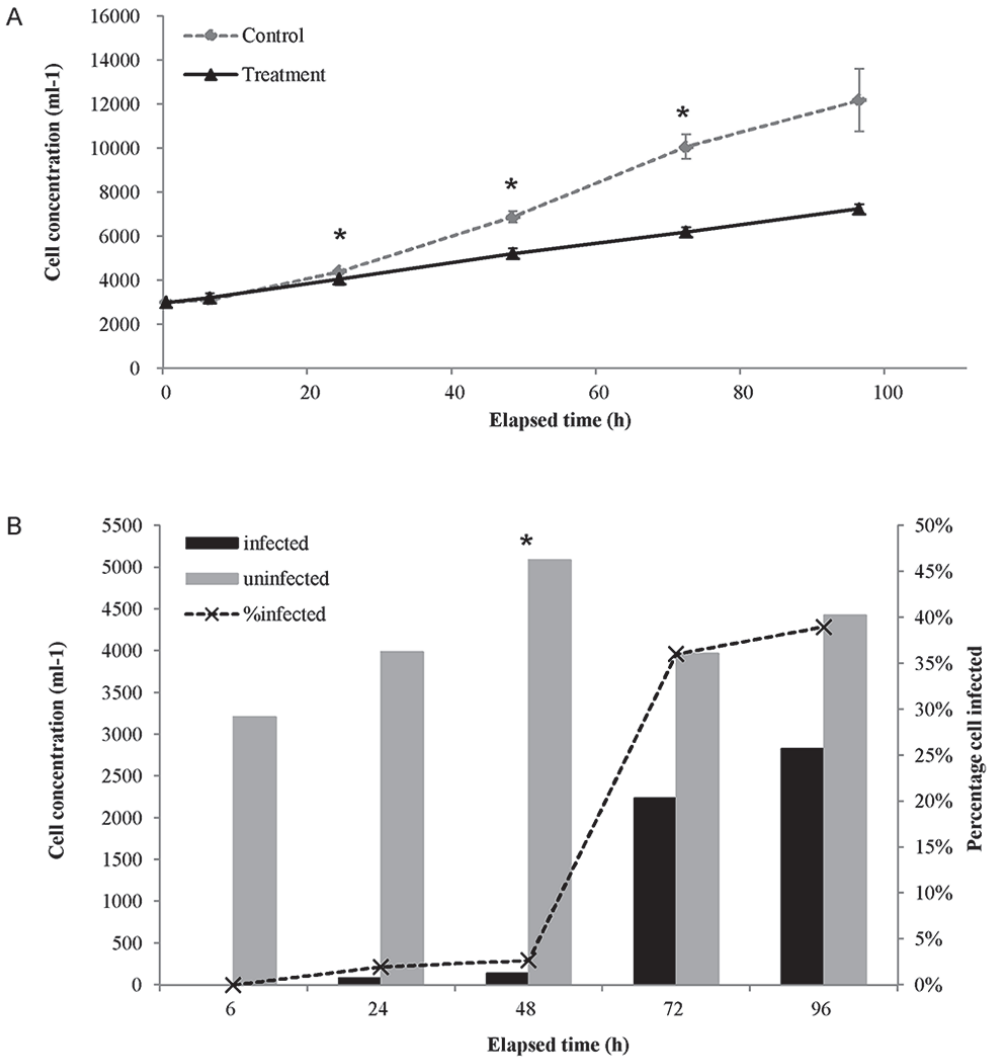


Fig. 4.3.1: Time course study of the host dinoflagellate *Alexandrium* infected with *Amoebophrya*. (A) Total *Alexandrium* cell concentrations in infected cultures (in black) and in healthy cultures (in gray). (B) Estimation of numbers of infected (black) and non-infected (gray) hosts in infected cultures, with percentages of infected cells (dashed black line). Asterisks mark significant differences in the growth rate μ and percentages of infected cells (ANOVA; $P < 0.05$).

Sequencing and assembly statistics

Sequencing the three different libraries over the infection cycle yielded the following results: (i) parasitoid dinospores (time point 0), 445,296 reads/106 Mbp; (ii) the initial infection/penetration stage mixed with host and parasitoid (6 h), 309,199 reads/106 Mbp; and (iii) the matu- ration stage mixed with host and parasitoid (96 h), 301,377 reads/ 101 Mbp. Assemblies of the reads yielded 17,780 contigs for 0 h and 17,680 and 16,495 contigs for 6

and 96 h, respectively (Table 4.3.1). All 931,259 raw reads were *de novo* assembled and yielded 44,674 contigs. A total of 14,336 contigs could be classified as putative parasitoid transcripts, while the remaining 30,326 contigs were likely nonparasitoid transcripts (comprising putative *Alexandrium* host transcripts) (Table 4.3.1). Of the *Amoebophrya*-specific contigs, 6,789 (47.4%) were larger than 500 bp, 6,972 (48.6%) had a size range of 499 to 250 bp, and 575 (4.0%) had a size range of 249 to 200 bp.

Table 4.3.1: Overview of the sequencing data and assembly

	T0	T6	T96	Total
Raw read data				
No. of useable reads	445,296	309,199	301,377	931,259
Average reads length (bp)	238	342	336	336
Total no. of bases	105,995,167	105,788,165	101,190,027	312,973,359
Assembly				
No. of contigs	17,780	17,680	16,495	44,674
N50[a]	546	485	540	539
Largest contig length (bp)	4,556	5,654	4,296	5,398
Average contig length (bp)	557	499	537	544
Total no. of bases	9,917,251	8,828,154	8,872,796	24,318,082
Contamination value				
No. of parasitoid contigs	12,227	4,291	11,911	14,336
Annotation				
No. of potential coding sequences	5,743	2,182	5,455	6,662

[a] N50, the contig length (bp) was calculated by summing the lengths of the biggest contigs until 50% of the total contig length was reached.

Functional annotation of Amoebophrya-specific EST contigs

In the combined EST set, I annotated all *Amoebophrya* contigs (14,336 contigs) and were able to assign a putative function to 6,662 (46.47%) contigs based on BLAST searches and Trinotate against the NCBI Nr, the Uniprot, Pfam, KOG, Gene Ontology and KEGG databases (Table 4.3.2). Annotated sequences were classified into functional categories according to KOGs (Fig. 4.3.2), whereby 33% of the genes were assigned to metabolism, 33% to cellular processes and signaling, 21% to information storage and processing and 13% to the category "poorly characterized". Within these categories, more than half of the functional annotated sequences fell into four subcategories: "[O] posttranslational modification, protein turnover, chaperones" (16%), "[J] translation, ribosomal structure and biogenesis" (12%), "[C] energy production and conversion" (10%), "[T] signal transduction mechanisms" (7%), "[E] amino acid transport and metabolism" (6%) and "[R] general function prediction only" (11%). A full list of the genes with annotations can be retrieved from the supplementary material (see Data set S4.3.1 in the supplemental material).

Table 4.3.2: Gene content and annotation summaries.

Contig category or database	No. of contigs	e-value or significance	% contigs identified
Contig category			
Long (> 200bp)	14,336		
Spliced leader (SL)	51		
Poly (A) tail	91		
Database			
All annotated	6,662		46.47
NCBI_Nr	2,938	e^{-6}	20.49
Uniprot	2,003	e^{-6}	13.97
Pfam	4,944	e^{-11}	34.49
KOG	1,174	e^{-7}	8.19
Gene Ontology	2,771	e^{-11}	19.33
KEGG enzyme	1,325	e^{-11}	9.24

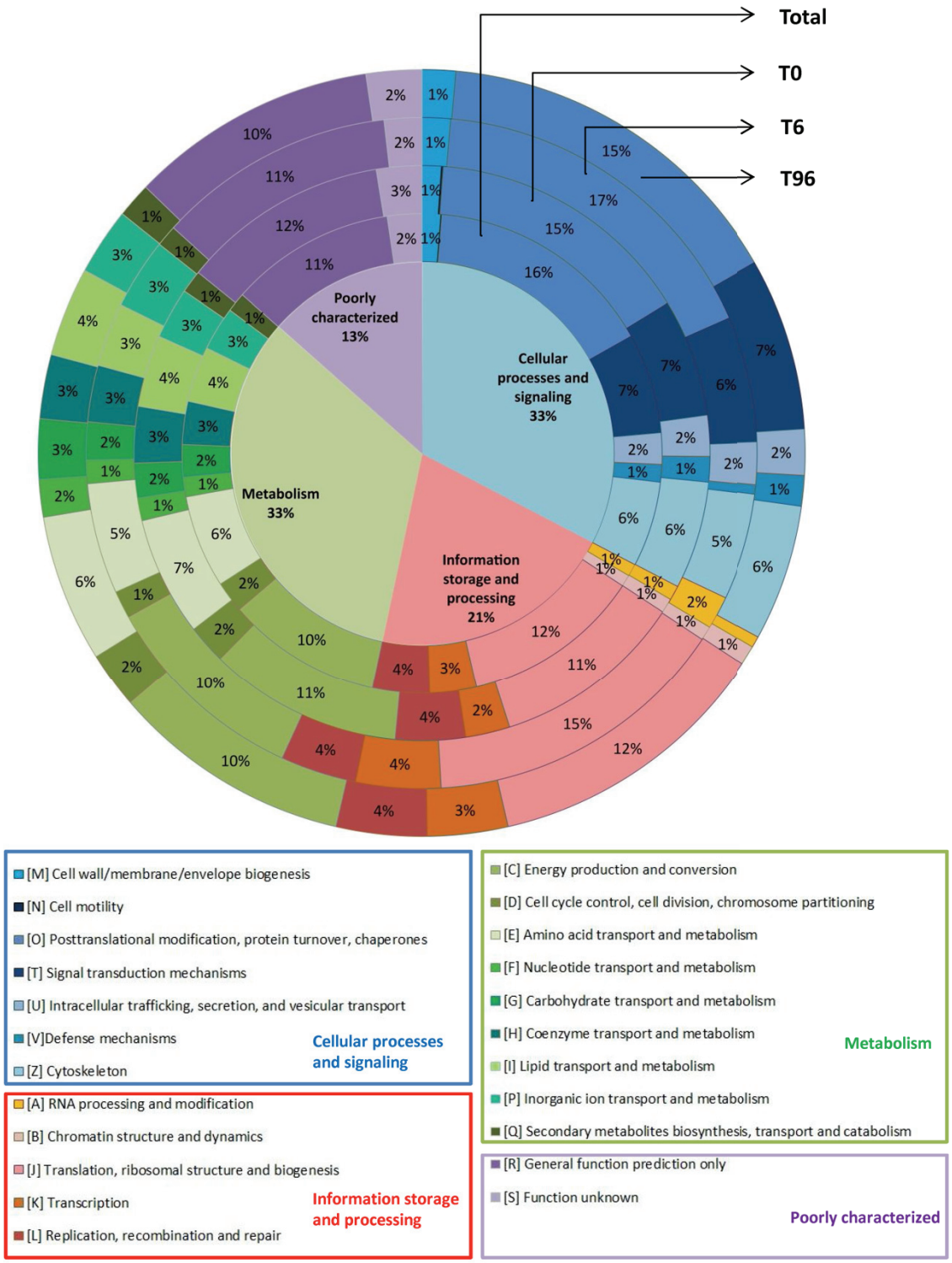


Fig 4.3.2: KOG category distribution of unique sequences of the expressed *Amoebophrya* genes. The color intensities within each group denote the subcategories. Percentages of the total number of genes grouped into these categories are given in each segment.

Comparison of three time points (0 h, 6 h and 96 h) of the infection cycle

All single reads from each of the three time points were mapped back to the *Amoebophrya* contigs (14,336 contigs). The read mapping indicated that 12,227 contigs were present at 0 h (pure dinospores) and that 4,291 and 11,911 contigs were present at 6-h initial penetration stage and 96-h maturation stage, respectively, with 3,587 contigs in common to all time points (Table 4.3.1 and Fig. 4.3.3). To investigate the potential gene products and pathways regulated during infection, I assessed the transcriptional changes between the time points. At the background of all 14,336 contigs, the significant enriched contigs were sorted by calculating the *P* value of < 0.05 from the hypergeometric distribution (Subramanian *et al.* 2005) (see Data set S4.3.1 in the supplemental material). Complete lists of the Pfam domains and families of over-represented genes (*P* value < 0.05) and KOG enrichments (*P* value < 0.05) are provided in Table S4.3.2 and Table S4.3.3, respectively. In the following, I describe our main findings.

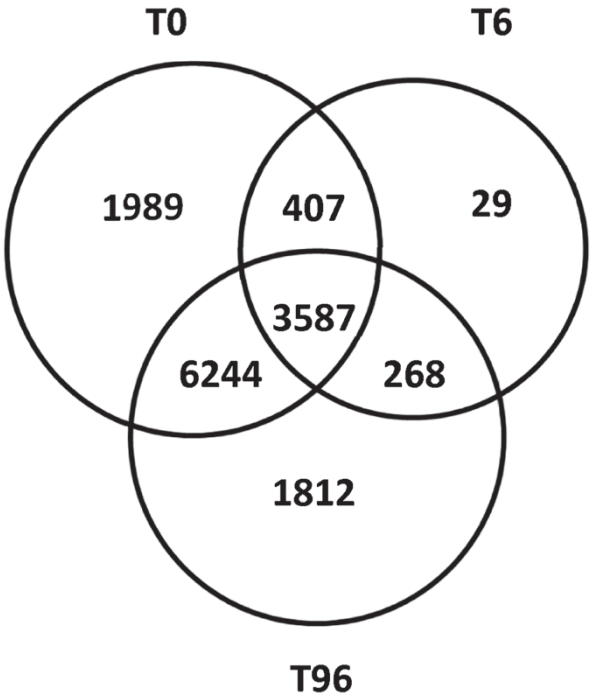


Fig. 4.3.3: Venn diagram of all 14,336 contigs (assembled from all time points) categorized with reads mapped from three time points: T0, pure dinospores; T6, initial penetration stage; T96, maturation stage. The number of raw reads mapping to each contig can be found in Data set S4.3.1 in the supplemental material.

Dinospores (0 h)

Significantly enriched protein families and genes in dinospores were mainly involved in processes associated with energy production (e.g. cytochrome, NADH dehydrogenase, inorganic pyrophosphatase), cell adhesion (e.g. carbohydrate-binding proteins, lectins), amino acid transport (e.g. aminotransferases, cysteine synthase, glycine/serine hydroxymethyltransferase) and ribosomal structure (e.g. ribosomal protein, GTPase). Four types of lectins were present in dinospores: ricin-type beta-trefoil lectin, C-type lectin, legume-like lectin and jacalin-like lectin (JRL).

Host encounter and initial penetration stage (6 h)

Most genes that were highly expressed at 6 h were related to cytoskeleton organization (e.g. beta tubulin, tubular mastigoneme protein, actin and related proteins), signal transduction (e.g. guanine nucleotide binding protein, serine threonine protein phosphatase, G protein) and stress response genes (e.g. heat shock protein 70 [HSP70]). For energy production, the expression of glycolysis-related genes (e.g. glucokinase) was comparably high and the genes related to mitochondrial metabolic pathways (e.g. citrate cycle [TCA cycle], electron transport chain, ATP synthease complex) were downregulated. Table 4.3.3 summarizes the presence of genes for mitochondrial pathways in *Amoebophrya*.

Maturation stage (96 h)

At the time point of 96 h, several oxidative pathway components, including peroxiredoxin, thioredoxin and glutaredoxin, were expressed during infection. The genes related to cell division, reconstruction and proliferation (e.g. chromosome segregation ATPases, meiosis-specific nuclear structural protein) were observed in this maturation stage. From 6 to 96 h, the absolute number of expressed stress response genes (including those encoding heat shock protein 90, DnaJ, Cpn10 and cold shock protein) increased.

There was an abundant range of proteases (in total 30 different Pfam domains and families) expressed by *Amoebophrya,* including ATP-dependent protease, cathepsin cysteine protease, subtilisin-like serine protease, cysteine protease, serine protease, and ubiquitin-specific proteases. Table S4.3.4 in the supplemental material lists the proteases identified in *Amoebophrya* data set and are discussed below.

Table 4.3.3: Genes related to mitochondrial metabolic pathways[a]

Functional Category	Definition	Name	E-value	Read count T0	T6	T96
TCA cycle	Aconitase	Contig_2081	2.00E-125	30	7	70
	Citrate synthase	Contig_2141	1.40E-50	19	0	17
	Citrate synthase	Contig_2724	1.50E-61	10	0	20
	Malate/lactate dehydrogenases	Contig_1233	8.10E-26	25	0	12
	Malate/lactate dehydrogenases	Contig_7089	3.10E-45	5	0	12
	NAD/NADP transhydrogenase alpha subunit	Contig_528	1.00E-12	123	0	105
	NAD/NADP transhydrogenase beta subunit	Contig_4871	3.60E-75	28	0	11
	Succinate dehydrogenase/fumarate reductase, Fe-S protein subunit	Contig_7660	2.30E-55	33	48	72
	Succinate dehydrogenase/fumarate reductase, flavoprotein subunit	Contig_5497	3.30E-23	26	1	21
	Succinate dehydrogenase/fumarate reductase, flavoprotein subunit	Contig_6096	5.50E-12	13	0	14
	Succinate dehydrogenase/fumarate reductase, flavoprotein subunit	Contig_4542	1.40E-17	8	0	0
	Succinyl-CoA synthetase, beta subunit	Contig_4953	5.80E-95	19	0	17
	Succinyl-CoA synthetase, beta subunit	Contig_11475	3.50E-23	6	0	0
Electron transport chain	Cytochrome *b*	Contig_5154	3.60E-08	9	0	17
	Cytochrome *b*	Contig_2214	4.50E-18	6	0	24
	Cytochrome *b*	Contig_11545	5.00E-18	4	4	6
	Cytochrome *c1*	Contig_14469	2.60E-30	9	3	0
	Cytochrome *c2*	Contig_7748	6.10E-44	28	0	47
	Electron transfer flavoprotein, alpha subunit	Contig_17848	9.20E-58	8	5	2
	Electron transfer flavoprotein, beta subunit	Contig_41980	6.80E-35	4	3	7
	NADH dehydrogenase, FAD-containing subunit	Contig_410	1.10E-24	73	0	38
Additional dehydrogenases	Glycerol-3-phosphate dehydrogenase	Contig_10030	1.80E-16	3	0	22
ATP synthase complex	F0F1-type ATP synthase, alpha subunit	Contig_3133	1.00E-175	5	0	0
	F0F1-type ATP synthase, alpha subunit	Contig_21601	1.30E-74	6	0	0
	F0F1-type ATP synthase, beta subunit	Contig_7952	1.80E-61	23	0	1
	F0F1-type ATP synthase, beta subunit	Contig_7654	1.00E-137	17	0	4
	F0F1-type ATP synthase, delta subunit	Contig_3817	8.80E-15	8	2	0
	F0F1-type ATP synthase, delta subunit	Contig_41749	1.00E-15	0	0	12
	F0F1-type ATP synthase, gamma subunit	Contig_41538	2.80E-13	0	0	6
	F0F1-type ATP synthase, subunit c/Archaeal/vacuolar-type H+-ATPase, subunit K	Contig_13688	1.30E-08	3	1	44
	F0F1-type ATP synthase, subunit c/Archaeal/vacuolar-type H+-ATPase, subunit K	Contig_13690	2.10E-11	1	0	24
ATPase	Archaeal/vacuolar-type H+-ATPase subunit A	Contig_20433	3.50E-82	11	4	1
	Archaeal/vacuolar-type H+-ATPase subunit A	Contig_43169	2.50E-52	0	0	5
	Archaeal/vacuolar-type H+-ATPase subunit A	Contig_44048	2.30E-90	2	3	2
	Archaeal/vacuolar-type H+-ATPase subunit C	Contig_8043	1.10E-09	7	0	0
	Archaeal/vacuolar-type H+-ATPase subunit D	Contig_6053	5.90E-43	2	4	4
	Archaeal/vacuolar-type H+-ATPase subunit F	Contig_1984	2.00E-11	8	0	0
	Archaeal/vacuolar-type H+-ATPase subunit I	Contig_1670	1.50E-10	2	0	6
	Archaeal/vacuolar-type H+-ATPase subunit I	Contig_1671	2.90E-10	0	0	3

[a] T0, pure dinospores; T6, initial penetration stage; T96, maturation stage; FAD, flavin adenine dinucleotide.

4.3.5 Discussion

Transcriptional activity of genes can be used to assess functional changes within organisms. Dinoflagellates use *trans*-splicing to generate mature mRNA species (Lidie & van Dolah 2007). This mechanism may influence the posttranscriptional fate of transcripts and thus might impair the correlation of the amount of transcripts to proteins. Despite this potential decoupling of functions from transcription, transcriptional changes indicate a reaction of an organism to environmental stimuli. My study demonstrates that the host-parasitoid interaction causes significant changes in parasitoid gene expression over the time course of the infection. The major findings are summarized in Fig. 4.3.4 where I depict the occurrence of genes related to cell adhesion, glycan-related enzymes and lectins, genes involved in energy metabolism and signal transduction, and genes encoding for cytoskeleton proteins, heat shock proteins and proteases.

