This article was downloaded by: *[Stiftung]*
On: *6 September 2010*
Access details: *Access Details: [subscription number 918151215]*
Publisher *Taylor & Francis*
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK

European Journal of Phycology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713725516

Genomic characterisation of the ichthyotoxic prymnesiophyte *Chrysochromulina polylepis,* and the expression of polyketide synthase genes in synchronized cultures

Uwe John[a]; Sara Beszteri[a]; Gernot Glöckner[b]; Rama Singh[c]; Linda Medlin[de]; Allan D. Cembella[a]
[a] Alfred-Wegener-Institute for Polar and Marine Research, Am Handelshafen 12, D-27570 Bremerhaven, Germany [b] Berlin Center for Genomics in Biodiversity Research, Botanischer Garten und Botanisches Museum Berlin-Dahlem, Freie Universität Berlin, 14195 Berlin, Germany [c] Institute for Marine Biosciences, National Research Council of Canada, Halifax, Nova Scotia, B3H 3Z1 Canada [d] UPMC Univ Paris 06, F-66651 Banyuls/mer, France [e] CNRS, UMR 7621, LOMIC, Observatoire Océanologique, F-66651 Banyuls/mer, France

First published on: 25 August 2010

To cite this Article John, Uwe , Beszteri, Sara , Glöckner, Gernot , Singh, Rama , Medlin, Linda and Cembella, Allan D.(2010) 'Genomic characterisation of the ichthyotoxic prymnesiophyte *Chrysochromulina polylepis,* and the expression of polyketide synthase genes in synchronized cultures', European Journal of Phycology, 45: 3, 215 — 229, First published on: 25 August 2010 (iFirst)
To link to this Article: DOI: 10.1080/09670261003746193
URL: http://dx.doi.org/10.1080/09670261003746193

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.informaworld.com/terms-and-conditions-of-access.pdf

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Eur. J. Phycol. (2010) **45(3)**: 215–229

Taylor & Francis
Taylor & Francis Group

Genomic characterisation of the ichthyotoxic prymnesiophyte *Chrysochromulina polylepis*, and the expression of polyketide synthase genes in synchronized cultures

UWE JOHN[1], SARA BESZTERI[1], GERNOT GLÖCKNER[2], RAMA SINGH[3], LINDA MEDLIN[4,5] AND ALLAN D. CEMBELLA[1]

[1]*Alfred-Wegener-Institute for Polar and Marine Research, Am Handelshafen 12, D-27570 Bremerhaven, Germany*
[2]*Berlin Center for Genomics in Biodiversity Research, Botanischer Garten und Botanisches Museum Berlin-Dahlem, Freie Universität Berlin, Konigin-Luise-Straße 6-8, 14195 Berlin, Germany*
[3]*Institute for Marine Biosciences, National Research Council of Canada, 1411 Oxford Street, Halifax, Nova Scotia, B3H 3Z1 Canada*
[4]*UPMC Univ Paris 06, UMR 7621, LOMIC, Observatoire Océanologique, F-66651 Banyuls/mer, France*
[5]*CNRS, UMR 7621, LOMIC, Observatoire Océanologique, F-66651 Banyuls/mer, France*

(Received 26 May 2009; revised 21 September 2009; accepted 2 March 2010)

Downloaded By: [Stiftung] At: 14:20 6 September 2010

The widely distributed prymnesiophyte species *Chrysochromulina polylepis* is prominent and well known for occasional formation of ichthyotoxic blooms. The chemical structure of the *C. polylepis* toxin(s) has not yet been elucidated, but the associated haemolytic activity, potent membrane disruption interactions and toxicity to finfish and protists have led to the suggestion that they may be similar to the prymnesins of *Prymnesium parvum*. Such polyether toxins are presumably formed partially or completely via polyketide biosynthetic pathways. In this genetic study of *C. polylepis*, we generated and analysed a genomic DNA and a normalized cDNA library. We estimated a genome size of approximately 230 mbp based upon analysis of >1000 genomic library clones. Of the cDNA library, 3839 clones were partially sequenced and annotated, representing approximately 2900 unique contigs. We detected several genes putatively related to toxin synthesis. Thirteen putative polyketide synthase (PKS)-related gene sequences were identified and phylogenetic analysis identified two of these as containing ketoacyl domains of the modular type I PKS. Semi-quantitative reverse-transcription polymerase chain reaction (RT-PCR) was used to follow the expression of PKS genes over the light/dark cycle of synchronized *C. polylepis* cultures. This is the first study showing the expression of PKS genes in marine microalgae, in this case in the toxigenic *C. polylepis*.

Key words: cell cycle, *Chrysochromulina*, expressed sequence tags (ESTs), gene expression, genomic characterization, ichthyotoxins, polyketide synthase (PKS), prymnesiophyte

Introduction

Harmful algal blooms (HABs) occur worldwide, with an apparent increase in frequency, intensity, environmental impact and geographic distribution in recent years (Hallegraeff, 2003). Among HAB taxa, the widely distributed prymnesiophyte genus *Chrysochromulina* from the class Prymnesiophyceae within the algal division Haptophyta (Jordan & Green, 1994; Edvardsen *et al*., 2000) is prominent and well known for occasional formation of ichthyotoxic blooms. The genus *Chrysochromulina* has close morphological and molecular phylogenetic relationships with the ichthyotoxic prymnesiophycean *Prymnesium* and shares the capacity to form HABs in coastal and brackish waters (reviewed in Edvardsen, 1996). Although several of over 50 described *Chrysochromulina* species are reported to be potentially toxic, most attention has been focussed on *C. polylepis* Manton et Parke, the source of a devastating toxic bloom that occurred in the Kattegat and Skagerrak region near the Norwegian coast in the late 1980s (Dahl *et al*., 1989, 2005; Granéli *et al*., 1993). This bloom resulted in extensive fish kills and caused severe ecological damage to wild biota with high economic losses at fish farms along the Norwegian and Swedish coasts (Nielsen *et al*., 1990; Skjoldal & Dundas, 1991; Gjøsæter *et al*., 2000).

The toxicity of *Chrysochromulina* in culture is highly species- and even strain-specific and has only been demonstrated in bioassays. For *C. polylepis*, there is evidence of allelochemical effects against other microalgae (Myklestad, 1995;

Correspondence to: Uwe John. E-mail: uwe.john@awi.de

ISSN 0967-0262 print/ISSN 1469-4433 online/10/00000215–229 © 2010 British Phycological Society
DOI: 10.1080/09670261003746193

Downloaded By: [Stiftung] At: 14:20 6 September 2010

Schmidt & Hansen, 2001), as well as related grazing inhibition (John *et al*., 2002). Certain zooplankton exposed to *C. polylepis* experienced reduced rates of growth and reproduction and enhanced mortalities (Nielson *et al*., 1990).

The mechanisms involved in the expression of toxicity in *C. polylepis* are still poorly understood. Several reports have indicated that the expression of toxicity (calculated as change of toxicity per cell) may be triggered by external factors, such as phosphorus deficiency (Edvardsen *et al*., 1990; Johansson & Graneli, 1999) and increased salinity (Edvardsen *et al*., 1996) but this could also be attributable to indirect effects on growth rate. However, maximal cell toxicity has been claimed to occur during exponential growth (Edvardsen *et al*., 1996; Schmidt & Hansen, 2001). Cell cycle analysis of synchronized cultures of *C. polylepis* showed that toxicity of cell extracts was discontinuous during the cell cycle, with maximal toxicity occurring during the light phase of the photoperiod (Eschbach *et al*., 2005).

The mode of action of the *C. polylepis* toxin(s) is apparently non-selective, but is associated with haemolytic activity, potent membrane disruption and toxicity to finfish and protists (Skjoldal & Dundas, 1991; Edvardsen *et al*., 1996; Gjøsæter *et al*., 2000). The observations of similar ichthyotoxicity and cytolytic effects on cells and tissues caused by exposure to other prymnesiophytes, such as *P. parvum*, led to speculations that they may be caused by similar toxins (John *et al*., 2002). Among the prymnesiophytes, the prymnesins (PRM 1 and PRM 2) isolated from *P. parvum* are the only structurally described toxins (Igarashi *et al*., 1998). The prymnesins are potent haemolytic components and are ichthyotoxic, particularly with Ca^{++} ion as supplement, but it is not clear whether or not prymnesins alone can account for all toxic and allelochemical effects because non-toxic species can also release compounds that can lyse cells (Eschbach *et al*., 2005).

Prymnesins are mixed linear polyether compounds and thus share some structural similarity to the polyether toxins produced by marine dinoflagellates, although the mode of action may not be analogous (Wright & Cembella, 1998). Based upon the results of stable isotope labelling studies of other polyether toxins, such as spirolides (MacKinnon *et al*., 2006), we assume that prymnesins are derived by similar polyketide biosynthetic pathways.