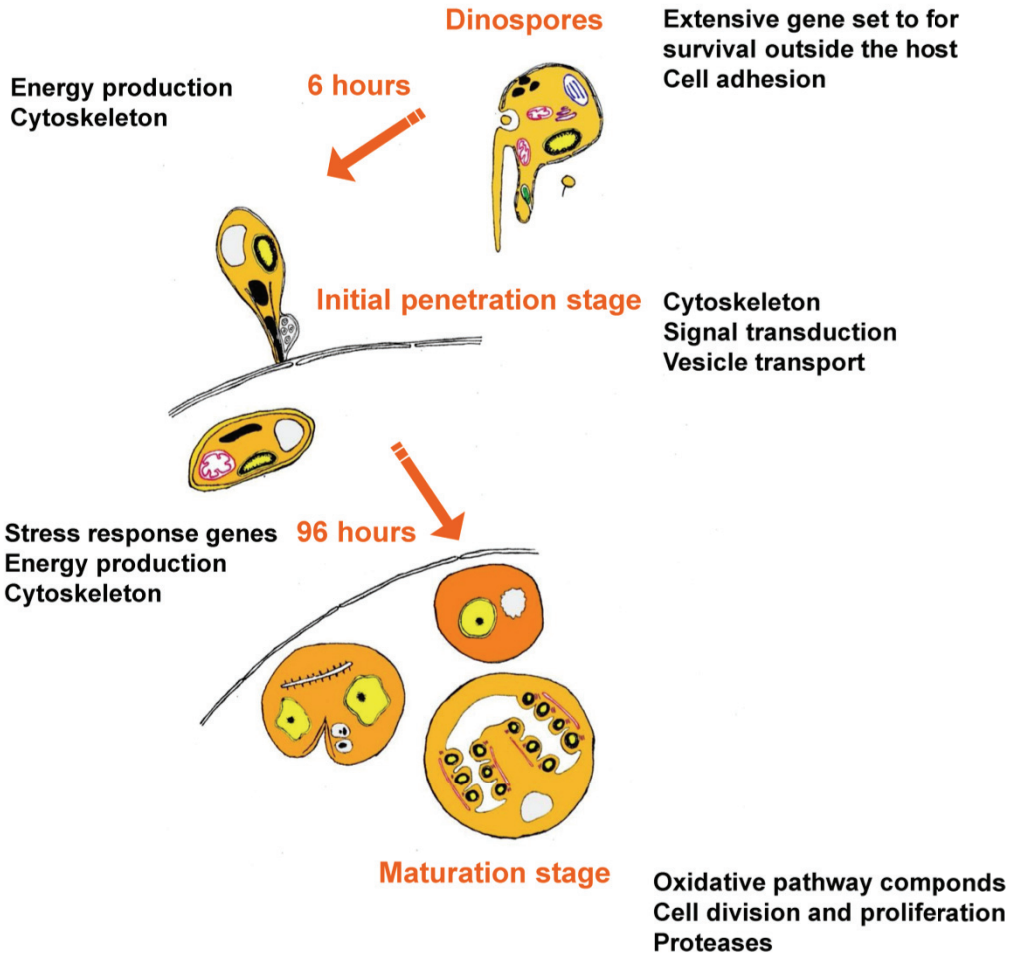


Fig. 4.3.4: Hallmarks of the biology of *Amoebophrya* during infection.

The importance of proteases for a parasitic lifestyle has been shown for *Perkinsus marinus,* a protozoan parasite of the oyster. The Perkinsozoa (*P. marinus*) are the earliest-diverging group of the dinoflagellates, with the phylogenetic position at the base of the dinoflagellate branch (Reece *et al.* 1997; Saldarriaga *et al.* 2003). Because of their close relationship, the Perkinsozoa are considered a key taxon to investigate parasitism within the Alveolata (Leander *et al.* 2003). In *P. marinus,* extracellular serine proteases are produced to degrade oyster protein in order to gain nutrients to support the parasite basic cell function (PEYRE *et al.* 1995), and to effect oyster host defenses (Garreis *et al.* 1996). Parasites with increased expression of low-molecular-weight protease (LMP: 30 to 45 kDa) appeared to be more infective of highly susceptible oysters (*Crassostrea virginica*) in host-supplemented medium (Earnhart *et al.* 2004; MacIntyre *et al.* 2003). Some of these serine proteases were isolated and characterized, and were found to be encoded by subtilisim-like gene(s) from the *P. marinus* genome (Brown & Reece 2003). In the data set, I also found high expression of serine protease and subtilisin-like serine protease, together with cysteine protease, cathepsin cysteine protease, and ubiquitin-specific proteases by *Amoebophrya*, which might be involved in host-parasitoid interactions of *Amoebophrya* in a manner similar to that seen with *Perkinsus* (see Data set S3.1 in the supplemental material; see also Table S4.3.4). Surprisingly, diverse proteases in *Amoebophrya* showed expression during all time points analysed. Thus, their expression appears not to be triggered by contact or entry of the parasitoid. Rather, the proteases have been already produced and stored to enable a rapid provision of these degrading enzymes after entering the host.

Dinospores (0 h)

This study represents the first broad-scale gene expression study of pure dinospores. In general, a large number of genes were expressed during this stage, indicating that the free-living and actively swimming dinospores, although not dividing, make use of an extensive and specific gene set. In particular, glycan-related enzymes and carbohydrate-binding proteins (lectins) serve as an important cellular surface recognition mechanism in host-parasite interactions (Joseph *et al.* 2010; Martel 2009; Roberts *et al.* 2006). Lectins of free-living, nonparasitic, heterotrophic dinoflagellates are involved in prey recognition and discrimination prior to phagocytosis (Wootton *et al.* 2007). This study revealed that two types of lectins were present in *Amoebophrya* dinospores: ricin-type

beta-trefoil lectin and C-type lectin. Pfam analysis detected similarities to two types of plant lectins: legume-like lectin and jacalin-like lectin. The jacalin-liked lectins (JRLs) are common name of β-prism-I lectins, which first identified from the seeds of jackfruit (*Artocarpus heterophyllus*) (Bunn-Moreno & Campos-Neto 1981). Lectins of this family have been found to exhibit a repertoire of functions due to their high sensitivity in recognizing cell surface carbohydrates (Raval *et al.* 2004). The expression of these surface lectins therefore most likely contributes to the attachment of the parasite to the host cell. Hence, parasitic dinoflagellates may use lectins in the same way as their apicomplexan counterpart, whereas the lectin function probably has also evolved towards prey recognition in heterotrophic dinoflagellate lineages (Wootton *et al.* 2007).

Similarities between 0 and 6 h of infection

During this stage of the parasitoid life cycle, the dinospores have a demand of energy due to active swimming for the purpose of finding and penetrating their host and maybe already for the initiation of the trophont development. To date, not much is known about the energy metabolism of *Amoebophrya*. A comparison with their sister lineage, apicomplexa (diverging ~800 to 900 million years ago), may help to understand the nutritional mode of these parasites (Baldauf 2008; Curtis *et al.* 2012; Hackett *et al.* 2005). The apicomplexa include many human and animal pathogens such as *Plasmodium* (causative agent of malaria) (Walker *et al.* 2011), *Toxoplasma* (parasitic toxoplasmosis disease) (Tenter *et al.* 2000) and *Crytosporidium* (diarrhea in mammals). In the sexual stage, *Plasmodium* lives in the mosquito host and shows greater activity in electron transport and oxidative phosphorylation (Danne *et al.* 2013; Learngaramkul *et al.* 1999; Plattner & Soldati-Favre 2008). While in the asexual stage, the parasite lives in a glucose-rich environment (human blood) and gets sufficient ATP through the glycolysis pathway alone (Seeber *et al.* 2008; Van Dooren *et al.* 2006). I detected genes for mitochondrial pathways in *Amoebophrya* at all three time stages investigated, albeit at different expression levels (Table 4.3.3). At 6 h, levels of expression of glucokinase were comparably high and many genes related to mitochondrial metabolic pathways were downregulated. This may suggests the importance of glycolysis and thus of anaerobic energy production in the initial penetration stage of *Amoebophrya.* At 0 and 96 h, the high expression of genes for the TCA cycle, electron transport chain and ATP synthase complex indicated that oxidative phosphorylation may play a key role in generating ATP in dinospores and mature trophonts of parasitoid *Amoebophrya*. Unlike apicomplexa, the parasitoid *Amoebophrya* may

need ATP synthesis from complete carbohydrate oxidation in mitochondria, congruent with the observation that mitochondria of the parasitoid *Amoebophrya* infecting *Akashiwo sanguinea* ranged in size from relatively large in dinospores to small in initial trophonts (at 12 and 36 h) and elongated in mature trophonts (at 48 h) (Miller *et al.* 2012).

I found the majority of sequences encoding key components of mitochondrial pathways in *Amoebophrya* (Table 4.3.3), but the sequence encoding mitochondrial pyruvate dehydrogenase (PDH) was missing. In general, the PDH complex links glycolysis to the TCA cycle by converting pyruvate to acetyl coenzyme A (acetyl-CoA), which is thought to be a key regulator in the carbon flux in mitochondria (Fernie *et al.* 2004). There is evidence that the apicomplexa and some dinoflagellate lost their mitochondrial PDH, stopping the conversion of pyruvate into acetyl-CoA (Butterfield *et al.* 2013; Danne *et al.* 2013; Seeber *et al.* 2008; Van Dooren *et al.* 2006). Therefore I expect that the source of acetyl-CoA in the mitochondrion of *Amoebophrya* is not pyruvate but might be found in degradation of branched-chain amino acids or β-oxidation of fatty acids. This shedding of metabolic functions of *Amoebophrya* may reflex its efficient parasitoid energy production, similar to that revealed in well-supported examples of apicomplexa and other parasitic protists. For instance, *Helicosporidium,* the obligate parasitic green alga, has lost nearly all genes associated with light harvesting and photosystems, but contains an almost complete pathway for carbon fixation (Pombert *et al.* 2014).

Initial penetration stage (6 h)

The majority of the genes overrepresented at the time point of 6 h were related to cytoskeletal organization and signal transduction, pointing to processes that enable recognition, adhesion, and penetration of the host. The high expression of cytoskeleton-related genes implies its strong role during infection. This is supported by the ultrastructure research by Miller et al (Miller *et al.* 2012), who observed a motility system based on microfilament polymerization that was used to enter the host cytoplasm.

The expression of heat shock protein 70 (HSP70) increased at 6 h of the infection (see Table S4.3.3 in the supplemental material). This observation is congruent with further literature data where it has been shown that the HSP70 gene is highly expressed in the active

to latent form in the parasitic apicomplexa *Toxoplasma gondii* (Lyons & Johnson 1995; Weiss *et al.* 1998), in *P. marinus,* and in the parasitic ciliate *Crytocaryon irritans* (Lokanathan *et al.* 2010) to overcome the stress from the host environment (Joseph *et al.* 2010). In addition, further heat shock proteins (HSP90, DnaJ, Cpn10) and a cold shock protein (CSD) were expressed in *Amoebophrya* both at the initial penetration stage at 6 h and at the late maturation stage at 96 h and also in dinospores (see Data set S4.3.1 in the supplemental material). Heat shock proteins are highly conserved proteins, which play an essential role in response to stress (Feder & Hofmann 1999; Parsell & Lindquist 1993; Peyretaillade *et al.* 1998) and many heat shock proteins, such as the small HSP (Horwitz 1992) and HSP90, are also chaperones essential for activating signaling proteins and protein folding in the eukaryotic cell (Young *et al.* 2001). Cold shock proteins are associated with posttranscriptional regulation in eukaryotes (Mihailovich *et al.* 2010), and have been found to play an important role in regulating translation in the dinoflagellate *Lingulodinium* (Beauchemin *et al.* 2012). Taken together, this data indicate that the heat shock gene products in *Amoebophrya* may be needed to overcome the host defense response.

Late infection and maturation stage (96 h)

Genes expressed after 96 h reflect the fast cell growth and nuclear division of the parasitoid during this life stage. In total, 7,388 genes, including a wealth of genes related to cell division, reconstruction and proliferation, are expressed at the 96-h stage, (see Data set S4.3.1 in the supplemental material). During this life stage, the parasitoid trophont undergoes karyokinesis, and each cell forms its own flagella (Coats & Park 2002).

Analogous to *P. marinus*, *Amoebophrya* sp. might experience an oxidative burst reaction of the host cell that is counteracted through the expression of antioxidant genes (Joseph *et al.* 2010). *P. marinus* uses superoxide dismutases (SOD) to protect itself from reactive oxygen intermediates (ROIs) generated by the host's oxidative enzymes. SODs were found at the 96-h maturation stage and at the dinospore stage in *Amoebophrya*. With the rapid growth of the parasitoid in the host cell, the major host defense against pathogens may thus be production of reactive oxygen species (ROS) with a toxic effect on the pathogen by damaging DNA, proteins and lipids (Imlay 2003; Miura *et al.* 2012).

4.3.6 Conclusions

This analysis shows that parasitism is at least partly driven by common mechanisms in different eukaryote groups. Differentially expressed genes are associated with different metabolic pathways at each time point such as those corresponding to proteases, the parasitoid-host cellular surface recognition mechanism, antioxidant defense, energy production such as glycolysis, TCA cycle, and other mitochondrial proteins. Further comparative analyses of the dinoflagellate parasitoid with the closely related Apicomplexa will show whether these similarities have a common evolutionary basis within the Alveolata.

4.4 Transcriptomic profiling of *Alexandrium fundyense* during physical interaction with or exposure to chemical signals from the parasite *Amoebophrya* sp.

4.4.1 Abstract

Toxic microalgae have their own pathogens, and understanding the way in which these microalgae respond to antagonistic attacks may provide information about their capacity to persist during harmful algal bloom events. Here, I compared the effects of the physical presence of the parasite *Amoebophrya* sp. and exposure to waterborne cues from cultures infected with this parasite, on gene expression by the toxic dinoflagellates, *Alexandrium fundyense*. Compared with control samples, a total of 14,882 *Alexandrium* genes were differentially expressed over the whole-parasite infection cycle at three different time points (0, 6, and 96 h). RNA sequencing analyses indicated that exposure to the parasite and parasitic waterborne cues produced significant changes in the expression levels of *Alexandrium* genes associated with specific metabolic pathways. The observed upregulation of genes associated with glycolysis, the tricarboxylic acid cycle, fatty acid β-oxidation, oxidative phosphorylation and photosynthesis suggests that parasite infection increases the energy demand of the host. The observed upregulation of genes correlated with signal transduction indicates that *Alexandrium* could be sensitised by parasite attacks. This response might prime the defense of the host, as indicated by the increased expression of several genes associated with defense and stress. This findings provide a molecular overview of the response of a dinoflagellate to parasite infection.

4.4.2 Introduction

Phytoplanktonic organisms are important primary producers that play an essential role in food webs and energy fluxes in marine ecosystems; however, certain toxic phytoplankton species can form harmful algal blooms. The dinoflagellate *Alexandrium* is one of the best known bloom-forming and toxin-producing species responsible for paralytic shellfish poisoning (PSP) outbreaks (Anderson *et al.* 2012a; Cembella 2003; Hallegraeff 1993), and some studies show that the associated toxins act as defense compounds against copepod grazers (Selander *et al.* 2006; Wohlrab *et al.* 2010; Yang *et al.* 2010). In addition to toxins, the dinoflagellates, *Alexandrium,* can also produce allelochemicals that affect species interactions due to cell lysis of potential grazers and/or competitors (John *et al.* 2014; Tillmann *et al.* 2008; Tillmann & John 2002). However, very little is known about the roles of the toxic and allelochemical compounds produced by *Alexandrium* as defense against pathogens (Anderson *et al.* 2012a).

Microalgae can be infected by a broad variety of organisms, such as viruses, bacteria and parasites (Chambouvet *et al.* 2008; Kim 2006; Velo-Suárez *et al.* 2013). The parasites *Amoebophrya ceratii* (Syndiniales) and the host *Alexandrium fundyense* (Gonyaulacales) are both Dinophyceae (Alveolata) (Guillou *et al.* 2008). Field studies show that a large variety of host species are infected by the parasite *Amoebophrya*, including a wide taxonomic range of harmful dinoflagellates (Li *et al.* 2014; Park *et al.* 2013; Siano *et al.* 2010). There is also evidence that *Amoebophrya* infections play a pivotal role in controlling host mortality and can regulate the dynamics of dinoflagellate blooms at high infection rates (Chambouvet *et al.* 2008; Coats *et al.* 1996); these effects coincide with the life cycle transition of the host *Alexandrium* from vegetative division to sexual fusion (Velo-Suárez *et al.* 2013).

Together with grazing by microzooplankton, parasite infection is an important top-down control mechanism for bloom-forming dinoflagellates (Montagnes *et al.* 2008). *Amoebophrya* is a model parasitic organism that can be cocultured with its host, *Alexandrium,* in the laboratory with an infective cycle of approximately 4 days (Lu *et al.* 2014). Infection by *Amoebophrya* is initiated by penetration of the parasitic dinospores into the host cells (Cachon 1964; Miller *et al.* 2012). Once inside the cytoplasm or nucleus (depending on the specific host and parasitic strains), the parasite starts to feed (the trophont stage). The trophont

increases in size until sequential nuclear divisions and flagellar replications ultimately form an intracellular and multicellular 'beehive' stage inside the cytoplasm or nucleus of the host cell (the sporocyte) (Cachon 1964). The mature sporocyte ruptures the cell wall of the host, and most develop into a short-lived vermiform stage that soon divides into numerous free-living infectious dinospores (Coats & Bockstahler 1994; Coats & Park 2002).

Chemical signals can determine feeding behaviour, habitat selection and induced defense in a wide range of aquatic organisms (Chivers & Smith 1998; Hay 2009). In marine species, the induction of defense can also be triggered in response to waterborne cues emitted by predators (Chivers & Smith 1998; Roberts *et al.* 2011; Toth & Pavia 2000). A recent study investigated the potential mechanisms involved in chemically mediated predator-prey interactions, and accumulating evidence suggests that the marine dinoflagellate *Alexandrium* can recognize and distinguish predators and respond when exposed to waterborne cues from conspecific, threatening copepod grazers (Roberts *et al.* 2011; Selander *et al.* 2006; Wohlrab *et al.* 2010). A single study has described the ability of the dinoflagellate *Alexandrium ostenfeldii* to form temporary cysts as a response to waterborne cues from the parasitic Perkinsozoa *Parvilucifera infectans* (Toth *et al.* 2004).

Relatively few studies have examined the transcriptomic responses of dinoflagellates infected with parasites (Bachvaroff *et al.* 2009). Consequently, little is currently known about the importance of chemical cues that may prime host responses towards parasites. The objective of this study was to investigate the molecular mechanisms that underlie the responses of *Alexandrium* to parasite infection and to discriminate these responses from those elicited by parasitic waterborne cues. To this end, I used RNA sequencing (RNA-seq) to compare the transcriptional responses of the host *Alexandrium* to the presence of the parasite *Amoebophrya* and waterborne chemical cues from this parasite. Annotation of the final data revealed that a large number of genes associated with this host-parasite interaction are involved in energy conversion metabolic pathways, signal transduction, and defense mechanisms.

4.4.3 Materials and Methods

Cultures

The *Amoebophrya* (AT5.2) parasite strain [the term parasitoid is also used (Lu *et al.* 2014)] was isolated from *Alexandrium* cells sampled from the Gulf of Maine, USA (Chambouvet *et al.* 2011b), and was used to infect the *Alexandrium fundyense* strain [formerly described as *A. tamarense* (John *et al.* 2014)] isolated from the North Sea coast of Scotland (Alex5; RCC3037) (Tillmann *et al.* 2009). To understand the host-parasite mechanisms, I needed a strain, which can be reliably infected with high rates. I tested the prevalence of the *Amoebophrya* strain to different population and strains of *Alexandrium* in our laboratory and identified no difference in the infection percentage of each population. Alex5; RCC3037 was chosen in this study, because its infection percentage by the parasite *Amoebophrya* strain was among the highest ones. This strain of *Alexandrium* has also the advantage that its genetic/genomic background is well studied (Alpermann *et al.* 2009; Wohlrab 2013) and it does not produce allelochemicals which may affect the host-parasite infection (Tillmann *et al.* 2009). The life cycle of the parasite was 4 days. All cultures were grown at 15°C in K-medium (Keller *et al.* 1987), with cool-white fluorescent lamps providing photon irradiation of 150 µmol m^{-2} s^{-1} on a light:dark cycle of 14 h:10 h.

Influence of parasitism on the host

The parasite infection experiment covered one complete parasite life cycle (4 days) and included three harvesting time points (0, 6, and 96 h). Triplicate exponential phase cultures of *Alexandrium* (400 ml) at a concentration of approximately 1,000 cells ml^{-1} were prepared in 500 ml Erlenmeyer flasks. Triplicate cultures containing the host only (400 ml *Alexandrium* at a concentration of 1,000 cells ml^{-1}) served as no-parasite controls (-P). Portions (100 ml) of the *Alexandrium* cultures were used for RNA extraction at the 0-h time point. The remaining cultures were used for the parasite infection experiment. Infection of the host culture was performed following the procedures of Coats and Park (2002). Infective parasite dinospores were harvested from infected host cultures on Day 4 by gravity filtration through a 10-µm pore size mesh. The harvested dinospores were examined under a microscope to ensure the absence of host cell contamination and were then inoculated immediately into the triplicate *Alexandrium* cultures at a parasite: host ratio of 10:1 (Lu *et al.*

2014). The cultures were incubated for 6 or 96 h after adding the dinospores. At each time point, samples collected from the triplicate cultures of *Alexandrium* parasite-infected treatment (+P) and no-parasite control (-P) were used for fixation and cell counting, parasite prevalence assessment, RNA extraction and sequencing, and PSP toxin analysis.

Influence of waterborne cues on the host

To examine the response of *Alexandrium* to waterborne cues from the parasites, and to discriminate the potential wounding impact from the response to lysed *Alexandrium* cells, three different incubations were performed: (i) parasite-infected waterborne medium (+WP): the host was treated with medium by gravity filtration through a 0.2-µm pore size mesh from *Alexandrium* cells that had been infected by parasites; (ii) lysed host cells (+A): the host was treated with medium from host cells that had been lysed using ultrasound for 2 min; and (iii) No-parasite *Alexandrium* waterborne medium; used as control (-WP): the host was treated with filtered medium from an exponentially growing control *Alexandrium* culture. *Alexandrium* cultures (400 ml) at a concentration of approximately 1,000 cells ml^{-1} were incubated in nine 500 ml Erlenmeyer flasks (three per incubation type). In each experiment, the medium with waterborne cues was replenished at the 24-, 48-, and 96-h time points. The experiment examining the effect of waterborne cues on the host *Alexandrium* also covered one complete parasite life cycle (4 days) and included the same three harvesting time points (0, 6, and 96 h) as the previous parasite infection experiment. At each time point, samples were collected from each culture and used for fixation and cell counting, RNA extraction and sequencing, and PSP toxin analysis.

Fixation and growth rate calculation

Samples (10 ml) from each experiment were fixed with Lugol's solution (10 g of potassium iodide and 5 g of iodine in 100 ml of distilled water) at a final concentration of 2% (Tillmann *et al.* 2009), and three 1 ml aliquots were counted under an inverted microscope (Axiovert 200M; Zeiss, Göttingen, Germany) after sedimentation in chambers. A minimum of 400 cells per sample were counted. Parasite infection and the release and survival of dinospores was followed by examining the persistence of the natural autofluorescence of the parasite under a microscope (Axiovert 200M) (Coats & Bockstahler 1994). The growth rate of *Alexandrium* was calculated using the following formula (Guillard 1973; Tillmann *et al.*

2009), where μ is the growth rate (d^{-1}), t is the sampling day, and N_1 and N_2 are the abundances of *Alexandrium* at t_1 and t_2, respectively:

$$\mu = \frac{1}{t_2 - t_1} ln \frac{N_2}{N_1}$$

RNA preparation, library construction, and sequencing

Total RNA was extracted from parasite-infected (+P), parasite control (-P), waterborne cue-treated (+WP), lysed host cell-treated (+A) and waterborne cue control cultures (-WP) at three time points (0, 6, and 96 h). The 100 ml samples were centrifuged at 4°C for 10 min. The supernatants were decanted, and the resulting cell pellets were resuspended immediately in 1 ml of hot (60°C) TriReagent (Sigma-Aldrich, Steinheim, Germany), according to the manufacturer's protocol. RNA purification, including on-column DNA digestion, was performed using the RNeasy Kit (Qiagen, Hilden, Germany). The quality and quantity of the RNA were determined using a NanoDrop ND-100 spectrometer (PeqLab, Erlangen, Germany) and a RNA Nano Chip assay on a 2100 Bioanalyzer device (Agilent Technologies, Böblingen, Germany). For the construction of the reference transcriptome, aliquots of the triplicate RNA samples from each experiment were pooled and sequenced as a 100-bp paired-end Illumina library. The raw reads were assembled to yield the *Alexandrium* transcript reference sequences. Each RNA sample was then sequenced independently as a 50-bp single-end Illumina library for the expression analyses (see Table S4.4.1 in the supplemental material).

Libraries were prepared using the TruSeq RNA Sample Preparation Kit (Illumina, San Diego, CA, USA), according to the manufacturer's protocol. To obtain longer molecules for paired-end sequencing, the fragmentation time was reduced to 4 min. The libraries were quality checked and quantified using a Bioanalyzer 2100 device and a DNA Chip assay, and then sequenced using a HiSeq2000/2500 instrument (high-output mode) in either single-read/50 cycle or paired-end/2 × 100 cycle mode (Bentley *et al.* 2008). Multiplexing was performed using three, four, or five libraries per lane. Sequence information was extracted using the CASAVA v1.8.2 software (Illumina) in FASTQ format. The analysis produced 33 datasets for single-end sequencing and six data sets for paired-end sequencing (see Table S4.4.1 in the supplemental material).

Analysis of RNA-seq data

No complete genomes of *Alexandrium* species are currently available; therefore, an *Alexandrium* reference transcriptome was constructed by merging sequence information from three sources: (i) a *de novo* transcriptome of *Alexandrium* (122, 219 contigs; Table S4.4.2 in the supplemental material), which was assembled using the CLC Genomics Workbench (CLC Bio, Aarhus, Denmark) with default settings from paired-end Illumina RNA-seq reads of six *Alexandrium* samples at the 6- and 96-h time points and single-end reads at the 0-h time point (Table S4.1 in the supplemental material); (ii) 44,024 expressed sequence tags of *Alexandrium* from our internal database (Wohlrab 2013); and (iii) an expressed sequence tag dataset containing 29,995 *Alexandrium* contigs from a previous host-parasite infection study (Lu *et al.* 2014). The overlapping and identical contigs from these data sources were merged to generate a total of 147,835 unique transcripts in the *Alexandrium* reference transcriptome (Fig. S4.4.1 in the supplemental material). All contigs were mapped to the parasite *Amoebophrya* genome sequence data (SRP067624) of the same strain. Matching sequences were excluded from the dataset.

Table S4.4.1 (in the supplemental material) provides an overview of the number of RNA-seq reads per *Alexandrium* sample examined. After quality control, a total of 1.3 billion reads were aligned to the *Alexandrium* reference transcriptome using the CLC Genomics Workbench to obtain the read counts using default settings. To determine differential gene expression, the read counts were analysed using the DESeq package in R (Anders & Huber 2010). Size factor estimation and normalisation were performed using the 'estimateSizeFactors' and 'estimateDispersions' functions, respectively. Differentially expressed contigs were detected by a negative binomial test using the 'nbinomTest' function. Transcripts with a false discovery rate-adjusted P-value < 0.05 were considered statistically significant and used for annotation.

Genes that were significantly differentially expressed between treated and untreated *Alexandrium* samples were annotated by homology searches following the Trinotate annotation suite guidelines (Grabherr *et al.* 2011; Haas *et al.* 2013). Putative orthologs were predicted from reciprocal best BLAST hits; peptide sequences were predicted using the Trinity TransDecoder package; and protein families, signal peptides, and transmembrane domains were identified using Pfam (Punta *et al.* 2012), SignalP (Petersen *et al.* 2011), and

TMHMM (Krogh *et al.* 2001), respectively. The data were compared to the eggNOG/GO (Powell *et al.* 2012), Gene Ontology (Ashburner *et al.* 2000) and KEGG (Kanehisa *et al.* 2012) databases. The differentially expressed *Alexandrium* genes were translated into amino acid sequences using the Virtual Ribosome package (http://www.cbs.dtu.dk/services/VirtualRibosome), and the Batch Web CD-search tool (http://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi) was used to assign eukaryotic orthologous groups (KOGs). Significant enrichments of the transcripts were tested by calculating the *P*-value from a hypergeometric distribution at the background level of all KOGs (Subramanian *et al.* 2005). KOGs were considered significantly enriched for a given experimental time point when the test statistics gave a P-value < 0.05.

PSP toxin analysis

Samples (50 ml) of treated and untreated cultures were used for PSP toxin analyses, which were performed as described previously (Krock *et al.* 2007). Briefly, the 50 ml cell culture was centrifuged for 15 min at 3,220 x g (4,000 rpm). The supernatant was discarded and the cell pellet was added to 1 ml of sterile seawater, transferred to a 2 ml tube, and centrifuged for 10 min at maximum speed. After removing the seawater, the pellet was transferred to a tube containing 0.9 g of Lysing Matrix D (Thermo Savant, Illkirch, France). The cells were homogenised by reciprocal shaking in a Bio101 FastPrep instrument (Thermo Savant) at speed 6.5 for 45 s. The samples were then centrifuged for 15 min at 4°C and 13,000 g. The supernatant was passed through a spin filter (pore size 0.45 mm) by centrifugation for 30 s at 3,000 g. The filtrate was analysed by high-performance liquid chromatography with fluorescence detection (Krock *et al.* 2007).

4.4.4 Results

Growth of Alexandrium

At late growth stages, the growth rates of *Alexandrium* treated with the *Amoebophrya* parasite (+P) or parasitic waterborne cues (+WP) were lower than those of their corresponding controls (parasite control;-P and waterborne cue control;-WP, respectively) (Fig. 4.4.1 and Table S4.4.3 in the supplemental material). The differences between the growth rates of *Alexandrium* in the three treatment groups (+P, +WP and +A) were significant (ANOVA: $f = 10.85$, $P < 0.05$). Compared with that of the parasite control culture (-P), the growth rate of the parasite-infected culture (+P) was reduced significantly at the 24-h (ANOVA: $f = 16.55$, $P < 0.05$), 72-h (ANOVA: $f = 43.41$, $P < 0.01$), and 96-h (ANOVA: $f = 109.5$, $P < 0.01$) time points. Similarly, compared with that of the waterborne cue control (-WP), the growth rate of the parasitic waterborne cue-treated culture (+WP) was reduced significantly at the 48-h (ANOVA: $f = 17.85$, $P < 0.05$), 72-h (ANOVA: $f = 16.25$, $P < 0.05$), and 96-h (ANOVA: $f = 12.68$, $P < 0.05$) time points. By contrast, the growth rate of the culture exposed to the lysed *Alexandrium* cells (+A) (adding more potential organic food supply) was significantly higher than that of the control at the early infection stages (6 h, ANOVA: $f = 21.21$, $P < 0.01$; 24 h, ANOVA: $f = 54.42$, $P < 0.01$) (Table S4.4.3 in the supplemental material).

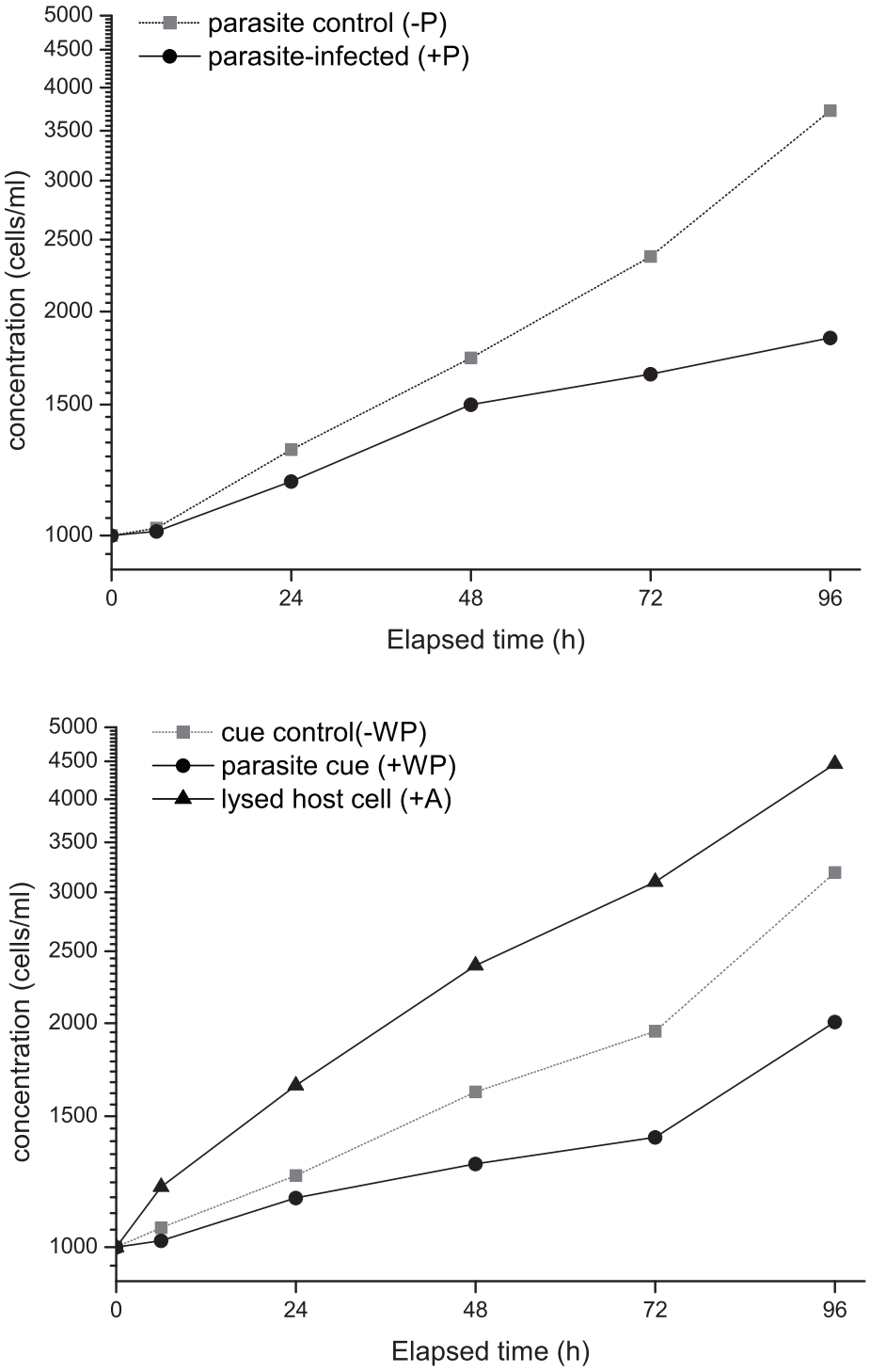


Fig. 4.4.1: Growth of *Alexandrium* in the treated (parasite-infected, +P; waterborne parasite cue-treated; +WP, and lysed host cell-treated; +A) and control (parasite control, -P; and cues control, -WP) groups at the indicated time points.

Differentially expressed genes

A total of 14,882 genes were significantly differentially expressed (adjusted *P*-value < 0.05) between the treated and corresponding control samples at the 6 and 96 h time points (Fig. 4.4.2). At the early infection stage (6 h), the parasite (+P) induced lower differentially expressed genes compared to other conditions. Interestingly, parasite cue (+WP) and cue control (-WP) induced a relatively important number of upregulated genes (123) compared to the host-lysed cell (+A) (40 upregulated and 83 downregulated).

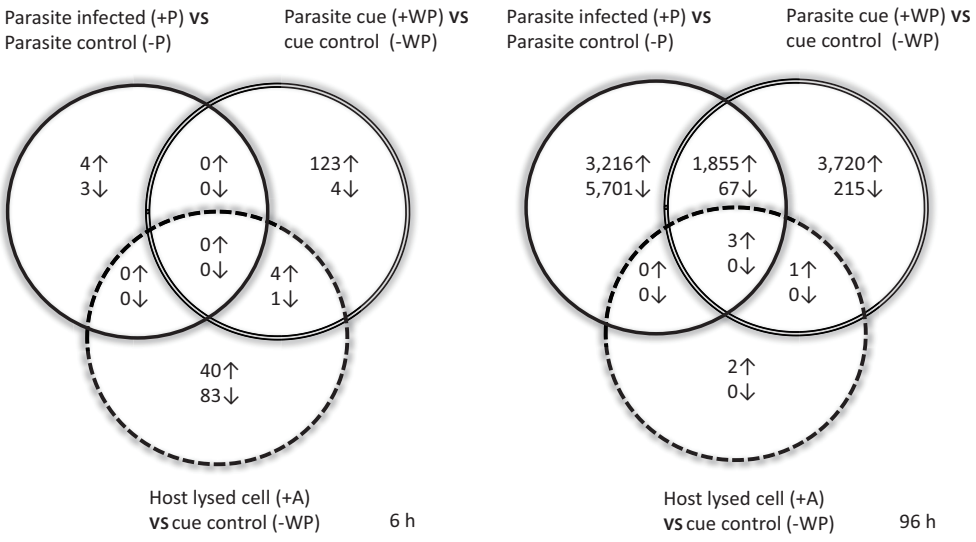
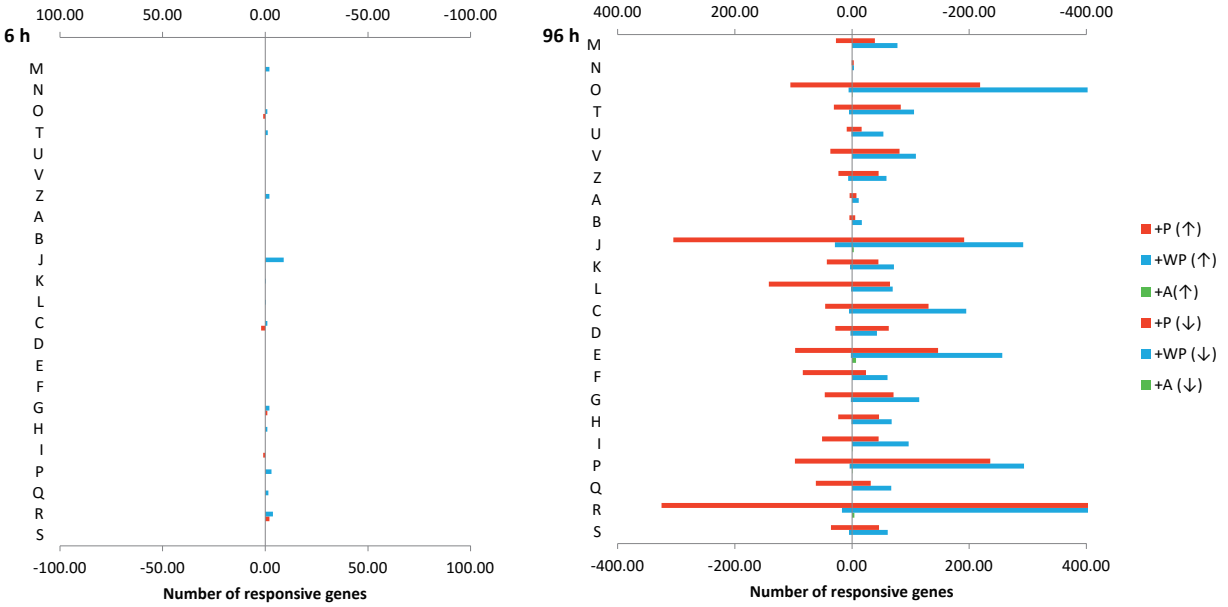


Fig. 4.4.2: Differential gene expression in *Alexandrium* treated with parasites (+P), waterborne cues from the parasites (+WP), or lysed host cells (+A).

At the late infection stage (96 h), a large number of genes were differentially expressed in *Alexandrium* treated with the parasite (+P) or parasitic waterborne cues (+WP), whereas treatment with lysed host cells (+A) produced few changes in gene expression (two genes only). Hosts infected by parasites (+P) responded differently compared to hosts induced by parasite cues (+WP), by having much more down-regulated genes (5,701 compared to 215, respectively) but similar numbers of up-regulated genes (with similar transcripts for about half of them between the two conditions).

Functional categorization of Alexandrium genes at the 6-h and 96-h time points

Of the 14,882 contigs corresponding to differentially expressed genes, a putative function was identified for 9,680 of them, among which 7,121 were classified into functional categories according to KOGs (Fig. 4.4.3). At the early infection stage (6 h), only a few or no up or downregulated genes were detected in the parasite-infected samples. Only 28 of the 123 upregulated genes could be classified into KOG categories. Of these most were found in the categories 'translation, ribosomal structure and biogenesis (32%)', 'inorganic ion transport and metabolism (11%)', and 'general function prediction (14%)' categories.



Cellular processes and signaling

[M] Cell wall/membrane/envelope biogenesis
[N] Cell motility
[O] Posttranslational modification, protein turnover, chaperones
[T] Signal transduction mechanisms
[U] Intracellular trafficking, secretion, and vesicular transport
[V] Defense mechanisms
[W] Extracellular structures
[Y] Nuclear structure
[Z] Cytoskeleton

Metabolism

[C] Energy production and conversion
[D] Cell cycle control, cell division, chromosome partitioning
[E] Amino acid transport and metabolism
[F] Nucleotide transport and metabolism
[G] Carbohydrate transport and metabolism
[H] Coenzyme transport and metabolism
[I] Lipid transport and metabolism
[P] Inorganic ion transport and metabolism
[Q] Secondary metabolites biosynthesis, transport and catabolism

Information storage and processing

[A] RNA processing and modification
[B] Chromatin structure and dynamics
[J] Translation, ribosomal structure and biogenesis
[K] Transcription
[L] Replication, recombination and repair
[R] General function prediction
[S] Function unknown

Fig. 4.4.3: KOG category distributions of *Alexandrium* gene sequences that were uniquely upregulated (↑) or downregulated (↓) in the parasite-infected (+P) or waterborne cue-treated (+WP) or host-lysed samples (+A) compared with the corresponding control samples at the 6-h and 96-h time points.