Organisms known to produce polyketides *via* polyketide synthases (PKSs) include bacteria, fungi, sponges, microalgae and higher plants (e.g. Proksch *et al*., 2002; Dittmann & Wiegand, 2006; John *et al*., 2008). A common element of all polyketide biosynthesis is the strictly ordered regulated generation by PKS (Staunton & Weissman, 2001; Moore & Hertweck, 2002; Cembella & John, 2006; John *et al*., 2008). The PKSs are large multi-domain enzymes or enzyme complexes closely related to fatty acid synthases (FASs). They are composed of the identical ancestral set of functional modules: ketoacyl synthase (KS), acyl transferase (AT), ketoacyl reductase (KR), dehydratase (DH), enoyl reductase (ER), acyl carrier protein (ACP) [or phosphopantetheine attachment site (PP)], and thioesterase (TE) domains. Whereas all units are needed for fatty acid production by FASs, the minimal structure of PKSs requires only ACP, KS and AT for the condensation reaction of acetate units. The other domains, if present, can catalyse the stepwise reduction of the initial carbonyl units (for details see Staunton & Weissman, 2001; Moore & Hertweck, 2002; Kusebauch *et al*., 2009).

The PKSs are involved not only in phycotoxin biosynthesis, but also in synthesis of other compounds with a diverse spectrum of functions in nature, ranging from chemical defence to complex cell-to-cell communication (Borejsza-Wysocki & Hrazdina, 1996; Ikeda & Omura, 1997; Börner & Dittmann, 2005). Starter units of PKSs can be short-chain (branched) fatty acids, different alicyclic and aromatic acids, and amino acids. Additionally, post-PKS tailoring events such as glycosylation, acylation, alkylation and oxidation further add to polyketide structural and functional diversity (Moore & Herweck, 2002).

Degenerate primer sequences and heterologous probes for PKS genes for fungi and bacteria have been published and can be effectively used for the isolation of gene sequences in these organisms (Lee *et al*., 2001, Nicholson *et al*. 2001; Ayuso-Sacido & Genilloud, 2005; Schumann & Hertweck, 2006). Yet these primers are specifically designed for certain phylogenetic groups and are not generally usable on a wider range of species. Thus, the isolation of PKS genes from species for which only limited genome information is available is still a challenge.

No whole genome sequences are available for the target genus *Chrysochromulina*, therefore we adopted a limited genomic approach based upon expressed sequence tag (EST) surveys targeting transcribed coding regions. ESTs are partial cDNA sequences that derive from single-pass sequencing. They provide a rich source of information that has been used for the identification of novel genes, gene mapping, comparative genomics and functional characterization of gene products (e.g. Rudd, 2003). After construction and sequencing of a cDNA library, an expressed sequence tag (EST) database was analysed, compared with other Prymnesiophyceae, and screened for the presence of putative PKS genes.

The aim of the present work was to enhance the genetic and molecular characterization of *C. polylepis* and prymnesiophytes in general. The focus was to: (i) obtain insights into the genomic structure of this member of the Prymnesiophyceae, (ii) identify putative PKS genes that might be involved in synthesis of toxins, and (iii) determine whether or not the expression of PKS genes is regulated in a cell cycle-dependent manner, in correlation with the toxicity. We demonstrated for the first time the expression of PKS genes in prymnesiophytes and provided one of the most detailed genomic studies on an ichthyotoxic species to date.

Materials and methods

Synchronization conditions and sampling procedure

Experiments were conducted with a toxic, haploid strain (B1511) of *Chrysochromulina polylepis* Manton & Parke. The strain was isolated by Bente Edvardsen, University of Oslo, from the Oslofjord, Norway (59°00′N; 10°45′E). *Chrysochromulina polylepis* was grown in enriched seawater medium IMR 1/2 (Eppley *et al.*, 1967) supplemented with 10 nM selenite in batch cultures at 15°C under cool-white fluorescent light with a photon flux density of 45 μmol s^{-1} m^{-2}, applied over a 14-h:10-h light–dark regime. Unialgal cultures in exponential growth phase were scaled up for the synchronization experiments by inoculation of *ca.* 1×10^4 cells ml^{-1} in sequence into 0.05, 0.5, 1, 5 and finally into 10-l flasks to reach a final concentration of 1.5×10^5 cells ml^{-1} before each transfer. Cultures of 5 and 10 l were gently aerated with sterile-filtered air to provide CO_2 and to achieve a homogeneous cell distribution. For the synchronization experiments, three parallel cultures were sampled during exponential growth over 24 h at 2-h intervals, at a starting cell concentration of about 4×10^4 cells ml^{-1}. Samples were collected from 10-l cultures via a silicone rubber tube with an inner diameter of 3 mm by gently applying a vacuum created by drawing on a 50-ml syringe. Samples were immediately stored on ice and, after determination of the cell numbers, processed for RNA extraction. Sample collection during the dark period was under a red darkroom safety light.