At the late infection stage (96 h), treatment of *Alexandrium* with parasites or parasitic waterborne cues resulted in the upregulation of a large number of genes (Fig. 4.4.3). The significantly enriched transcripts (those with *P*-values < 0.05) were sorted from the hypergeometric distribution (the complete list is provided in Table S4.4 in the supplemental material). The majority of the enriched and upregulated genes were assigned to the following KOG categories: 'post-translational modification, protein turnover, chaperones' [parasite-infected only: 10%, $P < 0.05$; parasite-infected and parasite waterborne cue-treated (common genes): 17%, $P < 0.01$; parasite waterborne cue-treated only: 14%, $P < 0.05$], 'translation, ribosomal structure and biogenesis' [parasite-infected only: 7%, $P < 0.01$; parasite-infected and parasite waterborne cue-treated (common genes): 23%, $P < 0.01$; parasite waterborne cue-treated only: 11%, $P < 0.01$], and 'general function prediction [parasite-infected only: 20%, $P < 0.01$; parasite-infected and parasite waterborne cue-treated (common genes): 14%, $P < 0.05$; parasite waterborne cue-treated only: 18%, $P < 0.01$] (Table S4.4.4 in the supplemental material).

Transcripts that were upregulated in response to parasite infection only were enriched in the 'signal transduction mechanisms' (8%, $P < 0.01$), 'cell cycle control, cell division, chromosome partitioning cell' (6%, $P < 0.01$), and 'transcription' (6%, $P < 0.05$) categories. The genes that were upregulated in response to treatment with threatening parasitic waterborne cues only were mainly included in the 'amino acid transport and metabolism' (5%, $P < 0.01$) and 'lipid transport and metabolism' (5%, $P < 0.01$) categories. Genes that were commonly up-regulated by parasite infection and waterborne cues at the 96-h time point were enriched in the 'energy production and conversion' (9%, $P < 0.01$) and 'replication, recombination and repair' (3%, $P < 0.01$) categories.

I observed a marked downregulation of transcription in response to parasite infection alone at the 96-h time point (5,701 transcripts). Most of these genes were included in the 'translation, ribosomal structure and biogenesis' (29%, $P < 0.01$), 'replication, recombination and repair' (11%, $P < 0.01$), 'nucleotide transport and metabolism' (9%, $P < 0.01$), 'post-translational modification, protein turnover, chaperones' (6%, $P < 0.01$), and 'general function prediction' (12%, $P < 0.01$) categories (the full list is provided in Table S4.4.4 in the supplemental material).

Genes of particular interest

The numbers and functional annotations of the genes that were differentially expressed at the 6- and 96-h time points are compared in Table 4.4.1. Among the functional categories identified, I examined transcriptional changes in genes associated with metabolic pathways for energy production, photosynthesis, signal transduction, reactive oxygen species (ROS) and defense mechanisms. A full list of the regulated genes, with annotations, is provided in Appendix S4.4.1 in the supplemental material.

Among the 1,855 genes that were commonly upregulated at the 96-h time point in the +P and +WP treatments, I observed a significant enrichment ($P < 0.01$) of those involved in energy supply. More specifically, these genes were involved in glycolysis, fatty acid β-oxidation, the tricarboxylic acid cycle, oxidative phosphorylation, and the glyoxylate cycle (Table 4.4.1). Consistent with the observation that the breakdown of fatty acids (e.g. their β-oxidation) is enhanced following parasitic infection, the genes involved in fatty acid biosynthesis (such as those encoding the malonyl-acyl carrier protein and S-malonyltransferase) were downregulated in *Alexandrium*. Notably, core components of the photosystems, including six subunits of the photosystem II complex and photosystem I P700 chlorophyll *a* apoprotein A, were induced in response to parasite infection alone (Table 4.4.1).

In *Alexandrium* infected with the parasite or exposed to the waterborne cues, I observed differential regulation of signal transduction-related genes ('signal transduction mechanisms' category; $P < 0.01$), such as serine/threonine kinases and genes involved in calcium, mitogen-activated protein kinase and Ras signalling, as well as secondary metabolism-related genes. The differentially expressed serine/threonine kinases are highlighted in Table 4.4.1 (the full list is provided in Table S4.4.5 in the supplemental material), along with further potential defense-related ROS scavenging enzymes (the full list is provided in Table S4.4.6 in the supplemental material).

Table 4.4.1: Gene content and annotation summaries. The numbers of differentially expressed genes with predicted functions in *Alexandrium* at 6 h and 96 h after infection with *Amoebophrya* or treatment with parasitic waterborne cues or lysed host cells.

		6 h						96 h							
		+P		+WP		+A		+P		Overlap of +P and +WP		+WP		+A	
		up	down	up	down	up	down	up	down	up	down	up	down	up	down
Significantly differentially expressed transcripts (adj$p$<0.05)		4	3	123	4	40	83	3216	5701	1855	67	3720	215	2	0
Annotated transcripts		0	1	73	2	27	26	1796	3267	1566	56	2935	123	2	0
Energy production	ko00010 Glycolysis / Gluconeogenesis	0	0	0	0	0	0	5	10	13	0	18	2	0	0
	ko00620 Pyruvate metabolism	0	0	0	0	0	0	5	8	7	0	14	0	0	0
	ko01200 Carbon metabolism	0	0	0	0	0	0	12	24	28	0	42	2	0	0
	ko01230 Biosynthesis of amino acids	0	0	0	0	0	0	11	23	17	0	28	4	1	0
	ko01212 Fatty acid metabolism	0	0	0	0	0	0	4	11	9	0	18	0	0	0
	ko00020 Citrate cycle (TCA cycle)	0	0	0	0	0	0	3	7	13	0	15	0	0	0
	ko00190 Oxidative phosphorylation	0	0	2	1	0	0	10	13	29	0	23	0	0	0
	ko00630 Glyoxylate and dicarboxylate metabolism	0	0	0	0	0	0	4	5	7	0	9	0	0	0
Photosynthesis	ko00195 Photosynthesis	0	0	0	0	0	0	9	3	0	0	0	1	0	0
	ko00710 Carbon fixation in photosynthetic organisms	0	0	0	1	0	0	3	5	9	0	10	2	0	0
	ko00860 Porphyrin and chlorophyll metabolism	0	0	1	0	0	0	3	5	0	0	2	0	0	0
	ko00030 Pentose phosphate pathway	0	0	0	0	0	0	2	5	6	0	8	1	0	0
Signal transduction	ko00260 Glycine, serine and threonine metabolism	0	0	0	0	0	0	5	6	7	0	13	1	0	0
	ko04020 Calcium signaling pathway	0	0	1	0	0	0	10	6	8	0	11	1	0	0
	ko04010 MAPK signaling pathway	0	0	1	0	0	0	13	14	9	0	13	0	0	0
	ko04014 Ras signaling pathway	0	0	0	0	0	0	6	6	9	0	7	1	0	0
	ko01110 Biosynthesis of secondary metabolites	0	0	2	0	0	0	28	52	52	0	96	3	1	0
	PF00069 Pkinases domain (Serine/threonine protein kinase)	0	0	1	0	0	0	51	46	24	0	39	2	0	0
	ROS scavenging enzymes	0	0	0	0	0	0	7	17	17	0	37	1	0	0
Cell cycle	ko04110 Cell cycle	0	0	0	0	0	0	9	23	10	1	18	2	0	0

MAPK, mitogen-activated protein kinase; TCA, tricarboxylic acid.

PSP toxin content of Alexandrium

At both the 6- and 96-h time points, the PSP toxin content (fmol $cell^{-1}$) of *Alexandrium* cells in the +P and +WP treatments was slightly lower than that of the corresponding control cultures (Fig. 4.4.4); however, the differences were not statistically significant (Students *t*-test, $P > 0.05$). The control cultures did not show a significant change in PSP toxin content over time.

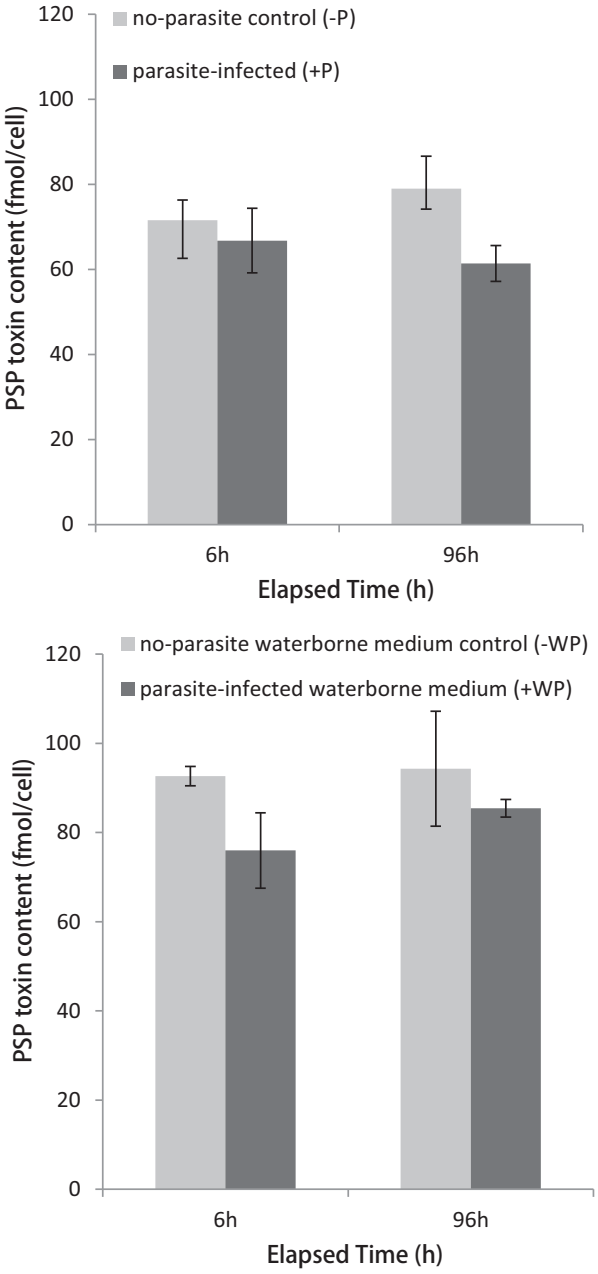


Fig. 4.4.4: PSP toxin content of *Alexandrium* in the parasite-infected (+P), parasite waterborne medium-treated (+WP), and corresponding control groups (-P and -WP) at the 6 h and 96 h time points. No significant differences were detected (ANOVA).

4.4.5 Discussion

The results presented herein demonstrate that the dinoflagellates, *Alexandrium,* can mount a strong response to the parasite, *Amoebophrya*, as revealed by changes in gene expression. At the early stage (6-h time point), neither parasite infection nor exposure to parasitic waterborne cues caused a marked regulation of genes (Fig. 4.4.2). It is likely that the level of threat from the parasite, in terms of time or intensity, was too low to trigger the main response of the host at the 6-h time point. Infection with parasites resulted in a massive downregulation of *Alexandrium* genes (5,701) after 96 h, whereas exposure to waterborne cues did not trigger a similar pattern of downregulation, suggesting that a direct interaction between the parasite and its host is necessary for this phenomenon. The observed downregulation of genes could be due to an ongoing degradation of the host cell structures; however, the simultaneous upregulation of more than 3,000 genes in the parasite-infected samples argues against this scenario. On the other hand, exposure to parasitic waterborne cues also elicited the upregulation of an approximately equal number of transcripts as direct parasite infection, indicating that the host is capable of sensing its parasite. The remarkable overlap of upregulated genes between these two conditions indicates that the stimulation of signal transduction chains by waterborne cues alone could prime the host's defense or induce host's adaptive responses to the parasite activity. While the observed downregulation of genes could be the consequence of host decay, the induction of genes is not likely a side effect of such and thus represents an active response to the parasites attack. The apicomplexan parasites, *Toxoplasma* and *Plasmodium* included, provide efficient infection strategies to subvert host cell processes, avoid clearance by the defense mechanisms and modulate the metabolic pathways of the host (Plattner & Soldati-Favre 2008). Therefore, this parasite-driven activity could result in expression changes related to host adaptive responses to parasite-initiated effects. I summarized my major findings in Fig. 4.4.5 and describe the differentially expressed genes involved in energy production, signal transduction, and defense mechanisms in the following sections.

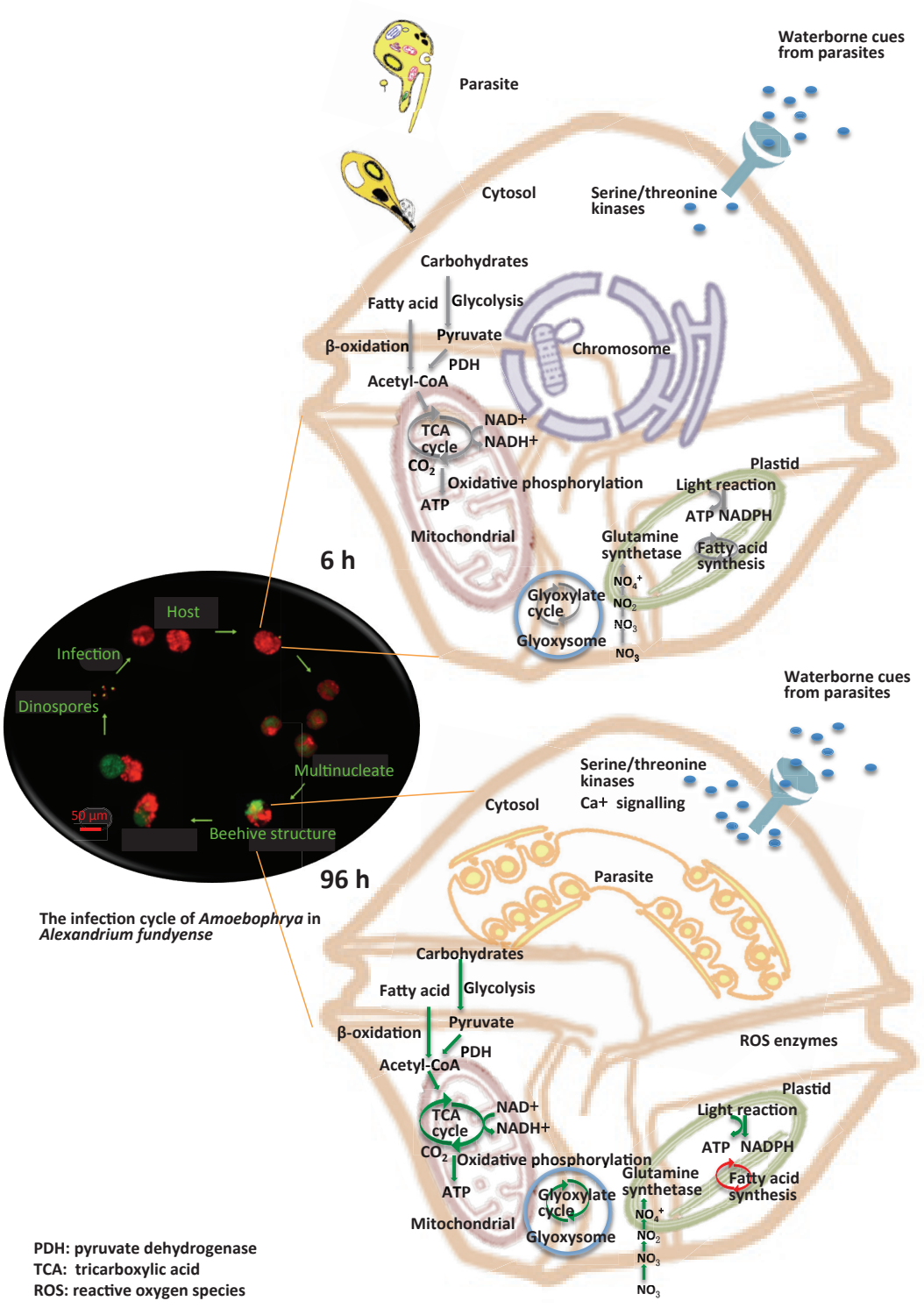


Fig. 4.4.5: Altered cellular activity in *Alexandrium fundyense* infected by the parasite *Amoebophrya*. The presence of the parasite and waterborne cues from the parasite results in up-regulation (green arrows) of energy production processes involved in photosynthesis, ATP synthesis through glycolysis and fatty acid production, and down-regulation (red arrows) of fatty acid synthesis. At the same time, the activation of calcium-mediated signal transduction and enzymes correlated with ROS production indicates that *Alexandrium* could perceive chemical cues from the parasite and induce defense mechanisms.

Functional categorization of processes in Alexandrium

Energy production

The increased expression of pathway components associated with energy production at the 96-h time point after infection of *Alexandrium* with the parasite or exposure to waterborne cues indicates an increased energy demand for defense in the host. A KEGG pathway analysis depicted a complete oxidation of carbohydrates and fatty acids by the tricarboxylic acid cycle in the mitochondria of *Alexandrium*, as well as ATP generation, and may reflect the costs necessary for defense, as observed in land plants (Livaja *et al.* 2008). Alternatively, the high demand of energy might also be induced by the parasite to get resources for its intracellular development and reproduction, in agreement with studies in *Toxoplasma gondii* (Danne *et al.* 2013; Tenter *et al.* 2000). And one could speculate that the ATP generation observed here might be induced by a rapid ATP depletion in host cells after the parasite infection. However, photosynthetic activity, which is increased after parasite infection but is not inducible by waterborne cues, might hint at the requirement for enhanced energy needs and oxygen production from light in defending the host cells.

Genes related to fatty acid biosynthesis, such as the malonyl-acyl carrier protein, were downregulated in parasite-infected cells at the 96-h time point. Acyl carrier protein is the core protein involved in fatty acid synthesis (Mazumdar & Striepen 2007). This results indicate that fatty acid synthesis in the plastid was inhibited, whereas β-oxidation of fatty acids was induced, either reflecting the increased energy requirement of the infected host or the response induced by the parasite to fulfil itself. This finding is consistent with that reported for the response of the coral *Acropora cervicornis* to pathogen infection (Libro *et al.* 2013), as well as those of other organisms affected by grazing (Flöthe *et al.* 2014; Wohlrab *et al.* 2010). In addition, the regulation of fatty acid coding genes, as major components of cell membranes, may be an indicator of cell growth machinery. I identified the downregulation of genes involved in cell cycle in response to parasite infection at 96 h (Table 4.4.1), which implies that parasite infection could inhibit host cell division process. By contrast, cell growth-related genes were observed to be upregulated at 96h when infected by parasite or exposed to waterborne cues. This phenomenon is likely to be induced by the parasite in order to make use of this resource for its own reproduction and growth, because simultaneously cell divison-

and proliferation-related genes were highly expressed in the parasite *Amoebohrya* at the time of fast cell growth and nuclear divison (Lu *et al.* 2014).

Signal transduction and defense mechanisms

Calcium signalling and the activities of several protein kinases seem to be important for dinoflagellates as direct responses to the parasite and may also act to prime the cell towards a parasite attack. Calcium and calmodulin, which were differentially regulated in parasite-infected *Alexandrium,* are key components of signal transduction pathways and are involved in stress responses of various environmental stress conditions (Scandalios 2005) but also in cell cycle control in plants and marine phytoplankton (Jingwen *et al.* 2006). The second messenger cAMP regulates cell cycle progression in the dinoflagellates, *Crypthecodinium cohnii* (Lam *et al.* 2001; Wurzinger *et al.* 2011); although it may be important for the dinoflagellate host-parasite interaction, little is currently known about the significance of calcium signalling in photosynthetic dinoflagellates compared with heterotrophic dinoflagellates or parasitic apicomplexan (Plattner *et al.* 2012; Verret *et al.* 2010). A recent study showed that the expression levels of calcium-dependent protein kinases and serine/threonine kinases are altered as a defensive response to copepod grazers (Wohlrab *et al.* 2010).

Exposure of *Alexandrium* to parasites or parasitic waterborne cues induced the upregulation of genes involved in the production of ROS (Table 4.4.1 and Table S4.4.6 in the supplemental material), pointing towards ROS production as a key defense mechanism or a response to ROS spreading in the cytosol due to internal membrane damage. ROS, such as oxygen ions and peroxides (H_2O_2), are a by-product of cell metabolism; however, their over-production in marine organisms can cause oxidative damage and irreversibly alter DNA, proteins, and lipids (Halliwell & Gutteridge 2015; Lesser 2006). A recent study showed that a high production of ROS by *Alexandrium catenella* under stress conditions may play the key role of fish gill damage in Chilean fjords (Mardones *et al.* 2015). As such, ROS can have direct negative effects on intracellular parasites. ROS production occurs in *Alexandrium* during temperature increases (Jauzein & Erdner 2013), and has been observed in diverse marine organisms exposed to environmental abiotic stressors such as UV or heat shock (Lesser 2012); hence, it is considered an indicator of stress-related pathways. Flores *et al.* (2012) found that the addition of ROS enzymes increased the survival of both the ciliate

Tiarina fusus and the heterotrophic dinoflagellate *Polykrikos kofoidii* after exposure to *Alexandrium*, suggesting that ROS may be indirectly correlated to the toxicity of *Alexandrium* to protists. Here, several ROS enzymes, including superoxide dismutase, peroxidase, and catalase, were upregulated in response to parasite infection or exposure to parasitic waterborne cues. Superoxide dismutase catalyses the conversion of the superoxide radical (O_2^-) to peroxides (H_2O_2), whereas peroxidase and catalase convert H_2O_2 to water (Apel & Hirt 2004). Photosynthesis is an additional source of ROS in plants (Foyer & Shigeoka 2011). My finding that photosynthesis-related genes were up-regulated after direct contact with the parasite, but were not induced by exposure to parasitic waterborne cues, suggests that ROS play an important role in the defense mechanism.