Determination of cell concentration

Culture samples (2 ml) were diluted in 18 ml sterile-filtered seawater pre-cooled to 15°C. Cells were counted with a Multisizer II particle counter (Coulter Electronics, Krefeld, Germany) equipped with a 100 μm aperture, within a size window of 5 to 12 μm, which excluded most background particles. Average cell size with standard deviation (SD) was calculated with the Coulter Multisizer Software. Calculation of specific growth rate (μ, unit per day) was performed using the formula:

$$\mu(d^{-1}) = (\ln(C_1) - \ln(C_0))/1$$

where C_1 is the cell concentration at time 1 and C_0 is the cell concentration at time 0.

Cell cycle analysis

Samples (20 ml) of *C. polylepis* culture were fixed with 0.25% glutaraldehyde, stained with 5 μM SytoxGreen (Molecular Probes, Leyden, The Netherlands) and subsequently analysed for relative DNA content using a FACS Vantage flow cytometer (Becton-Dickinson, San Jose, California) equipped with an Innova Enterprise II 621 laser. This procedure, including analysis of at least 1×10^4 cells per sample at 1 psi, was essentially as previously described (Eschbach *et al.*, 2001). Dot plots and histograms were created with the WinMDI 2.8 software (Joseph Trotter, Scripps Research Institute, La Jolla, California). Cell-cycle analysis was achieved with the Multicycle software (Phoenix Flow Systems, San Diego, California). The number of cells in a certain cell-cycle phase was expressed as the percentage of the total cell number in the sample. Duration of single cell-cycle phase was determined using an algorithm for synchronized cell cultures (Beck, 1978).

RNA extraction

Samples of $10–15 \times 10^6$ cells taken at 2-h intervals during a 24-h kinetic study were harvested by centrifugation at $5000 \times g$ for 15 min at 4°C. The cell pellets were resuspended in 500 μl RLT buffer (Qiagen, Hilden, Germany) and immediately frozen in liquid nitrogen. The resuspended pellets were stored at –80°C until use. Before RNA extraction the cells were mechanically disrupted with a Tissue Lyser (Qiagen, Hilden, Germany) for 30 s, at 30 1/s frequency. Homogenates were processed according to the manufacturer's instructions (RNeasy Plant Mini kit, Qiagen, Hilden, Germany) with few modifications. Briefly, the samples were loaded on a QIAshredder column (Qiagen, Hilden, Germany), and centrifuged for 10 min at 8000 *g*. A second purification step with the Qiagen RNeasy Cleanup kit was conducted, including the on-column DNA digestion with RNAse-free DNAse (Qiagen, Hilden, Germany). The extracted total RNA was checked for integrity by gel electrophoresis and quantified by spectrophotometry. The extracted RNA was stored at –80°C until use.

cDNA library generation

The cDNA synthesis was generated by Vertis biotechnology (Freising-Weihenstephan, Germany) from 2.4 μg total RNA. In brief, oligonucleotide primers were attached to the 5′- and 3′-ends of the cDNA to allow PCR-amplification and directional cloning of the cDNA as well. The *Not*I/*Asc*I-sites of the plasmid vector pFDX3840 (supplied by Prof. Dr Ralf Reski, Freiburg) were used for directional cloning. All PCR amplification steps were performed with a long and accurate (LA) PCR system as described by Baskaran *et al.* (1996).

Downloaded By: [Stiftung] At: 14:20 6 September 2010

Downloaded By: [Stiftung] At: 14:20 6 September 2010

Normalization of the cDNA was performed according to Ko (1990) with several modifications. Specifically, the cDNA for normalization was not sheared but rather was used full length. Normalization was achieved by two consecutive cycles of denaturation and reassociation of the cDNA, resulting in N1- and N2-cDNA. With the N2-cDNA, a Cot-value of approximately 90 was achieved. Reassociated double stranded-cDNA was separated from the remaining single stranded-cDNA (normalized cDNA) by passing the mixture over a hydroxylapatite column (Ausubel *et al.*, 1987). The hydroxylapatite-purified single-stranded (ss)-cDNA was amplified with 12 (N1) and 14 (N2) LA-PCR cycles and size-fractionated on 1% agarose gels. The fractions >750 bp were cut out from the gel and isolated by electroelution. The eluted fragments were cloned directionally into the vector pFDX3840 bearing ampicillin antibiotic resistance. The plasmids were transformed into *E. coli* TOP 10F (Invitrogen, Karlsruhe, Germany).

EST sequencing

A total of 3839 clones from a normalized *C. polylepis* cDNA library were picked using a Q-pix (Genetix, Hampshire, UK) colony picker. Following overnight growth of picked clones at 37°C in LB medium containing 50 $\mu g\,ml^{-1}$ ampicillin, stocks were made by adding 7.5% glycerol. The stocks were frozen immediately in liquid nitrogen and stored at –80°C. The template for sequencing was generated directly from glycerol-stock cells with the Templiphi technology (GE Healthcare, Sunnyvale, CA, USA). Sequencing reactions were performed with ET-terminator chemistry (GE Healthcare, Sunnyvale, CA, USA) and M13 forward or M13 reverse primer (Table 1). Cleaned sequencing products were analysed on MegaBACE 1000 or 4000 sequencers (GE Healthcare, Sunnyvale, CA, USA).

EST annotation

The scf traces generated by the MegaBACE 1000/4000 were processed with the pre-gap and gap modules in the Staden software package (http://staden.sourceforge.net/). Vector sequences were removed and the traces were clipped for quality using pre-gap. The Gap4 module of the Staden package was used to assemble overlapping reads and generate sequence files in fasta format. The fasta files were blasted using blastx (Altschul *et al.*, 1997) against the Genbank non-redundant protein database (nr) with an e-value threshold of 0.001. Contigs that did not report a significant hit were blasted using PSI-BLAST against the SwissProt (Bairoch *et al.*, 2005) database. Expectation value threshold was set at 0.1 and three rounds were allowed.

Table 1. Primer sequences for the semi-quantitative PCR procedure.

Primer name	Sequence 5′–3′
M13 forward	GTT TTC CCA GTC ACG ACG TTG
M13 reverse	TGA GCG GAT AAC AAT TTC ACA CAG
GAPDH forward	TCA ACG ACG CCA AAT ACA ATG
GAPDH reverse	ACC CTT CGT GAT GCC GTA GT
Cytochrome forward	ATG GCC ACC ACG AAA TCC T
Cytochrome reverse	ATA CCT CGC CTC TGA ATG CAA
PKS7 forward	GGT GTT CAA GCT GCT GAT GC
PKS7 reverse	TGC CTG CAT ACC CAA ATG AG

ESTs with e-values below 0.001 were examined manually; those with an identity of 30–50% over a region of similarity of at least 50 amino acids were analysed further in the SwissProt and non-redundant (nr) databases of NCBI – BLASTX. The sequences were grouped according to their functional categories (http://www.ncbi.nlm.nih.gov/COG/, http://www.ebi.ac.uk/interpro/).

Blast analysis

The contigs from the assembly were compared to custom made and publicly available databases. We retrieved the coding sequences (CDS) of *Arabidopsis thaliana* from TAIR (http://www.arabidopsis.org/), the *Emiliania huxleyi* CDS from the JGI website (http://www.jgi.doe.gov/), and the CDS of *Thalassiosira pseudonana*, *Phaeodactylum cornutum*, *Physcomitrella patens*, and *Cyanidioschyzon merolae* from the respective NCBI web sites (http://www.ncbi.nlm.nih.gov/) in August 2009. Additionally, we used the databases Swissprot (swissprot version 54) and the complete refseq set from the NCBI database (version 33). Comparisons between EST sequences were done as tblastx searches; the protein databases were searched using blastx.

Blast parameters used (WUBlast2.0):

B 120	maximum number of database sequences for which *any* alignments will be reported
E 1E-15	the expectation threshold for reporting database hits
W 11	seed word length for the ungapped blast algorithm.

Genomic library

DNA of *C. polylepis* was extracted using the plant DNA extraction kit following the instructions of the manufacturer (Qiagen, Hilden, Germany). Genomic libraries with a target insert size of 1.5 kb were constructed from total DNA as described previously (Glöckner *et al.*, 2004). Colonies containing bacterial vectors with cDNA inserts were grown in LB medium. The Qiagen magnetic bead protocol was used for the plasmid preparation.

Custom primers and the cycle sequencing method were used for sequencing. The sequencing reaction products were separated on ABI3700 96 capillary machines.

Quality clipping was done with phred (Ewing & Green, 1998) and vector removal with phrap (http://www.phrap.org).

Downloaded By: [Stiftung] At: 14:20 6 September 2010

Genomic data analysis

The phrap (http://www.phrap.org/) assembler enabled us to obtain a minimal contig set, mainly by joining paired end reads. Previous results have shown that an automated assembly alone does not join all possible overlaps caused by accumulating sequencing errors at the end region of the sequences. This is particularly the case for contigs with low coverage. Thus, a final manual curation was done on the automatically generated contig set.

Additionally to the above-mentioned global Blast databases swissprot and refseq we used organism-specific databases to pinpoint the potential phylogenetic distribution of blast matches. Databases were generated using predicted proteins from *Cyanidioschizon*, *Chlamydomonas*, *Physcomitrella*, and *Arabidopsis*. Blast hits occurring against bacterial proteins only were assumed to be derived from culture contaminations. However, in many eukaryote genomes 'bacterial' genes have been identified, which might result from recent lateral gene transfer. Potential chloroplast- and mitochondria-specific genes were assigned according to their respective best hit to organelle-specific genes.