PSP toxin distribution

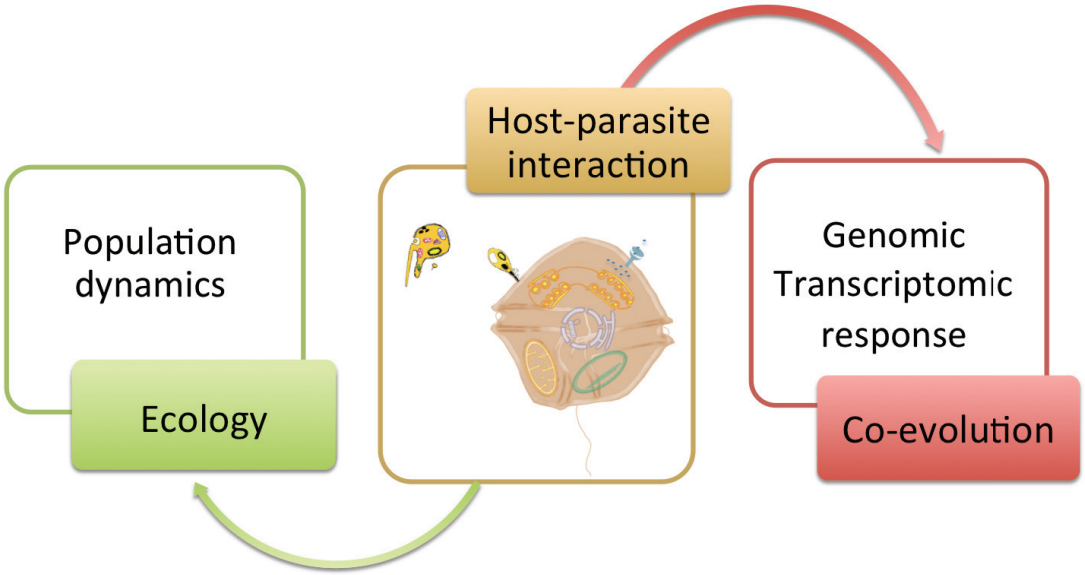
The absence of a significant change in the PSP toxin content of *Alexandrium* after parasite infection or exposure to parasitic waterborne cues (Fig. 4.4.4) indicates that there is no effect of the parasite on PSP toxin production in this species. To date, PSP toxins have been shown to act as potential defense compounds against metazoan (copepod) grazers (Selander *et al.* 2006; Wohlrab *et al.* 2010; Yang *et al.* 2010), but do not act as defense compounds against unicellular heterotrophs (Tillmann & Hansen 2009; Tillmann & John 2002). The phylogeny and mode of action ('intracellular grazer') of *Amoebophrya* more closely resemble those of unicellular protistan grazers than metazoan copepods; therefore, it is likely that PSP toxins do not serve as defense compounds towards protistan parasites in *Alexandrium*. Bai *et al.* (2007) performed similar experiments using *Amoebophrya* sp. and a toxic strain of *Karlodinium veneficum* and found that the parasite did not actively catabolize the host's toxins. However, little is currently known about the impact of parasitism on *Alexandrium* toxin production, and this study investigated the effect on only one strain of *Alexandrium.* Repeated exposure of different *Alexandrium* strains and analysis of the effects of PSP toxins during parasite infection remain to be explored. In addition, I observed the up-regulation of genes involved in secondary metabolism in response to parasite infection or parasitic waterborne cues (Table 4.4.1). Aside from the known PSP toxins and the unknown allelochemicals (Ma *et al.* 2009; Tillmann & John 2002), these secondary metabolites may also be involved in the defense against parasites during the infection cycle.

4.4.6 Conclusions

The results presented herein reveal that the dinoflagellate *Alexandrium fundyense* undergoes specific alterations in gene expression in response to infection or exposure to waterborne cues from the parasite *Amoebophrya.* A large number of genes were downregulated, mainly due to parasite infection and damage to host cell structures. By contrast, the upregulation of genes in cells treated with parasitic waterborne cues affected the host's defense mechanisms, in particular energy production involved in photosynthesis, ATP synthesis through glycolysis and fatty acid production, calcium-mediated signal transduction, and ROS production. Taken together, these data suggest that dinoflagellate parasite infection and chemical cues from the parasites can trigger a powerful defense response in dinoflagellate hosts. It would be beneficial to have further protein biochemistry or proteomics analysis in order to investigate the changes on the transcriptome level measured in this study reflected in protein mechanisms.

5. Synthesis

This thesis work has explored the genetic characteristics of the parasite *Amoebophrya* sp., as well as the responses of its host *Alexandrium fundyense* towards the parasite infection. Data were presented at the cellular, genomic and transcriptomic level. In this synthesis, the major discussion is elaborated to provide novel views on host-parasite relationships, conceptualised as an ecological and co-evolutionary interdependency within a complex adaptive system.



The following questions were addressed in this thesis to further complete our understanding of the host-parasite interactions:

- Are there variations in susceptibility or adaptations of different *Alexandrium* geographical populations towards parasite infection?
- If PSP toxins produced by *Alexandrium* are acting as a chemical defensive mechanism against infectivity and if so how are they regulated?
- What are the genomic characteristics of the basal dinoflagellate parasite *Amoebophrya* and the parasite's survival strategies?
- Which processes and genes are regulated in the parasite over the time course of the infection?
- How is the induced defense/stress mechanisms expressed in the host *Alexandrium* at a transcriptomic level?
- Do waterborne cues from the parasite trigger similar responses of the host?

5.1 Effects of parasite infection on host's population dynamics

The studies described in **Chapter I** provide the prevalence of the *Amoebophrya* sp. strains to different populations of *Alexandrium fundyense* from three very distant geographic origins. The *Alexandrium* populations, established with about 30 strains were isolated from the North Sea, the Gulf of Maine and Alaska. Two clones of parasites were from the Gulf of Maine area. A strong negative effect of parasitism on the host shows that the parasite has great potential in controlling blooms of *Alexandrium*. No significant differences in growth rate and lytic rate of *Alexandrium* were observed among the three populations, implying that there was no adaptation of the host towards parasite infection on an intra-specific level. Furthermore, no difference in the parasite prevalence (the percentage of infected cells in the host population) of each population could be identified, but the infection percentages were highly variable on the intra-population level (2%-36%).

Alex5 (RCC3037, isolated in North Sea, 2009) was chosen as the host for further studies in this thesis (**Chapter II – IV)**, since its infection percentage by the parasite *Amoebophrya* was among the highest. Furthermore, its genetic and genomic background is well studied (Alpermann *et al.* 2009; Wohlrab 2013) and it is not known to produce allelochemicals that could affect the host-parasite infection (Tillmann *et al.* 2009).

The results in this thesis show that neither PSP toxin content nor toxin profile changed in different populations (**Chapter I**), over time after parasite infection or exposure to parasitic waterborne cues (**Chapter IV**). Intracellular toxin contents of dinoflagellates are known to be affected by a variety of abiotic factors (e.g. temperature, salinity, light, and nutrients) (John & Flynn 2000; Ogata *et al.* 1987; White 1978). Production of PSP toxins has been described to act as a defensive response to grazing by predators or competitors (Selander *et al.* 2006; Wohlrab *et al.* 2010), but not as a defensive strategy against unicellular heterotrophs (Tillmann & Hansen 2009; Tillmann & John 2002). I hypothesized that parasitism by *Amoebophrya* species appears to represent another biological factor that influences dinoflagellate toxin content, but the results indicated that it is likely there is no apparent direct effect of the parasite on PSP toxin production, and hence PSP toxins do not serve as defense compounds towards parasites in *Alexandrium*.

Bai *et al.* (2007) described a positive correlation to karlotoxin concentrations in *Karlodinium veneficum* infected by *Amoebophrya*, but they concluded that the hosts' toxin content may be coupled to other factors, such as an increased cell size. Moreover, such as correlation does not necessarily reflect a direct connection. A recent study of *Alexandrium fundyense* infected by *Amoebophrya* displayed similar results to my study in that PSP toxin contents did not vary significantly over the infection cycle (Kim & Park 2016). In conclusion, the presented findings in this thesis tend to show that PSP toxins are not a likely way for hosts to defend themselves against parasitoids, although further testing with more strains is required.

This study (**Chapter I & IV**) yielded the following insights:

- Parasite pressure leads to a negative impact on host population dynamics;
- No adaptation of host was displayed at the intra-specific level, but instead showed wide variation on an inter-population level;
- PSP toxins do not seems to serve as defense compounds against parasitism in *Alexandrium*.

5.2 Genomic features of parasitism

The sequenced genome of the parasite *Amoebophrya* in size around 90 Mbp (**Chapter II**), which is smaller compared to that of the coral symbionts, *Symbiodinium minutum* (616 Mbp assembly of the 1,500 Mbp genome) (Shoguchi *et al.* 2013) and *S. kawagutii* (935 Mbp assembly of the 1,180 Mbp genome) (Lin *et al.* 2015). The relatively small genome size in the syndinean dinoflagellate *Amoebophrya* may be due to the parasitic life style with consequent dependence on the host metabolism and resources. The genome analysis (**Chapter II**) provides insights into the parasitic features of the organism and some characteristics of the Alveolata/dinoflagellate genome evolution. Furthermore, since the genome sequences of the parasite were available (**Chapter II**), it was possible to differentiate genes from the parasite versus the host during the parasite infection cycle to be able to study the mechanisms and processes of parasite infection and host responses at the molecular level (**Chapter III & IV**).

The first significant genomic feature of *Amoebophrya* is the lack of a mitochondrial genome (**Chapter II**) despite the presence of this organelle at certain life stages, Miller *et al.* (2012) observed that the parasite *Amoebophrya* from the dinoflagellate *Akashiwo sanguinea* had two mitochondria in dinospores. The mitochondria were present but inconspicuous from 12 to 36 hours and prolonged as mature trophonts by 48 hours post-infection (Miller *et al.* 2012). Dinoflagellates inherited a drastically reduced mitochondrial genome, and is likely the most gene-impoverished of any aerobic eukaryote, containing only three protein-coding genes (cob, cox1, and cox3), two highly fragmented rRNAs, and no tRNAs (reviewed by Wallerand Jackson, 2009). The search for mitochondrial encoded genes (coxI and cytb) or mitochondrial rRNA yielded no match in the *Amoebophrya* genome from the analysis in **Chapter II**. This appears to be the first aerobic eukaryote species that is known to have an aerobic life style and mitochondria, but no mitochondrial genome.

In addition, the predicted proteome of *Amoebophrya* was further searched for five oxidative phosphorylation complexes forming the respiratory chain in mitochondria. The results in **Chapter II** demonstrate that oxidative phosphorylation complexes I and III were lost, and this is consistent with previous studies, which showed the same mitochondrial condition, e.g., lack of complexes I and III in *Chromera velia* an aerobic phototrophic relative of Apicomplexa (Flegontov *et al.* 2015).

Another feature of the *Amoebophrya* genome is the secondary loss of the photosynthetic plastid (**Chapter II**). There is evidence that the ancestor of parasitic dinoflagellates and apicomplexa possessed a photosynthetic plastid and became secondarily non-photosynthetic (Janouškovec *et al.* 2010). Most heterotrophic dinoflagellates retain reduced forms of a plastid organelle and even the parasitic apicomplexans contain a relict non-photosynthetic plastid called the apicoplast (reviewed by Keeling, indicating the essential metabolic functions of plastids despite photosynthesis. In contrast, no plastid organelle has been reported in an ultrastructural study of *Amoebophrya* species (Miller *et al.* 2012). To data, the apicomplexan *Cryptosporidium*, a protozoan pathogen that causes acute gastroenteritis and diarrhoea worldwide, is the only known case of plastid loss (Xu *et al.* 2004; Zhu *et al.* 2000). In *Amoebophrya*, the few remaining plastidial genes (e.g. porphobilinogen synthase [HemB], porphobilinogen deaminase [HemC], and uroporphyrinogen III

decarboxylase [HemE]) are encoded independently by the plastid and reallocated to the cytosol, and some genes (e.g. uroporphyrinogen III synthase [HemD] and coproporphyrinogen oxidase [HemF]) are clearly plastid derived (**Chapter II**). This observation, in concert with a new research reporting that the dinoflagellate *Hematodinium*, a marine parasite of crustaceans, also had a complete secondary loss of the plastid organelle by retention of cytosolic pathways for synthesis of fatty acids and tetrapyrroles, and retained few genes of plastid origin (Gornik *et al.* 2015).

An inability to synthesize several amino acids was also observed in the genome as one of typical parasitic features. A novel type of metabolic regulation observed in *Amoebophrya* was that tryptophan synthesis is physically interlocked at the genetic level with the shikimate pathway (**Chapter II**). The shikimate pathway is a seven step metabolic route and produces precursors for the biosynthesis of aromatic amino acids (including phenylalanine, tyrosine, and tryptophan), as well as secondary metabolites to which approximately 20% of photosynthetically fixed carbon is directed in vascular plants (Herrmann 1995). In *Amoebophrya*, the seventh building block, aroG, was inserted into the tryptophan synthesis gene (**Chapter II**). Furthermore, the shikimate pathway appeared to be required for the synthesis of tryptophan only, since the absent of the synthesis of phenylalanine and tyrosine in *Amoebophrya* (**Chapter II**).

The major genomic features of the parasite *Amoebophrya* (**Chapter II**) were as follows:

- No mitochondrial genome;
- Secondary loss of the photosynthetic plastid organelle;
- Regulation of synthesis of tryptophan (the only aromatic amino acid that can be synthesized by *Amoebophrya*) is achieved via a unique shikimate pathway.

5.3 Transcriptomic responses of parasitism

Parasite infection has been well studied on morphological and physiological levels, and yet the molecular processes of infection are poorly understood. One aim of this study was to investigate cDNA libraries from the parasite *Amoebophrya* during infection (**Chapter III)**. Three cDNA libraries were generated at different life stages: the pure dinospore stage (0 h), the initial infection/penetration stage (6 h), and the maturation stage (96 h). Since the genome sequences of the parasite were available (**Chapter II**), it was possible to determine the presence of genes from the parasite in all libraries. By analyzing the expressed sequence tags (ESTs) obtained from different life history stages, results from **Chapter III** provide insights into the infection mechanisms of the parasite *Amoebophrya* in the host *Alexandrium*, and identify processes and genes that may be relevant to the transition of free-living organisms to parasites.

In general, the results in **Chapter III** demonstrate that the host-parasite interaction causes significant changes in parasite gene expression over the time course of the infection. In particular, glycan-related enzymes and carbohydrate-binding proteins (lectins) were present in *Amoebophrya* at the beginning of the infection that most likely contributes to the attachment of the parasite to the host cell. The function of lectins has been evolved in prey recognition and discrimination prior to phagocytosis in heterotrophic dinoflagellate lineages (Wootton *et al.* 2007), therefore lectins might serve as an important cellular surface recognition mechanism in host-parasite interactions (Joseph *et al.* 2010; Martel 2009; Roberts *et al.* 2006).

The majority of the genes over-represented in the parasite during the infection were related to cytoskeletal organization, signal transduction and stress responses, indicating the processes that enable recognition, adhesion, and penetration of the host, as well as the responses towards the host defense (**Chapter III**). First, the expression of several heat shock proteins (HSP70, HSP90, DnaJ, Cpn10) and cold shock proteins increased in *Amoebophrya* during the infection. Heat shock proteins are highly conserved proteins and highly expressed in the active to latent form in the parasitic apicomplexa *Toxoplasma gondii*, in *Perkinsus marinus*, a protozoan parasite of the oyster, and in the parasitic ciliate *Cryptocaryon irritans* to overcome the stress from the host environment (Joseph *et al.* 2010; Lokanathan *et al.* 2010;

Weiss *et al.* 1998). Cold shock proteins are associated with posttranscriptional regulation in eukaryotes and have been found to play an important role in regulating translation in the dinoflagellate *Lingulodinium* (Beauchemin *et al.* 2012; Mihailovich *et al.* 2010). These results (**Chapter III**) indicate that the stress related gene products in *Amoebophrya* might be needed to overcome the response of host defense.

Second, a wealth of genes (7,388) related to cell division reconstruction and proliferation are highly expressed in *Amoebophrya* at late infection and maturation stage (**Chapter III**). With the rapid growth of the parasite in the host cell, the host produces reactive oxygen species (ROS) (**Chapter IV**) with a toxic effect on the pathogen (Miura *et al.* 2012). ROS, such as oxygen ions and peroxides (H_2O_2), are a by-product of cell metabolism; however, their overproduction in marine organisms can cause oxidative damage and irreversibly alter DNA, proteins and lipids (Lesser 2012). The parasite *Amoebophrya* experiences this oxidative burst reaction of the host and might use superoxide dismutases (SOD) to protect itself from reactive oxygen intermediates (ROIs) generated by the host's oxidative enzymes. Similar scenario was illustrated in protozoan parasite *Perkinsus marinus* (Joseph *et al.* 2010) and the expression of SODs related genes were found in *Amoebophrya* during infection (**Chapter III**). Taken together, the parasites may evolve to enhance their infectivity under the pressure from the host environment.

The results in **Chapter III** indicate that host-parasite interaction causes significant changes in parasite gene expression over the time course of the infection.

- Particular surface lectins are expressed at the beginning of the infection cycle, which likely mediate the attachment to the host cell;
- In a later phase, signal transduction-related genes together with transmembrane transport and cytoskeleton proteins point to a high integration of processes involved in host recognition, adhesion, and invasion;
- At the final maturation stage, cell division- and proliferation-related genes were highly expressed, reflecting the fast cell growth and nuclear division of the parasitoid.

5.4 Effects of parasite infection on *Alexandrium*

Protist parasites and their hosts exert selection pressure on each other, as a result of a continuous conflict between the divergent interests of each partner. Such kind of long-term adaptations may lead to innovative traits for parasites to infect hosts (e.g., a high integration of processes involved in host recognition and a fast cell growth and nuclear division after infection described in the results in **Chapter III**), and hosts to resist infection by opposite processes (**Chapter IV**). Whereas most molecular studies searching gene expression changes were concerned with one part of either the parasite or the host, to my knowledge there are few studies available on the molecular processes underlying parasite infection and also parasite-induced defense of the host.

The results of **Chapter IV** identified many genes involved in the host's induced defense and/or feedback responses towards parasite infection. Compared with control samples, a total of 14,882 *Alexandrium* genes were differentially expressed over the whole parasite infection cycle at three time points (0, 6, and 96 h), the same time points as reported in **Chapter III**. The differential expression of certain genes during parasite infection suggests their involvement in the interaction between the host and the parasite.

The significantly higher expression of pathway components associated with energy production after infection of *Alexandrium* with the parasite or exposure to waterborne cues from the parasite indicates an increased energy demand for defense in the host (**Chapter IV**). On the one hand, A KEGG pathway analysis investigated a complete oxidation of carbohydrates and fatty acids by the tricarboxylic acid cycle in the mitochondria of *Alexandrium*, as well as ATP generation, and may reflect necessary costs for defense. On the other hand, the high demand of energy might also be induced by the parasite to get resources for its intracellular development and reproduction, as observed in the apicomplexan parasite *Toxoplasma gondii* (Danne *et al.* 2013).

Signal transduction, for instance, calcium signalling and the activities of several protein kinases, seem to be important for *Alexandrium* as direct responses to the parasite and may also act to prime the cell towards a parasite attack (**Chapter IV**). Calcium and calmodulin are important components of signal transduction pathways and are involved in

stress responses to various environmental conditions, as well as in cell cycle control in plants and marine phytoplankton (Jingwen *et al.* 2006; Scandalios 2005). A recent study showed that the expression levels of calcium-dependent protein kinases and serine/threonine kinases are changed as a defensive response to copepod grazers (Wohlrab *et al.* 2010).

One objective of this study was to examine if the waterborne cues from the parasite could prime the host response and to discriminate these responses from those elicited by the parasite infection. The results (**Chapter IV**) demonstrate that the host *Alexandrium* mounts a strong response to the parasitic waterborne cues, as revealed by the upregulation of an equal large number (~3,000) of transcripts as direct parasite infection. The remarkable overlap of upregulated genes between these two conditions (parasite infection and waterborne cues from the parasite) indicates that the stimulation of signal transduction chains by waterborne cues alone could prime the host's defense or induce host's adaptive responses to the parasite activity.