We estimate that most bacteria-derived clones could be detected, because a very diverse set of bacterial genomes has already been completely determined. However, we cannot rule out additional culture contaminations (viruses, other eukaryotes, etc.) because available data are scarce.

Genome size estimation was done under the primary assumption that protist genomes generally code for around 10 000 protein-coding genes (Armbrust *et al*., 2004; Derelle *et al*., 2006). The average 'gene space' was determined by dividing the number of sequenced prymnesiophyte nuclear contigs by the number of genes found. This number was then multiplied by the expected number of genes corrected for potential unidentified unknown genes. This unidentifiable, species-specific number of genes has been estimated to be about 40% (e.g. Eichinger *et al*., 2005; Marsden *et al*., 2006), if no genome of the same taxon is completely determined. Species-specific gene family extensions would lead to a slight underestimation of genome size. Thus, the number given is the lower threshold value.

Analysis of candidates

In total 13 putative PKS sequences were identified, based on similarity to known PKS genes among the *C. polylepis* EST dataset. These sequences were further analysed with the NRPS-PKS software tools (Ansari *et al*., 2004) for the prediction of the domain organization. For most putative PKS ESTs, a ClustalX (Thompson *et al*., 1997) alignment was generated using sequences from the BLASTX database. However, the KS domain of PKS4 (930 bp) and PKS7 (1509 bp) was analysed with the alignment from John *et al*. (2008). The resulting sequence dataset was aligned with Kalign (Lassmann & Sonnhammer, 2005) and PHYML (Guindon & Gascuel, 2003; see John *et al*., 2008) was used for Maximum Likelihood phylogenetic analysis with 1000 bootstrap runs. GAPDH and cytochrome f were also identified and analysed as described herein and served as controls in the gene expression studies.

Gene expression

In the semi-quantitative RT-PCR approach the genes PKS7, GAPDH, and cytochrome f were analysed. PCR primers were designed with Primer3 software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) (Table 1). To obtain reliable data using the semi-quantitative RT-PCR approach it is crucial to identify the proper amount of total RNA used for reverse transcription, the dilution factor, the amount of cDNA template for the PCR reaction and the optimal number of PCR cycles. Therefore a step-by-step optimization was performed before running the final set of samples. Finally, reverse transcription (RT) was performed with the Omniscript RT-PCR Kit (Qiagen, Hilden, Germany) with some modifications. One µg of DNAse-treated (Qiagen, Hilden, Germany) total RNA was reverse-transcribed at 42°C, with 50 pmol oligo-VNdT_{18} primer and 1 mM of dATP, dGTP, dCTP and dTTP each in a total volume of 20 µl. After reverse transcription the 20 µl reaction was diluted 1:5 with low TE buffer (10 mmol Tris-HCL, 10 mM EDTA pH 8). A 2 µl aliquot was used for PCR amplification by the Hot-MasterTaq procedure (Eppendorf, Hamburg, Germany) with 10 mmol dNTPs. The PCR cycle conditions were 1 min at 94°C denaturation and different numbers of cycles consisting of 30 s at 94°C denaturation, 30 s at 60°C annealing, and 45 s at 72°C elongation. As the final settings, 30 cycles for PKS7, 24 cycles for GAPDH, and 20 cycles for cytochrome f were run. For the semi-quantitative PCR reaction, equal aliquots of each PCR reaction (10 µl) were separated on 2% agarose gels using Tris–borate buffer containing 1.0 M Tris, 0.9 M boric acid and 0.01 M EDTA and photographed after ethidium bromide staining. Gels were analysed with 1D Image analysis software (Kodak Digital Science, Jahnsdorf, Germany). Analyses were made semi-quantitatively on the basis of net band intensities.

Results

Genome characteristics

In total, 1056 clones containing genomic DNA fragments were sequenced, yielding 781.5 kb of assembled data (from 950 kb raw data). Of these, 761 952 bp of genomic DNA were counted as of *C. polylepis* origin. However, 0.83% of this sequence data appears to belong to the chloroplast genome and 0.08% to the mitochondrial genome, leaving 754 991 bp, which are part of the *C. polylepis* nuclear genome. Using Blast we were able to identify putative genes on 3.26% of the sequence contigs, resulting in a gene space (basepairs/identifiable gene) of 45 580 bp and a minimum estimate of genome size of approximately 230 MB. Also 2.5% potential bacterial and 0.59% viral potential

contamination was identified (all summarized in Table 2).

EST database

From the normalized cDNA library of clones that were 5′-end sequenced, 2927 sequences passed the quality control as successful reads. Assembly of the resulting ESTs yielded 476 contigs and 1724 singletons. The sequences reached an average length of 684 bp and exhibited an average GC content of 57.27%. Highly significant matches were most frequently obtained with sequences from unicellular eukaryotes, animals, fungi, and plants. However, significant matches to sequences from prokaryotes were also observed.

Grouping of the ESTs according to the COG categories for putative cellular function assigned about 75% (1679) of the ESTs into the category 'poorly characterized', which includes the genes with no hit (1526 sequences) and genes with unknown functions (153 sequences) (Fig. 1). The relatively high number of ESTs without assignable function may be attributable to the fact that only partial sequence information was available for *Chrysochromulina* and only a few prymnesiophyte genome sequences are represented in gene data banks. Furthermore, some EST sequences may be derived from non-coding transcribed parts of the genes, thereby obscuring their function. The percentage distribution of sequences falling into different functional categories (521 genes) is plotted in Fig. 2: 10% (224) of the sequences encode proteins involved in metabolism, 9% (191) are involved in cellular processes, 4% (98) are related to information storage and processing and <1% (15) are predicted to encode cell structural proteins. Furthermore, 1% (27) of the represented sequences encoded proteins that are apparently involved in cell defence and toxicity. Table 3 and Fig. 2 show a detailed view of the contig distribution among the functional categories.

Table 2. Genome survey of sequencing results for *Chrysochromulina polylepis*.

Bacterial contamination	2.52%
Net bp *Chrysochromulina* (including chloroplast + mitochondria)	761 952
Chloroplast	0.83%
Mitochondrial	0.08%
Transposon	1.01%
Virus	0.59%
Gene hit	3.26%
Nuclear genome (bp)	754 991
Nuclear hits	3.29%
Gene space (bp/identifiable gene)	45 580 bp
Genome size (MB; 5000 identifiable genes)	227.9

EST library comparison

We performed BLAST-based comparisons to the two EST sets of *P. parvum* and *Isochrysis galbana*, respectively, together with the whole coding potential of *Emiliana huxleyi*. Approximately 100 genes shared between *C. polylepis* and *P. parvum* are not present in the non-toxigenic *I. galbana* data set (see Fig. 3*a*). Depending on the threshold used, 781 (threshold 100) or 1248 (threshold 120) EST contigs had no counterpart in other genomes. We then analysed which part of the data set is common among only distantly related species. First we compared our data set to complete genomes of diatoms (*Thalassiosira pseudonana* and *Phaeodactylum tricornutum*) and to proteins from higher plants. This analysis revealed that a considerable fraction of the shared genes is present in higher plants and not in diatoms despite a comparable evolutionary history (secondary endosymbiosis; Fig. 3*b*). Furthermore, 937 contigs have similarities to genes in other genomes as we found by comparing them to the complete NCBI protein reference set (not shown). A third comparison to individual

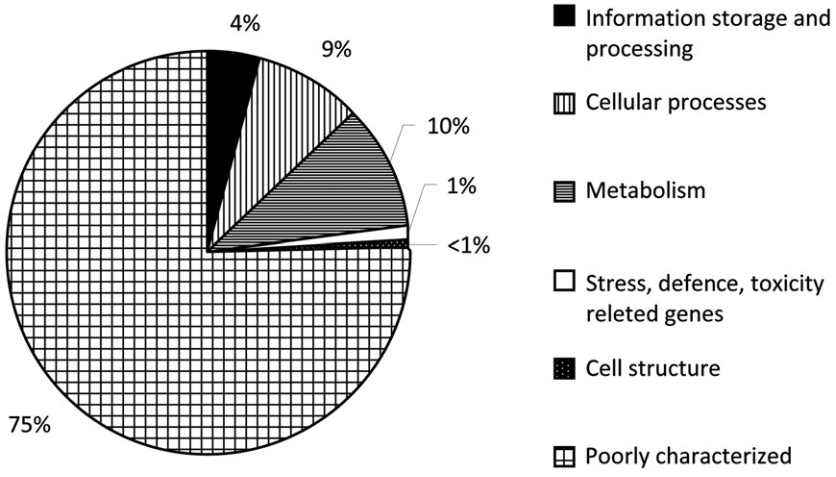

Fig. 1. Functional characterization of 2200 annotated cDNA contig sequences of *Chrysochromulina polylepis*.

Downloaded By: [Stiftung] At: 14:20 6 September 2010

Downloaded By: [Stiftung] At: 14:20 6 September 2010

complete genomes of photosynthetic species (the alga *Cyanidioschyzon merolae*, the moss *Physcomitrella patens*, and *Arabidopsis thaliana*) confirms that most of the detectable similarities of our data set are found with land plants and not to unrelated algal species (Fig. 3*c*).