Co-evolution therefore seems to have driven the ability of the host *Alexandrium* to recognize parasites based on their waterborne cues. This reciprocal selection may lead to continuous changes of both parasite infectivity and host resistance, thus cause cyclic changes through negative frequency-dependent selection, also called *Red Queen* dynamics. Evidence for co-evolutionary dynamics from plankton populations is very limited. The exact nature of the underlying dynamics of host-parasite co-evolution is yet under debate and may be determined by recurrent selective sweeps (i.e., arms race dynamics) or winnerless coevolution (i.e., *Red Queen* dynamics). A recent cross-infection study revealed a high potential for *Red Queen* dynamics between the dinoflagellate host *Alexandrium minutum* and parasite *Parvilucifera.*

RNA sequencing analyses in **Chapter IV** indicated that exposure to the parasite or parasitic waterborne cues produced significant changes in the expression levels of *Alexandrium* genes associated with specific metabolic pathways.

- At the early stage, neither parasite infection nor exposure to parasitic waterborne cues triggers a marked regulation of genes;

- At the final stage, infection with parasites resulted in a massive down-regulation of *Alexandrium* genes, mainly due to parasite infection and degradation of host cell structures;
- The up-regulation of genes in cells treated with parasitic waterborne cues affected the host's defense mechanisms, in particular energy production involved in photosynthesis, ATP synthesis through glycolysis and fatty acid production;
- Signal transduction chains by waterborne cues alone could prime the host's defenses or induce host's adaptive responses to the parasite activity.

5.5 Concluding remarks

The infection process of this host-parasite relationship with its co-evolutionary dynamics was considered in this thesis, together with the exploration of the genetic basis that highlights how the parasite *Amoebophrya* evolved to control toxic *Alexandrium* population dynamics. On the one hand, the parasites exert a strong negative effect on the population of the bloom-forming dinoflagellate at an ecological level, but a selection by the parasites alone is not strong enough to make host populations adapt. On the other hand, I hypothesized that the toxin production of the hosts *Alexandrium* may act as a defensive mechanism against parasite infectivity. In contrast, this seems not to be true since there is no effect of the parasite infection or exposure to parasitic waterborne cues on host's PSP toxin production.

The parasite genome (<100Mb) sequence revealed several novel features of parasitic dinoflagellates, such as a complete loss of the mitochondrial genome, an inability of certain amino acids generation, an unusual metabolic regulation as the physical link of the shikimate pathway and tryptophan synthesis in *Amoebophrya*. All these functional losses of the parasitic syndinean dinoflagellates, together with a reduction of genome size compared to core dinoflagellates, point out the parasite dependence on the host and highlight the phylogenetic placement of *Amoebophrya*. A transcriptomic dataset of the parasite *Amoebophrya* from the infection cycle of this parasite-host system shows that at early infection processes related to the attachment to the host cell, followed by integrative processes involved in host recognition, adhesion, and invasion. During maturation, cell division and proliferation related genes reflect fast cell growth of the parasite.

The host *Alexandrium* reacts differently towards parasite infection and respective parasite waterborne cues, but both treatments exhibited significant changes in gene expression associated with the signal transduction and metabolic pathways as energy production for its increased energy demand demonstrated in genes related to defense mechanisms. Co-evolution seems to have driven the ability of the host *Alexandrium* to recognize parasites based on their waterborne cues.

5.6 Future perspectives

There are interesting questions arising from the presented findings and a lot of additional work will be required for a complete understanding of the host-parasite interaction.

Although the studies presented in this thesis contribute to a better understanding of the host resistance and parasite infectivity on a population-wide distribution, several aspects concerning variations in susceptibility of different strains of the parasite *Amoebophrya* remain to be elucidated. I found this natural host-parasite infection in the North Sea during an expedition, but failed in cell isolation and translation in the lab. Therefore only two clones of the parasite *Amoebophrya* from Gulf of Maine area were used in my study. Clearly infection experiments with different parasite strains and preferably more host populations would be needed in order to better answer the question: Are there variations in susceptibility of different *Alexandrium* genotypes and/or in infectivity of *Amoebophrya* parasites?

In the marine environment, the recent increase of diseases caused by parasitic syndinean may have been facilitated by alterations of the environment. Therefore, future investigations on the influence of environmental conditions on host-parasite infection should be performed with different abiotic factors (e.g. different temperatures and pH values) to characterize whether host genotypes have changed, or would be expected to respond to climate change events, in their susceptibility under different environmental conditions.

A host's susceptibility to parasite infection may change over generations, as demonstrated in serial passage experiments (Little *et al.* 2006). The next questions would be: Whether or not susceptibility to infection and the genotypes of *Alexandrium* change over several generations? Whether or not phenotypic changes result in an altered genotype? The short generation time of the host-parasite system (~four days) enables us to conduct serial passage of the parasite over many generations in a relatively short time. The experiment could be designed as follows: expose three *Alexandrium* genotypes of different susceptibilities to a parasite *Amoebophrya* genotype and over multiple passages. For the last (20^{th}) passage, the parasite would be inoculated with the native *Alexandrium* genotype and an *Alexandrium* alternative genotype that has not been used in the serial passage experiment, but is of known

susceptibility to that parasite genotype. In the end, the virulence of parasite infection, the gene expression profile and the genotype changes would be tested to indicate whether phenotypes, genotypes and susceptibility to parasite infection changes are in concordance.

Amoebophrya was previously suggested as a biological control agent for harmful algal bloom (HAB) organisms due to its high host specificity and virulence (Taylor 1968). But this idea was questioned considering its uncertainties about host specificity and pathogen stability (Nishitani *et al.* 1985; Salomon & Imai 2006), as well as the limitation of parasites in regulating certain species of marine dinoflagellates (Coats & Bockstahler 1994; Salomon *et al.* 2003b). In my study, the negative effects of parasitism on host growth indicates that, in principal, *Amoebophrya* potentially could be used to control blooms of *Alexandrium,* but the effects on multi-clonal cultures and possible negative impacts to other members in the ecosystem are still unknown. Widely accepted theories based on abundant empirical evidence that release from natural enemies favors the success of organisms introduced into new territories (Salomon & Imai 2006; Torchin *et al.* 2003). Especially in an open system such as the oceans, application of living biological pest control organisms is much more dangerous and less predictable than biological control on land (Secord 2003). Therefore, the proposed use of parasites as a biological control of HABs is still an open and complex issue and certainly not ready for field trials.

References

Adams KL, Palmer JD (2003) Evolution of mitochondrial gene content: gene loss and transfer to the nucleus. *Molecular Phylogenetics and Evolution* **29**, 380-395.

Alpermann TJ, Beszteri B, John U, Tillmann U, Cembella AD (2009) Implications of life‐history transitions on the population genetic structure of the toxigenic marine dinoflagellate *Alexandrium tamarense*. *Molecular Ecology* **18**, 2122-2133.

Alpermann TJ, John UE, Medlin LK, *et al.* (2006) Six new microsatellite markers for the toxic marine dinoflagellate *Alexandrium tamarense*. *Molecular Ecology Notes* **6**, 1057-1059.

Alpermann TJ, Tillmann U, Beszteri B, Cembella AD, John U (2010) Phenotypic variation and genotypic diversity in a planktonic population of the toxigenic marine dinoflagellate *Alexandrium tamarense* (Dinophyceae). *Journal of Phycology* **46**, 18-32.

Altschul SF, Madden TL, Schäffer AA, *et al.* (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389-3402.

Alves-de-Souza C, Varela D, Iriarte JL, González HE, Guillou L (2012) Infection dynamics of Amoebophryidae parasitoids on harmful dinoflagellates in a southern Chilean fjord dominated by diatoms. *Aquatic Microbial Ecology* **66**, 183-197.

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

Anderson DM (1998) Physiology and bloom dynamics of toxic *Alexandrium* species, with emphasis on life cycle transitions. *Nato Asi Series G Ecological Sciences* **41**, 29-48.

Anderson DM (2009) Approaches to monitoring, control and management of harmful algal blooms (HABs). *Ocean & Coastal Management* **52**, 342-347.

Anderson DM, Alpermann TJ, Cembella AD, *et al.* (2012a) The globally distributed genus *Alexandrium*: Multifaceted roles in marine ecosystems and impacts on human health. *Harmful Algae* **14**, 10-35.

Anderson DM, Cembella AD, Hallegraeff GM (2012b) Progress in understanding harmful algal blooms: paradigm shifts and new technologies for research, monitoring, and management. *Annual Review of Marine Science* **4**, 143-176.

Anderson DM, Kulis DM, Binder BJ (1984) Sexuality and cyst formation in the dinoflagellate *Gonyaulax tamarensis*: cyst yield in batch cultures. *Journal of Phycology* **20**, 418-425.

Anderson DM, Kulis DM, Doucette GJ, Gallagher JC, Balech E (1994) Biogeography of toxic dinoflagellates in the genus *Alexandrium* from the northeastern United States and Canada. *Marine Biology* **120**, 467-478.

Anderson DM, Lindquist NL (1985) Time-course measurements of phosphorus depletion and cyst formation in the dinoflagellate *Gonyaulax tamarensis* Lebour. *Journal of Experimental Marine Biology and Ecology* **86**, 1-13.

Anderson RM, May R (1982) Coevolution of hosts and parasites. *Parasitology* **85**, 411-426.

Apel K, Hirt H (2004) Reactive oxygen species: metabolism, oxidative stress, and signal transduction. *Annual Review of Plant Biology* **55**, 373-399.

Ashburner M, Ball CA, Blake JA*, et al.* (2000) Gene Ontology: tool for the unification of biology. *Nature Genetics* **25**, 25-29.

Bachvaroff TR, Gornik SG, Concepcion GT*, et al.* (2014) Dinoflagellate phylogeny revisited: Using ribosomal proteins to resolve deep branching dinoflagellate clades. *Molecular Phylogenetics and Evolution* **70**, 314-322.

Bachvaroff TR, Place AR (2008) From stop to start: tandem gene arrangement, copy number and trans-splicing sites in the dinoflagellate *Amphidinium carterae*. *PloS One* **3**, e2929.

Bachvaroff TR, Place AR, Coats DW (2009) Expressed sequence tags from *Amoebophrya* sp. infecting *Karlodinium veneficum*: comparing host and parasite sequences. *Journal of Eukaryotic Microbiology* **56**, 531-541.

Bai X, Adolf JE, Bachvaroff T, Place AR, Coats DW (2007) The interplay between host toxins and parasitism by *Amoebophrya*. *Harmful Algae* **6**, 670-678.

Baker AC (2003) Flexibility and specificity in coral-algal symbiosis: diversity, ecology, and biogeography of *Symbiodinium*. *Annual Review of Ecology, Evolution, and Systematics*, 661-689.

Baldauf SL (2008) An overview of the phylogeny and diversity of eukaryotes. *Journal of Systematics and Evolution* **46**, 263-273.

Balech E (1995) The genus *Alexandrium Halim* (Dinoflagellata). Sherkin Island Marine Station, Cork, Ireland 151 pp.

Beauchemin M, Roy S, Daoust P*, et al.* (2012) Dinoflagellate tandem array gene transcripts are highly conserved and not polycistronic. *Proceedings of the National Academy of Sciences* **109**, 15793-15798.

Bentley DR, Balasubramanian S, Swerdlow HP*, et al.* (2008) Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **456**, 53-59.

Bergelson J, Kreitman M, Stahl EA, Tian D (2001) Evolutionary dynamics of plant R-genes. *Science* **292**, 2281-2285.

Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W (2011) Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* **27**, 578-579.

Boorstein RJ, Chiu L-N, Teebor GW (1989) Phylogenetic evidence of a role for 5-hydroxymethyluracil-DNA glycosylase in the maintenance of 5-methylcytosine in DNA. *Nucleic Acids Research* **17**, 7653-7661.

Borodovsky M, Lomsadze A (2011) Eukaryotic Gene Prediction Using GeneMark.hmm-E and GeneMark-ES. *Current Protocols in Bioinformatics*, editoral board, Andreas D. Baxevanis, *et al*. **CHAPTER**, Unit-4.610.

Bouligand Y, Norris V (2001) Chromosome separation and segregation in dinoflagellates and bacteria may depend on liquid crystalline states. *Biochimie* **83**, 187-192.

Brosnahan ML, Kulis DM, Solow AR*, et al.* (2010) Outbreeding lethality between toxic Group I and nontoxic Group III *Alexandrium tamarense* spp. isolates: predominance of heterotypic encystment and implications for mating interactions and biogeography. *Deep Sea Research Part II: Topical Studies in Oceanography* **57**, 175-189.

Brown GD, Reece KS (2003) Isolation and characterization of serine protease gene (s) from *Perkinsus marinus*. *Diseases of Aquatic Organisms* **57**, 117-126.

Bunn-Moreno M, Campos-Neto A (1981) Lectin (s) extracted from seeds of *Artocarpus integrifolia* (jackfruit): potent and selective stimulator (s) of distinct human T and B cell functions. *The Journal of Immunology* **127**, 427-429.

Burger G, Zhu Y, Littlejohn TG*, et al.* (2000) Complete sequence of the mitochondrial genome of *Tetrahymena pyriformis* and comparison with *Paramecium aurelia* mitochondrial DNA. *Journal of Molecular Biology* **297**, 365-380.

Burki F, Shalchian-Tabrizi K, Minge M*, et al.* (2007) Phylogenomics reshuffles the eukaryotic supergroups. *PloS One* **2**, e790.

Bustillos-Guzmán JJ, Band-Schmidt CJ, Durán-Riveroll LM*, et al.* (2015) Paralytic toxin profile of the marine dinoflagellate *Gymnodinium catenatum* Graham from the Mexican Pacific as revealed by LC-MS/MS. *Food Additives & Contaminants: Part A* **32**, 381-394.

Butterfield ER, Howe CJ, Nisbet RER (2013) An analysis of dinoflagellate metabolism using EST data. *Protist* **164**, 218-236.

Cachon J (1964) Contribution à l'étude des péridiniens parasites. Cytologie, cycles évolutifs. *Annales des Sciences Naturelles - Zoologie et Biologie Animale* **6**, 1-158.

Cachon J, Cachon M (1970) Ultrastructure des Amoebophryidae (*Péridiniens Duboscquodinida*) II Systèmes atractophoriens et microtubulaires; leur intervention dans la mitose. *Protistologica* **6**, 57-70.

Carruthers V, Sibley L (1997) Sequential protein secretion from three distinct organelles of *Toxoplasma gondii* accompanies invasion of human fibroblasts. *European Journal of Cell Biology* **73**, 114.

Cembella AD (1998) Ecophysiology and metabolism of paralytic shellfish toxins in marine microalgae. In: *Physiological Ecology of Harmful Algal Blooms* (eds. Anderson DM, Cembella AD, Hallegraeff GM), pp. 381-403. Springer Verlag, Berlin, Heidelberg, New York.

Cembella AD (2003) Chemical ecology of eukaryotic microalgae in marine ecosystems. *Phycologia* **42**, 420-447.

Chambouvet A, Alves-de-Souza C, Cueff V, *et al.* (2011a) Interplay between the parasite *Amoebophrya* sp. (Alveolata) and the cyst formation of the red tide dinoflagellate *Scrippsiella trochoidea*. *Protist* **162**, 637-649.

Chambouvet A, Laabir M, Sengco M, Vaquer A, Guillou L (2011b) Genetic diversity of Amoebophryidae (Syndiniales) during *Alexandrium catenella/tamarense* (Dinophyceae) blooms in the Thau lagoon (Mediterranean Sea, France). *Research in Microbiology* **162**, 959-968.

Chambouvet A, Morin P, Marie D, Guillou L (2008) Control of toxic marine dinoflagellate blooms by serial parasitic killers. *Science* **322**, 1254-1257.

Chan Y-H, Wong JT (2007) Concentration-dependent organization of DNA by the dinoflagellate histone-like protein HCc3. *Nucleic Acids Research* **35**, 2573-2583.

Chatton E (1912) Diagnoses préliminaires des péridiniers parasites nouveaux. *Bulletin de la Société Zoologique de France.* **37**, 85-93.

Chivers DP, Smith RJF (1998) Chemical alarm signalling in aquatic predator-prey systems: a review and prospectus. *Ecoscience*, 338-352.

Coats DW (1999) Parasitic life styles of marine dinoflagellates. *Journal of Eukaryotic Microbiology*, 402-409.

Coats DW, Adam E, Gallegos CL, Hedrick S (1996) Parasitism of photosynthetic dinoflagellates in a shallow subestuary of Chesapeake Bay, USA. *Aquatic Microbial Ecology* **11**, 1-9.

Coats DW, Bockstahler KR (1994) Occurrence of the parasitic dinoflagellate *Amoebophrya ceratii* in Chesapeake Bay populations of *Gymnodinium sanguineum*. *Journal of Eukaryotic Microbiology* **41**, 586-593.

Coats DW, Park MG (2002) Parasitism of photosynthetic dinoflagellates by three strains of *Amoebophrya* (Dinophyta): parasite survival, infectivity, generation time, and host specificity. *Journal of Phycology* **38**, 520-528.

Coffroth MA, Santos SR (2005) Genetic diversity of symbiotic dinoflagellates in the genus *Symbiodinium*. *Protist* **156**, 19-34.

Cowman AF, Crabb BS (2002) The *Plasmodium falciparum* genome - a blueprint for erythrocyte invasion. *Science* **298**, 126-128.

Curtis BA, Tanifuji G, Burki F*, et al.* (2012) Algal genomes reveal evolutionary mosaicism and the fate of nucleomorphs. *Nature* **492**, 59-65.

Danne JC, Gornik SG, MacRae JI, McConville MJ, Waller RF (2013) Alveolate mitochondrial metabolic evolution: dinoflagellates force reassessment of the role of parasitism as a driver of change in apicomplexans. *Molecular Biology and Evolution* **30**, 123-139.

Destombe C, Cembella A, Murphy C, Ragan M (1992) Nucleotide sequence of the 18S ribosomal RNA genes from the marine dinoflagellate *Alexandrium tamarense* (Gonyaulacales, Dinophyta). *Phycologia* **31**, 121-124.

Dobson A, Lafferty KD, Kuris AM, Hechinger RF, Jetz W (2008) Homage to Linnaeus: How many parasites? How many hosts? *Proceedings of the National Academy of Sciences* **105**, 11482-11489.

Dubremetz J, Garcia-Réguet N, Conseil V, Fourmaux MN (1998) Invited review apical organelles and host-cell invasion by Apicomplexa. *International Journal for Parasitology* **28**, 1007-1013.

Earl D, vonHoldt B (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**, 359-361.

Earnhart CG, Vogelbein MA, Brown GD, Reece KS, Kaattari SL (2004) Supplementation of *Perkinsus marinus* cultures with host plasma or tissue homogenate enhances their infectivity. *Applied and Environmental Microbiology* **70**, 421-431.

Emanuelsson O, Brunak S, von Heijne G, Nielsen H (2007) Locating proteins in the cell using TargetP, SignalP and related tools. *Nature Protocols* **2**, 953-971.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**, 2611-2620.

Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10**, 564-567.

Falkowski PG, Katz ME, Knoll AH*, et al.* (2004) The evolution of modern eukaryotic phytoplankton. *Science* **305**, 354-360.

Fast NM, Xue L, Bingham S, Keeling PJ (2002) Re‐examining alveolate evolution using multiple protein molecular phylogenies. *Journal of Eukaryotic Microbiology* **49**, 30-37.

Feder ME, Hofmann GE (1999) Heat-shock proteins, molecular chaperones, and the stress response: evolutionary and ecological physiology. *Annual Review of Physiology* **61**, 243-282.

Fensome RA, Saldarriaga JF, Taylor MF (1999) Dinoflagellate phylogeny revisited: reconciling morphological and molecular based phylogenies. *Grana* **38**, 66-80.

Fernie AR, Carrari F, Sweetlove LJ (2004) Respiratory metabolism: glycolysis, the TCA cycle and mitochondrial electron transport. *Current Opinion in Plant Biology* **7**, 254-261.

Figueroa RI, Dapena C, Bravo I, Cuadrado A (2015) The hidden sexuality of *Alexandrium Minutum*: an example of overlooked sex in dinoflagellates. *PloS One* **10**, e0142667.

Figueroa RI, Garces E, Bravo I (2007) Comparative study of the life cycles of *Alexandrium tamutum* and *Alexandrium minutum* (Gonyaulacales, Dinophyceae) in culture. *Journal of Phycology* **43**, 1039-1053.

Figueroa RI, Garcés E, Camp J (2010a) Reproductive plasticity and local adaptation in the host–parasite system formed by the toxic *Alexandrium minutum* and the dinoflagellate parasite *Parvilucifera sinerae*. *Harmful Algae* **10**, 56-63.

Figueroa RI, Rengefors K, Bravo I, Bensch S (2010b) From homothally to heterothally: Mating preferences and genetic variation within clones of the dinoflagellate *Gymnodinium catenatum*. *Deep Sea Research Part II: Topical Studies in Oceanography* **57**, 190-198.

Finn RD, Clements J, Eddy SR (2011) HMMER web server: interactive sequence similarity searching. *Nucleic Acids Research*, gkr367.

Flegontov P, Lukes J (2012) Mitochondrial genomes of photosynthetic Euglenids and Alveolates. In: *Mitochondrial Genome Evolution* (ed. Maréchal-Drouard L), pp. 127-153.

Flegontov P, Michálek J, Janouškovec J*, et al.* (2015) Divergent mitochondrial respiratory chains in phototrophic relatives of apicomplexan parasites. *Molecular Biology and Evolution* **32**, 1115-1131.

Flores HS, Wikfors GH, Dam HG (2012) Reactive oxygen species are linked to the toxicity of the dinoflagellate *Alexandrium* spp. to protists. *Aquatic Microbial Ecology* **66**, 199-209.

Flöthe CR, Molis M, Kruse I, Weinberger F, John U (2014) Herbivore-induced defence response in the brown seaweed *Fucus vesiculosus* (Phaeophyceae): temporal pattern and gene expression. *European Journal of Phycology* **49**, 356-369.

Foyer CH, Shigeoka S (2011) Understanding oxidative stress and antioxidant functions to enhance photosynthesis. *Plant physiology* **155**, 93-100.

Fritz L, Nass M (1992) Development of the endoparasitic dinoflagellate *Amoebophrya ceratii* within host dinoflagellate species. *Journal of Phycology* **28**, 312-320.

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.

Garcés E, Alacid E, Bravo I, Fraga S, Figueroa RI (2013) *Parvilucifera sinerae* (Alveolata, Myzozoa) is a generalist parasitoid of dinoflagellates. *Protist* **164**, 245-260.

Garreis KA, La Peyre JF, Faisal M (1996) The effects of *Perkinsus marinus* extracellular products and purified proteases on oyster defence parameters *in vitro*. *Fish & Shellfish Immunology* **6**, 581-597.

Gautier A, Michel-Salamin L, Tosi-Couture E, McDowall A, Dubochet J (1986) Electron microscopy of the chromosomes of dinoflagellates in situ: confirmation of Bouligand's liquid crystal hypothesis. *Journal of Ultrastructure and Molecular Structure Research* **97**, 10-30.

Gawryluk RM, Gray MW (2010) An ancient fission of mitochondrial cox1. *Molecular Biology and Evolution* **27**, 7-10.

Gayoso AM, Fulco VK (2006) Occurrence patterns of *Alexandrium tamarense* (Lebour) Balech populations in the Golfo Nuevo (Patagonia, Argentina), with observations on ventral pore occurrence in natural and cultured cells. *Harmful Algae* **5**, 233-241.

Gómez F (2012) A quantitative review of the lifestyle, habitat and trophic diversity of dinoflagellates (Dinoflagellata, Alveolata). *Systematics and Biodiversity* **10**, 267-275.

Gornik SG, Cassin AM, MacRae JI*, et al.* (2015) Endosymbiosis undone by stepwise elimination of the plastid in a parasitic dinoflagellate. *Proceedings of the National Academy of Sciences* **112**, 5767-5772.

Grabherr MG, Haas BJ, Yassour M*, et al.* (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* **29**, 644-652.

Gray MW, Lang BF, Burger G (2004) Mitochondria of protists. *Annual Review of Genetics* **38**, 477-524.

Guillard R (1973) Division rates. *Handbook of Phycological Methods* **1**, 289-312.

Guillou L, Viprey M, Chambouvet A*, et al.* (2008) Widespread occurrence and genetic diversity of marine parasitoids belonging to *Syndiniales* (Alveolata). *Environmental Microbiology* **10**, 3349-3365.

Gunderson JH, John SA, Chanson Boman W, Coats DW (2002) Multiple strains of the parasitic dinoflagellate *Amoebophrya* exist in Chesapeake Bay. *Journal of Eukaryotic Microbiology* **49**, 469-474.

Haas BJ, Papanicolaou A, Yassour M*, et al.* (2013) *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols* **8**, 1494-1512.

Hackett JD, Anderson DM, Erdner DL, Bhattacharya D (2004) Dinoflagellates: a remarkable evolutionary experiment. *American Journal of Botany* **91**, 1523-1534.

Hackett JD, Scheetz TE, Yoon HS*, et al.* (2005) Insights into a dinoflagellate genome through expressed sequence tag analysis. *BMC Genomics* **6**, 80.

Hallegraeff GM (1993) A review of harmful algal blooms and their apparent global increase. *Phycologia* **32**, 79-99.

Halliwell B, Gutteridge JM (2015) Free radicals in biology and medicine, Oxford University Press, USA.

Hartl DL, Clark AG, Clark AG (1997) Principles of population genetics, Sinauer associates Sunderland.

Harvell C, Kim K, Burkholder J*, et al.* (1999) Emerging marine diseases--climate links and anthropogenic factors. *Science* **285**, 1505-1510.

Hay ME (2009) Marine chemical ecology: chemical signals and cues structure marine populations, communities, and ecosystems. *Annual Review of Marine Science* **1**, 193.

Herrmann KM (1995) The shikimate pathway: early steps in the biosynthesis of aromatic compounds. *The Plant Cell* **7**, 907.

Hikosaka K, Watanabe Y-i, Tsuji N*, et al.* (2010) Divergence of the mitochondrial genome structure in the apicomplexan parasites, Babesia and Theileria. *Molecular Biology and Evolution* **27**, 1107-1116.

Horwitz J (1992) Alpha-crystallin can function as a molecular chaperone. *Proceedings of the National Academy of Sciences* **89**, 10449-10453.

Howe CJ, Nisbet RE, Barbrook AC (2008) The remarkable chloroplast genome of dinoflagellates. *Journal of Experimental Botany* **59**, 1035-1045.

Imlay JA (2003) Pathways of oxidative damage. *Annual Review of Microbiology* **57**, 395-418.

Jackson AP (2015) The evolution of parasite genomes and the origins of parasitism. *Parasitology* **142**, S1-S5.

Jackson C, Gornik S, Waller R (2012) The mitochondrial genome and transcriptome of the basal dinoflagellate *Hematodinium* sp.: character evolution within the highly derived mitochondrial genomes of dinoflagellates. *Genome Biology and Evolution* **4**, 59-72.

Jackson CJ, Norman JE, Schnare MN*, et al.* (2007) Broad genomic and transcriptional analysis reveals a highly derived genome in dinoflagellate mitochondria. *BMC Biology* **5**, 41.

Jaeckisch N, Yang I, Wohlrab S*, et al.* (2011) Comparative genomic and transcriptomic characterization of the toxigenic marine dinoflagellate *Alexandrium ostenfeldii*. *PloS One* **6**, e28012.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**, 1801-1806.

Janouškovec J, Horák A, Oborník M, Lukeš J, Keeling PJ (2010) A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids. *Proceedings of the National Academy of Sciences* **107**, 10949-10954.

Janson S, Gisselson L-Å, Salomon PS, Granéli E (2000) Evidence for multiple species within the endoparasitic dinoflagellate *Amoebophrya ceratii* as based on 18S rRNA gene-sequence analysis. *Parasitology Research* **86**, 929-933.

Jauzein C, Erdner DL (2013) Stress‐related responses in *Alexandrium tamarense* cells exposed to environmental changes. *Journal of Eukaryotic Microbiology* **60**, 526-538.

Jeong HJ, Du Yoo Y, Kim JS*, et al.* (2010) Growth, feeding and ecological roles of the mixotrophic and heterotrophic dinoflagellates in marine planktonic food webs. *Ocean Science Journal* **45**, 65-91.

Jeong HJ, Yoo Y, Park JY*, et al.* (2005) Feeding by phototrophic red-tide dinoflagellates: five species newly revealed and six species previously known to be mixotrophic. *Aquatic Microbial Ecology* **40**, 133-150.

Jingwen L, Nianzhi J, Huinnong C (2006) Cell cycle and cell signal transduction in marine phytoplankton. *Progress in Natural Science* **16**, 671-678.

John E, Flynn K (2000) Growth dynamics and toxicity of *Alexandrium fundyense* (Dinophyceae): the effect of changing N: P supply ratios on internal toxin and nutrient levels. *European Journal of Phycology* **35**, 11-23.

John U, Fensome RA, Medlin LK (2003) The application of a molecular clock based on molecular sequences and the fossil record to explain biogeographic distributions within the *Alexandrium tamarense* "species complex" (Dinophyceae). *Molecular Biology and Evolution* **20**, 1015-1027.

John U, Litaker RW, Montresor M*, et al.* (2014) Formal revision of the *Alexandrium tamarense* species complex (Dinophyceae) taxonomy: the introduction of five species with emphasis on molecular-based (rDNA) classification. *Protist* **165**, 779-804.

John U, Tillmann U, Hülskötter J*, et al.* (2015) Intraspecific facilitation by allelochemical mediated grazing protection within a toxigenic dinoflagellate population. *Proceedings of the Royal Society of London B: Biological Sciences* **282**, 20141268.

Joseph SJ, Fernandez-Robledo JA, Gardner MJ*, et al.* (2010) The Alveolate *Perkinsus marinus*: biological insights from EST gene discovery. *BMC Genomics* **11**, 228.

Kamikawa R, Nishimura H, Sako Y (2009) Analysis of the mitochondrial genome, transcripts, and electron transport activity in the dinoflagellate *Alexandrium catenella* (Gonyaulacales, Dinophyceae). *Phycological Research* **57**, 1-11.

Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M (2012) KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Research* **40**, D109-D114.

Katoh K, Misawa K, Kuma Ki, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* **30**, 3059-3066.

Kearse M, Moir R, Wilson A*, et al.* (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647-1649.

Keeling PJ (2010) The endosymbiotic origin, diversification and fate of plastids. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **365**, 729-748.

Keeling PJ (2013) The number, speed, and impact of plastid endosymbioses in eukaryotic evolution. *Annual Review of Plant Biology* **64**, 583-607.

Keeling PJ, Burger G, Durnford DG*, et al.* (2005) The tree of eukaryotes. *Trends in Ecology & Evolution* **20**, 670-676.

Keller MD, Selvin RC, Claus W, Guillard RRL (1987) Media for the culture of oceanic ultraphytoplankton. *Journal of Phycology* **23**, 633-638.

Kim K-Y, Kim C-H (2007) Phylogenetic relationships among diverse dinoflagellate species occurring in coastal waters off Korea inferred from large subunit ribosomal DNA sequence data. *Algae* **22**, 57-67.

Kim K-Y, Yoshida M, Fukuyo Y, Kim C-H (2002) Morphological observation of *Alexandrium tamarense* (Lebour) Balech, *A. catenella* (Whedon et Kofoid) Balech and one related morphotype (Dinophyceae) in Korea. *Algae* **17**, 11-19.

Kim S (2006) Patterns in host range for two strains of *Amoebophrya* (Dinophyta) Infecting thecate dinoflagellates: *Amoebophyra* spp. ex *Alexandrium affine* and ex *Gonyaulax polygramma*. *Journal of Phycology* **42**, 1170-1173.

Kim S, Gil Park M, Yih W, Coats DW (2004) Infection of the bloom-forming thecate dinoflagellates *Alexandrium affine* and *Gonyaulax spinifera* by two strains of *Amoebophrya* (Dinophyta). *Journal of Phycology* **40**, 815-822.

Kim S, Park MG (2016) Effect of the endoparasite *Amoebophrya* sp. on toxin content and composition in the paralytic shellfish poisoning dinoflagellate *Alexandrium fundyense* (Dinophyceae). *Harmful Algae* **51**, 10-15.

Kochin BF, Bull JJ, Antia R (2010) Parasite evolution and life history theory. *PLoS Biology* **8**, e1000524.

Kohli GS, John U, Van Dolah FM, Murray SA (2016) Evolutionary distinctiveness of fatty acid and polyketide synthesis in eukaryotes. *The ISME Journal*.

Koumandou VL, Nisbet RER, Barbrook AC, Howe CJ (2004) Dinoflagellate chloroplasts–where have all the genes gone? *TRENDS in Genetics* **20**, 261-267.

Krock B, Seguel CG, Cembella AD (2007) Toxin profile of *Alexandrium catenella* from the Chilean coast as determined by liquid chromatography with fluorescence detection and liquid chromatography coupled with tandem mass spectrometry. *Harmful Algae* **6**, 734-744.

Krogh A, Larsson B, von Heijne G, Sonnhammer ELL (2001) Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *Journal of Molecular Biology* **305**, 567-580.

Kühn SF (1998) Infection of *Coscinodiscus* spp. by the parasitoid nanoflagellate *Pirsonia diadema*: II. Selective infection behaviour for host species and individual host cells. *Journal of Plankton Research* **20**, 443-454.

Kuris AM (1974) Trophic interactions: similarity of parasitic castrators to parasitoids. *Quarterly Review of Biology* **49**, 129.

LaJeunesse TC, Lambert G, Andersen RA, Coffroth MA, Galbraith DW (2005) *Symbiodinium* (pyrrhophyta) genome sizes (DNA content) are smallest among dinoflagellates. *Journal of Phycology* **41**, 880-886.

Lam CMC, New DC, Wong JTY (2001) cAMP in the cell cycle of the dinoflagellate *Crypthecodinium cohnii* (Dinophyta). *Journal of Phycology* **37**, 79-85.

Landsberg JH (2002) The effects of harmful algal blooms on aquatic organisms. *Reviews in Fisheries Science* **10**, 113-390.

Leander BS, Clopton RE, Keeling PJ (2003) Phylogeny of gregarines (Apicomplexa) as inferred from small-subunit rDNA and β-tubulin. *International Journal of Systematic and Evolutionary Microbiology* **53**, 345-354.

Learngaramkul P, Petmitr S, Krungkrai SR, Prapunwattana P, Krungkrai J (1999) Molecular characterization of mitochondria in asexual and sexual blood stages of *Plasmodium falciparum*. *Molecular Cell Biology Research Communications* **2**, 15-20.

Leblond JD, Dahmen JL (2012) Mono-and digalactosyldiacylglycerol composition of dinoflagellates. V. The galactolipid profile of *Alexandrium tamarense* (Dinophyceae) during the course of infection by the parasitic syndinian dinoflagellate *Amoebophrya* sp. *European Journal of Phycology* **47**, 490-497.

Lesser MP (2006) Oxidative stress in marine environments: biochemistry and physiological ecology. *Annual Review of Physiology* **68**, 253-278.

Lesser MP (2012) Oxidative stress in tropical marine ecosystems. *Oxidative Stress in Aquatic Ecosystems* **1**, 9-19.

Lewitus AJ, Horner RA, Caron DA*, et al.* (2012) Harmful algal blooms along the North American west coast region: History, trends, causes, and impacts. *Harmful Algae* **19**, 133-159.

Li C, Song S, Liu Y, Chen T (2014) Occurrence of *Amoebophrya* spp. infection in planktonic dinoflagellates in Changjiang (Yangtze River) Estuary, China. *Harmful Algae* **37**, 117-124.

Li L, Stoeckert CJ, Roos DS (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Research* **13**, 2178-2189.

Libro S, Kaluziak ST, Vollmer SV (2013) RNA-seq profiles of immune related genes in the staghorn coral *Acropora cervicornis* infected with white band disease. *PloS One* **8**, e81821.

Lidie KB, van Dolah FM (2007) Spliced leader RNA-mediated trans-splicing in a dinoflagellate, *Karenia brevis*. *Journal of Eukaryotic Microbiology* **54**, 427-435.

Lilly EL, Halanych KM, Anderson DM (2007) Species boundaries and global biogeography of the *Alexandrium tamarense* complex (Dinophyceae). *Journal of Phycology* **43**, 1329-1338.

Lin S (2011) Genomic understanding of dinoflagellates. *Research in Microbiology* **162**, 551-569.

Lin S, Cheng S, Song B*, et al.* (2015) The *Symbiodinium kawagutii* genome illuminates dinoflagellate gene expression and coral symbiosis. *Science* **350**, 691-694.

Lin S, Zhang H, Hou Y, Zhuang Y, Miranda L (2009) High-level diversity of dinoflagellates in the natural environment, revealed by assessment of mitochondrial cox1 and cob genes for dinoflagellate DNA barcoding. *Applied and Environmental Microbiology* **75**, 1279-1290.

Little TJ, Watt K, Ebert D (2006) Parasite - host specificity: experimental studies on the basis of parasite adaptation. *Evolution* **60**, 31-38.

Liu L, Hastings J (2006) Novel and rapidly diverging intergenic sequences between tandem repeats of the luciferase genes in seven dinoflagellate species. *Journal of Phycology* **42**, 96-103.

Livaja M, Palmieri MC, von Rad U, Durner J (2008) The effect of the bacterial effector protein harpin on transcriptional profile and mitochondrial proteins of *Arabidopsis thaliana*. *Journal of Proteomics* **71**, 148-159.

Llewellyn LE (2006) Saxitoxin, a toxic marine natural product that targets a multitude of receptors. *Natural Product Reports* **23**, 200-222.

Lokanathan Y, Mohd-Adnan A, Wan KL, Nathan S (2010) Transcriptome analysis of the *Cryptocaryon irritans* tomont stage identifies potential genes for the detection and control of cryptocaryonosis. *BMC Genomics* **11**, 76.

Lowe CD, Keeling PJ, Martin LE*, et al.* (2010) Who is *Oxyrrhis marina*? Morphological and phylogenetic studies on an unusual dinoflagellate. *Journal of Plankton Research*, fbq110.

Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research* **25**, 955-964.

Lu Y, Wohlrab S, Glöckner G, Guillou L, John U (2014) Genomic insights in processes driving the infection of *Alexandrium tamarense* by the parasitoid *Amoebophrya* sp. *Eukaryotic Cell* **13**, 1439-1449.

Lu Y, Wohlrab S, Groth M*, et al.* (2016) Transcriptomic profiling of *Alexandrium fundyense* during physical interaction with or exposure to chemical signals from the parasite *Amoebophrya*. *Molecular Ecology*.

Lukeš J, Leander BS, Keeling PJ (2009) Cascades of convergent evolution: the corresponding evolutionary histories of euglenozoans and dinoflagellates. *Proceedings of the National Academy of Sciences* **106**, 9963-9970.

Lyons R, Johnson A (1995) Heat shock proteins of *Toxoplasma gondii*. *Parasite Immunology* **17**, 353-359.

Ma H, Krock B, Tillmann U, Cembella A (2009) Preliminary characterization of extracellular allelochemicals of the toxic marine dinoflagellate *Alexandrium tamarense* using a *Rhodomonas salina* bioassay. *Marine Drugs* **7**, 497-522.

MacIntyre E, Earnhart C, Kaattari S (2003) Host oyster tissue extracts modulate *in vitro* protease expression and cellular differentiation in the protozoan parasite, *Perkinsus marinus*. *Parasitology* **126**, 293-302.

Mardones JI, Dorantes-Aranda JJ, Nichols PD, Hallegraeff GM (2015) Fish gill damage by the dinoflagellate *Alexandrium catenella* from Chilean fjords: Synergistic action of ROS and PUFA. *Harmful Algae* **49**, 40-49.

Maréchal E, Azzouz N, de Macedo CS*, et al.* (2002) Synthesis of chloroplast galactolipids in apicomplexan parasites. *Eukaryotic Cell* **1**, 653-656.

Maroney P, Denker JA, Darzynkiewicz E, Laneve R, Nilsen TW (1995) Most mRNAs in the nematode *Ascaris lumbricoides* are trans-spliced: a role for spliced leader addition in translational efficiency. *Rna* **1**, 714-723.

Martel CM (2009) Conceptual bases for prey biorecognition and feeding selectivity in the microplanktonic marine phagotroph *Oxyrrhis marina*. *Microbial Ecology* **57**, 589-597.

Mazumdar J, Striepen B (2007) Make it or take it: fatty acid metabolism of apicomplexan parasites. *Eukaryotic Cell* **6**, 1727-1735.

Mazzillo FFM, Ryan JP, Silver MW (2011) Parasitism as a biological control agent of dinoflagellate blooms in the California Current System. *Harmful Algae* **10**, 763-773.

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

Mendez GS, Delwiche CF, Apt KE, Lippmeier JC (2015) Dinoflagellate gene structure and intron splice sites in a genomic tandem array. *Journal of Eukaryotic Microbiology* **62**, 679-687.

Mihailovich M, Militti C, Gabaldón T, Gebauer F (2010) Eukaryotic cold shock domain proteins: highly versatile regulators of gene expression. *Bioessays* **32**, 109-118.

Miller JJ, Delwiche CF, Coats DW (2012) Ultrastructure of *Amoebophrya* sp. and its changes during the course of infection. *Protist* **163**, 720-745.

Miura C, Sugawara K, Neriya Y*, et al.* (2012) Functional characterization and gene expression profiling of superoxide dismutase from plant pathogenic phytoplasma. *Gene* **510**, 107-112.

Montagnes DJ, Chambouvet A, Guillou L, Fenton A (2008) Responsibility of microzooplankton and parasite pressure for the demise of toxic dinoflagellate blooms. *Aquatic Microbial Ecology* **53**, 211.

Morey JS, Monroe EA, Kinney AL*, et al.* (2011) Transcriptomic response of the red tide dinoflagellate, *Karenia brevis*, to nitrogen and phosphorus depletion and addition. *BMC Genomics* **12**, 346.

Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M (2007) KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Research* **35**, W182-W185.

Morris SC (1981) Parasites and the fossil record. *Parasitology* **82**, 489-509.

Mungpakdee S, Shinzato C, Takeuchi T*, et al.* (2014) Massive gene transfer and extensive RNA editing of a symbiotic dinoflagellate plastid genome. *Genome Biology and Evolution* **6**, 1408-1422.

Nagai S, Lian C, Hamaguchi M*, et al.* (2004) Development of microsatellite markers in the toxic dinoflagellate *Alexandrium tamarense* (Dinophyceae). *Molecular Ecology Notes* **4**, 83-85.

Nash EA, Barbrook AC, Edwards-Stuart RK*, et al.* (2007) Organization of the mitochondrial genome in the dinoflagellate *Amphidinium carterae*. *Molecular Biology and Evolution* **24**, 1528-1536.

Nishitani L, Erickson G, Chew K (1985) Role of the parasitic dinoflagellate *Amoebophrya ceratii* in control of *Gonyaulax catenella* populations. *Toxic Dinoflagellates.* Elsevier Sci. Pub. Co. Inc. New York, 225-230.

Ogata T, Ishimaru T, Kodama M (1987) Effect of water temperature and light intensity on growth rate and toxicity change in *Protogonyaulax tamarensis*. *Marine Biology* **95**, 217-220.

Okamoto N, Horák A, Keeling PJ (2012) Description of two species of early branching dinoflagellates, *Psammosa pacifica* ng, n. sp. and *P. atlantica* n. sp. *PloS One* **7**.

Orlova TY, Selina MS, Lilly EL, Kulis DM, Anderson DM (2007) Morphogenetic and toxin composition variability of *Alexandrium tamarense* (Dinophyceae) from the east coast of Russia. *Phycologia* **46**, 534-548.

Palenchar JB, Bellofatto V (2006) Gene transcription in trypanosomes. *Molecular and Biochemical Parasitology* **146**, 135-141.

Park MG, Cooney SK, Kim JS, Coats DW (2002a) Effects of parasitism on diel vertical migration, phototaxis/geotaxis, and swimming speed of the bloom-forming dinoflagellate *Akashiwo sanguinea*. *Aquatic Microbial Ecology* **29**, 11-18.

Park MG, Cooney SK, Yih W, Coats DW (2002b) Effects of two strains of the parasitic dinoflagellate *Amoebophrya* on growth, photosynthesis, light absorption, and quantum yield of bloom-forming dinoflagellates. *Marine Ecology Progress Series* **227**, 281-292.

Park MG, Kim S, Shin E-Y, Yih W, Coats DW (2013) Parasitism of harmful dinoflagellates in Korean coastal waters. *Harmful Algae* **30, Supplement 1**, S62-S74.

Park MG, Yih W, Coats DW (2004) Parasites and phytoplankton, with special emphasis on dinoflagellate infections. *Journal of Eukaryotic Microbiology* **51**, 145-155.

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

Parsell D, Lindquist S (1993) The function of heat-shock proteins in stress tolerance: degradation and reactivation of damaged proteins. *Annual Review of Genetics* **27**, 437-496.

Pérez-Martınez X, Antaramian A, Vázquez-Acevedo M*, et al.* (2001) Subunit II of cytochrome c oxidase in Chlamydomonad algae is a heterodimer encoded by two independent nuclear genes. *Journal of Biological Chemistry* **276**, 11302-11309.

Petersen TN, Brunak S, von Heijne G, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods* **8**, 785-786.

PEYRE JF, SCHAFHAUSER DY, RIZKALLA EH, FAISAL M (1995) Production of serine proteases by the oyster pathogen *Perkinsus marinus* (Apicomplexa) *in vitro*. *Journal of Eukaryotic Microbiology* **42**, 544-551.

Peyretaillade E, Broussolle V, Peyret P*, et al.* (1998) Microsporidia, amitochondrial protists, possess a 70-kDa heat shock protein gene of mitochondrial evolutionary origin. *Molecular Biology and Evolution* **15**, 683-689.

Plattner F, Soldati-Favre D (2008) Hijacking of host cellular functions by the Apicomplexa. *Annual Review of Microbiology* **62**, 471-487.

Plattner H, Sehring IM, Mohamed IK*, et al.* (2012) Calcium signaling in closely related protozoan groups (Alveolata): Non-parasitic ciliates (*Paramecium*, *Tetrahymena*) vs. parasitic Apicomplexa (*Plasmodium*, *Toxoplasma*). *Cell Calcium* **51**, 351-382.

Pombert J-F, Blouin NA, Lane C, Boucias D, Keeling PJ (2014) A lack of parasitic reduction in the obligate parasitic green alga *Helicosporidium*. *PLoS Genet* **10**, e1004355.

Poulin R (2011) Evolutionary ecology of parasites. Princeton university press.

Powell S, Szklarczyk D, Trachana K*, et al.* (2012) eggNOG v3. 0: orthologous groups covering 1133 organisms at 41 different taxonomic ranges. *Nucleic Acids Research* **40**, D284-D289.

Price PW (1980) Evolutionary biology of parasites. Princeton University Press.

Pritchard A, Seilhamer J, Mahalingam R*, et al.* (1990) Nucleotide sequence of the mitochondrial genome of Paramecium. *Nucleic Acids Research* **18**, 173-180.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.

Punta M, Coggill PC, Eberhardt RY*, et al.* (2012) The Pfam protein families database. *Nucleic Acids Research* **40**, D290-D301.

Råberg L, Alacid E, Garces E, Figueroa R (2014) The potential for arms race and Red Queen coevolution in a protist host–parasite system. *Ecology and Evolution* **4**, 4775-4785.

Ramakrishnan M, Sowjanya TN, Raj KB, Kasbekar DP (2011) A factor in a wild isolated *Neurospora crassa* strain enables a chromosome segment duplication to suppress repeat-induced point mutation. *Journal Biosciences* **36**, 817-821.

Raval S, Gowda SB, Singh DD, Chandra NR (2004) A database analysis of jacalin-like lectins: sequence–structure–function relationships. *Glycobiology* **14**, 1247-1263.

Reece KS, Siddall ME, Burreson EM, Graves JE (1997) Phylogenetic analysis of *Perkinsus* based on actin gene sequences. *The Journal of parasitology*, 417-423.

Richards TA, Dacks JB, Campbell SA*, et al.* (2006) Evolutionary origins of the eukaryotic shikimate pathway: gene fusions, horizontal gene transfer, and endosymbiotic replacements. *Eukaryotic Cell* **5**, 1517-1531.

Roberts EC, Legrand C, Steinke M, Wootton EC (2011) Mechanisms underlying chemical interactions between predatory planktonic protists and their prey. *Journal of plankton research*, fbr005.

Roberts EC, Zubkov MV, Martin‐Cereceda M, Novarino G, Wootton EC (2006) Cell surface lectin‐binding glycoconjugates on marine planktonic protists. *FEMS Microbiology Letters* **265**, 202-207.

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes* **4**, 137-138.

Roy S, Morse D (2013) Transcription and maturation of mRNA in dinoflagellates. *Microorganisms* **1**, 71-99.

Sala-Rovira M, Geraud M, Caput D*, et al.* (1991) Molecular cloning and immunolocalization of two variants of the major basic nuclear protein (HCc) from the histone-less eukaryote *Crypthecodinium cohnii* (Pyrrhophyta). *Chromosoma* **100**, 510-518.

Saldarriaga JF, McEwan ML, Fast NM, Taylor F, Keeling PJ (2003) Multiple protein phylogenies show that *Oxyrrhis marina* and *Perkinsus marinus* are early branches of the dinoflagellate lineage. *International Journal of Systematic and Evolutionary Microbiology* **53**, 355-365.

Saldarriaga JF, Taylor F, Keeling PJ, Cavalier-Smith T (2001) Dinoflagellate nuclear SSU rRNA phylogeny suggests multiple plastid losses and replacements. *Journal of Molecular Evolution* **53**, 204-213.

Salomon P, Imai I (2006) Pathogens of harmful microalgae. In: *Ecology of Harmful Algae*, pp. 271-282. Springer.

Salomon PS, Janson S, Granéli E (2003a) Multiple species of the dinophagous dinoflagellate genus *Amoebophrya* infect the same host species. *Environmental Microbiology* **5**, 1046-1052.

Salomon PS, Janson S, Granéli E (2003b) Parasitism of *Dinophysis norvegica* by *Amoebophrya* sp. in the Baltic Sea. *Aquatic Microbial Ecology* **33**, 163-172.

Salomon PS, Stolte W (2010) Predicting the population dynamics in *Amoebophrya* parasitoids and their dinoflagellate hosts using a mathematical model. *Marine Ecology Progress Series* **419**, 1-10.

Sano J, Kato KH (2009) Localization and copy number of the protein-coding genes actin, α-tubulin, and Hsp90 in the nucleus of a primitive dinoflagellate, *Oxyrrhis marina*. *Zoological Science* **26**, 745-753.

Satou Y, Hamaguchi M, Takeuchi K, Hastings KEM, Satoh N (2006) Genomic overview of mRNA 5′-leader trans-splicing in the ascidian *Ciona intestinalis*. *Nucleic Acids Research* **34**, 3378-3388.

Scandalios J (2005) Oxidative stress: molecular perception and transduction of signals triggering antioxidant gene defenses. *Brazilian Journal of Medical and Biological Research* **38**, 995-1014.

Scholin CA, Herzog M, Sogin M, Anderson DM (1994) Identification of group- and strain-specific genetic markers for globally distributed *Alexandrium* (dinophyceae). II. Sequence analysis of a fragment of the LSU rRNA gene. *Journal of Phycology* **30**, 999-1011.

Secord D (2003) Biological control of marine invasive species: cautionary tales and land-based lessons. *Biological Invasions* **5**, 117-131.

Seeber F, Limenitakis J, Soldati-Favre D (2008) Apicomplexan mitochondrial metabolism: a story of gains, losses and retentions. *Trends in Parasitology* **24**, 468-478.

Selander E, Thor P, Toth G, Pavia H (2006) Copepods induce paralytic shellfish toxin production in marine dinoflagellates. *Proceedings of the Royal Society of London B: Biological Sciences* **273**, 1673-1680.

Shilo M (1967) Formation and mode of action of algal toxins. *Bacteriological Reviews* **31**, 180.

Shoguchi E, Shinzato C, Hisata K, Satoh N, Mungpakdee S (2015) The large mitochondrial genome of *Symbiodinium minutum* reveals conserved noncoding sequences between dinoflagellates and apicomplexans. *Genome Biology and Evolution* **7**, 2237-2244.

Shoguchi E, Shinzato C, Kawashima T*, et al.* (2013) Draft assembly of the *Symbiodinium minutum* nuclear genome reveals dinoflagellate gene structure. *Current Biology* **23**, 1399-1408.

Siano R, Alves-de-Souza C, Foulon E*, et al.* (2010) Distribution and host diversity of *Amoebophryidae* parasites across oligotrophic waters of the Mediterranean Sea. *Biogeosciences Discussions* **7**, 7391-7419.

Skovgaard A (2014) Dirty tricks in the plankton: Diversity and Role of Marine Parasitic Protists: diversity and role of marine parasitic protists. *Acta Protozoologica* **53**, 51-62.

Skovgaard A, Massana R, Balagué V, Saiz E (2005) Phylogenetic position of the copepod-infesting parasite *Syndinium turbo* (Dinoflagellata, Syndinea). *Protist* **156**, 413-423.

Slamovits CH, Keeling PJ (2008a) Plastid-derived genes in the nonphotosynthetic alveolate *Oxyrrhis marina*. *Molecular Biology and Evolution* **25**, 1297-1306.

Slamovits CH, Keeling PJ (2008b) Widespread recycling of processed cDNAs in dinoflagellates. *Current Biololgy* **18**, R550-552.

Slamovits CH, Saldarriaga JF, Larocque A, Keeling PJ (2007) The highly reduced and fragmented mitochondrial genome of the early-branching dinoflagellate *Oxyrrhis marina* shares characteristics with both apicomplexan and dinoflagellate mitochondrial genomes. *Journal of Molecular Biology* **372**, 356-368.

Smayda TJ (2002) Turbulence, watermass stratification and harmful algal blooms: an alternative view and frontal zones as "pelagic seed banks". *Harmful Algae* **1**, 95-112.

Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688-2690.

Stanke M, Steinkamp R, Waack S, Morgenstern B (2004) AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic Acids Research* **32**, W309-W312.

Steele RE, Rae PM (1980) Ordered distribution of modified bases in the DNA of a dinoflagellate. *Nucleic Acids Research* **8**, 4709-4726.

Stentiford GD, Shields JD (2005) A review of the parasitic dinoflagellates *Hematodinium* species and *Hematodinium*-like infections in marine crustaceans. *Diseases of Aquatic Organisms* **66**, 47.

Subramanian A, Tamayo P, Mootha VK*, et al.* (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 15545-15550.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* **30**, 2725-2729.

Taylor F (1968) Parasitism of the toxin-producing dinoflagellate *Gonyaulax catenella* by the endoparasitic dinoflagellate *Amoebophrya ceratii*. *Journal of the Fisheries Board of Canada* **25**, 2241-2245.

Teng CY, Dang Y, Danne JC, Waller RF, Green BR (2013) Mitochondrial genes of dinoflagellates are transcribed by a nuclear-encoded single-subunit RNA polymerase. *PloS One* **8**, e65387.

Tenter AM, Heckeroth AR, Weiss LM (2000) *Toxoplasma gondii*: from animals to humans. *International Journal for Parasitology* **30**, 1217-1258.

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**, 4673-4680.

Thompson JN (1994) The coevolutionary process. University of Chicago Press.

Thompson JN (1998) Rapid evolution as an ecological process. *Trends in Ecology & Evolution* **13**, 329-332.

Thompson JN (1999) The evolution of species interactions. *Science* **284**, 2116-2118.

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 TL, da Purificação RC, Krock B, Cembella A (2009) Intra-population clonal variability in allelochemical potency of the toxigenic dinoflagellate *Alexandrium tamarense*. *Harmful Algae* **8**, 759-769.

Tillmann U, Hansen PJ (2009) Allelopathic effects of *Alexandrium tamarense* on other algae: evidence from mixed growth experiments. *Aquatic Microbial Ecology* **57**, 101-112.

Tillmann U, Hesse K-J, Tillmann A (1999) Large-scale parasitic infection of diatoms in the Northfrisian Wadden Sea. *Journal of Sea Research* **42**, 255-261.

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.

Torchin ME, Lafferty KD, Dobson AP, McKenzie VJ, Kuris AM (2003) Introduced species and their missing parasites. *Nature* **421**, 628-630.

Toth GB, Norén F, Selander E, Pavia H (2004) Marine dinoflagellates show induced life-history shifts to escape parasite infection in response to water-borne signals. *Proceedings of the Royal Society of London, Series B: Biological Sciences* **271**, 733-738.

Toth GB, Pavia H (2000) Water-borne cues induce chemical defense in a marine alga (*Ascophyllum nodosum*). *Proceedings of the National Academy of Sciences* **97**, 14418-14420.

Tourancheau AB, Tsao N, Klobutcher LA, Pearlman RE, Adoutte A (1995) Genetic code deviations in the ciliates: evidence for multiple and independent events. *The EMBO Journal* **14**, 3262-3267.

Trapnell C, Pachter L, Salzberg SL (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105-1111.

Vaidya AB, Mather MW (2009) Mitochondrial evolution and functions in malaria parasites. *Annual Review of Microbiology* **63**, 249-267.

Vale P (2010) New saxitoxin analogues in the marine environment: developments in toxin chemistry, detection and biotransformation during the 2000s. *Phytochemistry Reviews* **9**, 525-535.

Van Dooren GG, Stimmler LM, McFadden GI (2006) Metabolic maps and functions of the *Plasmodium* mitochondrion. *FEMS Microbiology Reviews* **30**, 596-630.

Van Valen L (1973) A new evolutionary law. *Evolutionary theory* **1**, 1-30.

Vanhamme L, Pays E (1995) Control of gene expression in trypanosomes. *Microbiological Reviews* **59**, 223-240.

Velo-Suárez L, Brosnahan ML, Anderson DM, McGillicuddy DJ, Jr. (2013) A quantitative assessment of the role of the parasite *Amoebophrya* in the termination of *Alexandrium fundyense* blooms within a small coastal embayment. *PloS One* **8**, e81150.

Verret F, Wheeler G, Taylor AR, Farnham G, Brownlee C (2010) Calcium channels in photosynthetic eukaryotes: implications for evolution of calcium-based signalling. *New Phytologist* **187**, 23-43.

Walker G, Dorrell RG, Schlacht A, Dacks JB (2011) Eukaryotic systematics: a user's guide for cell biologists and parasitologists. *Parasitology* **138**, 1638-1663.

Waller RF, Jackson CJ (2009) Dinoflagellate mitochondrial genomes: stretching the rules of molecular biology. *Bioessays* **31**, 237-245.

Waller RF, Keeling PJ (2006) Alveolate and chlorophycean mitochondrial cox2 genes split twice independently. *Gene* **383**, 33-37.

Waller RF, McFadden GI (2005) The apicoplast: a review of the derived plastid of apicomplexan parasites. *Current Issues in Molecular Biology* **7**, 57-80.

Wang L, Zhuang Y, Zhang H, Lin X, Lin S (2014) DNA barcoding species in *Alexandrium tamarense* complex using ITS and proposing designation of five species. *Harmful Algae* **31**, 100-113.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population-structure. *Evolution* **38**, 1358-1370.

Weiss LM, Ma YF, Takvorian PM, Tanowitz HB, Wittner M (1998) Bradyzoite development in *Toxoplasma gondii* and the hsp70 stress response. *Infection and Immunity* **66**, 3295-3302.

Wetzel J, Kingsford C, Pop M (2011) Assessing the benefits of using mate-pairs to resolve repeats in de novo short-read prokaryotic assemblies. *BMC Bioinformatics* **12**, 1.

White A (1978) Salinity effects on growth and toxin content of *Gonyaulax excavata*, a marine dinoflagellate causing paralytic shellfish poisoning. *Journal of Phycology* **14**, 475-479.

Wisecaver JH, Hackett JD (2011) Dinoflagellate genome evolution. *Annual Review of Microbiology* **65**, 369-387.

Wohlrab S (2013) Characterization of grazer-induced responses in the marine dinoflagellate Alexandrium tamarense. Universität Bremen.

Wohlrab S, Iversen MH, John U (2010) A molecular and co-evolutionary context for grazer induced toxin production in *Alexandrium tamarense*. *PloS One* **5**, e15039.

Woolhouse ME, Webster JP, Domingo E, Charlesworth B, Levin BR (2002) Biological and biomedical implications of the co-evolution of pathogens and their hosts. *Nature Genetics* **32**, 569-577.

Wootton EC, Zubkov MV, Jones DH*, et al.* (2007) Biochemical prey recognition by planktonic protozoa. *Environmental Microbiology* **9**, 216-222.

Wurzinger B, Mair A, Pfister B, Teige M (2011) Cross-talk of calcium-dependent protein kinase and MAP kinase signaling. *Plant Signaling & Behavior* **6**, 8-12.

Wyatt T, Jenkinson IR (1997) Notes on *Alexandrium* population dynamics. *Journal of Plankton research* **19**, 551-575.

Xu P, Widmer G, Wang Y, *et al.* (2004) The genome of *Cryptosporidium hominis*. *Nature* **431**, 1107-1112.

Yang I, John U, Beszteri S, *et al.* (2010) Comparative gene expression in toxic versus non-toxic strains of the marine dinoflagellate *Alexandrium minutum*. *BMC Genomics* **11**, 248.

Young JC, Moarefi I, Hartl FU (2001) Hsp90: a specialized but essential protein-folding tool. *Journal of Cell Biology* **154**, 267-274.

Zhang H, Campbell DA, Sturm NR, Lin S (2009) Dinoflagellate spliced leader RNA genes display a variety of sequences and genomic arrangements. *Molecular Biology and Evolution* **26**, 1757-1771.

Zhang H, Hou Y, Miranda L, *et al.* (2007) Spliced leader RNA trans-splicing in dinoflagellates. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 4618-4623.

Zhang H, Lin S (2008) mRNA editing and spliced-leader RNA *trans*-splicing groups *Oxyrrhis, Noctiluca, Heterocapsa*, and *Amphidinium* as basal lineages of dinoflagellates. *Journal of Phycology* **44**, 703-711.

Zhang Z, Green B, Cavalier-Smith T (1999) Single gene circles in dinoflagellate chloroplast genomes. *Nature* **400**, 155-159.

Zhu G, Marchewka MJ, Keithly JS (2000) *Cryptosporidium parvum* appears to lack a plastid genome. *Microbiology* **146**, 315-321.

Erklärung

Yameng Lu
Borriesstr.17
27570 Bremerhaven

Erklärung gemäß §6(5) der Promotionsordnung der Universität Bremen für die mathematischen, natur- und ingenieurwissenschaftlichen Fachbereiche vom 14. März 2007

Hiermit erkläre ich, Yameng Lu, dass ich die Doktorarbeit mit dem Titel:

“Control of texigenic dinoflagellates through parasitism: Implications for host-parasite coevolution”

selbstständig verfasst und geschrieben habe und außer den angegebenen Quellen keine weiteren Hilfsmittel verwendet habe.