Kinetics of synchronous growth

Synchronous growth of *C. polylepis* clone B1511 was achieved after reaching early exponential phase within 24 h in the 10-l cultures (Fig. 4*a*). Maximum cell concentration ($\sim 1.4 \times 10^5$ cells ml^{-1}) was attained within 6 days after inoculation. *Chrysochromulina polylepis* cultures exhibited a stepwise increase of cell number during the 24 h sampling period on Day 4 (Fig. 4*b*). The cell number increased from the middle of dark period until early light period. The mean growth rate of *C. polylepis* over the 6 days was $\mu = 0.57$ and was similar over the 24 h sampling period ($\mu = 0.53$).

Cell cycle analysis

Flow cytometric determination of the relative DNA content of *C. polylepis* nuclei revealed successive cell-cycle phases typical of eukaryotic cells (Fig. 5). Single distinct peaks for G1, S and G2+M phases were obtained, respectively. The G2 and M phases cannot be resolved by flow cytometry since cells in these two phases contain the same amount of DNA. As expected, *C. polylepis* cell division proceeded through the typical transitions of the eukaryotic cell cycle: DNA synthesis (S) began 2 hours before the beginning of the dark period and was completed 2 hours before the end of the dark period (Fig. 5). Cell division (G2+M) started and was completed during the dark period. An increase in the number of cells in S and G2+M phases was always synchronous with a decrease in the number of cells in G1 phase.

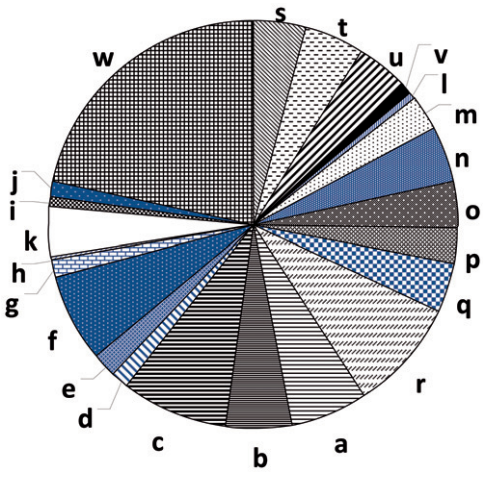

Fig. 2. Distribution of contigs of the normalized cDNA library by eukaryotic categories of orthologous groups (COG) classes; see Table 3 for meanings of letter codes.

PKS candidate genes and gene expression

A total of 13 potential PKS sequences were identified from the EST dataset. Analysis with the NRPS-PKS software identified seven fragments (PKS1, PKS2, PKS3, PKS4, PKS5, PKS6, PKS7) as parts of a KS domain, three fragments (PKS8, PKS9, PKS10) that code for a KR domain, and three fragments (PKS11, PKS12, PKS13) as fragments from an AT domain. To prove their origin the sequences were compared with known PKS sequences of several organisms. Confirmation of their domain function (data not shown) was achieved *via* analysis with the PhyloGena software (Hanekamp *et al.*, 2007) for most ESTs of potential PKS. In particular, the PKS4 and PKS7 fragments

Table 3. Eukaryotic orthologous gene group (COG) classes and respective numbers of *Chrysochromulina polylepis* unigenes.

Class	Metabolism	N° of contigs
a	Carbohydrate metabolism	43
b	Energy production and conversion	37
c	Amino acid transport and metabolism	59
d	Nucleotide transport and metabolism	11
e	Coenzyme transport and metabolism	14
f	Lipid metabolism	48
g	Secondary metabolites biosynthesis, transport and catabolism	10
h	Intracellular trafficking	2
	ESTs in total	**224**
	Cell structure	
i	Extracellular structures	6
j	Cytoskeleton	9
	ESTs in total	**15**
	Stress, defence, and toxicity	
k	Stress, defence and toxicity related genes	27
	ESTs in total	**27**
	Cellular processes	
l	Chromatin structure and dynamics	5
m	Cell division and chromosome partitioning	18
n	Post-translational modification, protein turnover, chaperones	33
o	Cell wall, membrane, envelope biogenesis	24
p	Cell motility	21
q	Inorganic ion transport and metabolism	28
r	Signal transduction	62
	ESTs in total	**191**
	Information storage and processing	
s	Transcription	30
t	Translation, ribosomal structure and biogenesis	33
u	DNA replication, recombination and repair	30
v	RNA processing	5
	ESTs in total	**98**
	Poorly characterized	
w	General function prediction only	153*

*Part of the total of 1679 ESTs with unknown function.
Abbreviation: EST, expressed sequence tag.

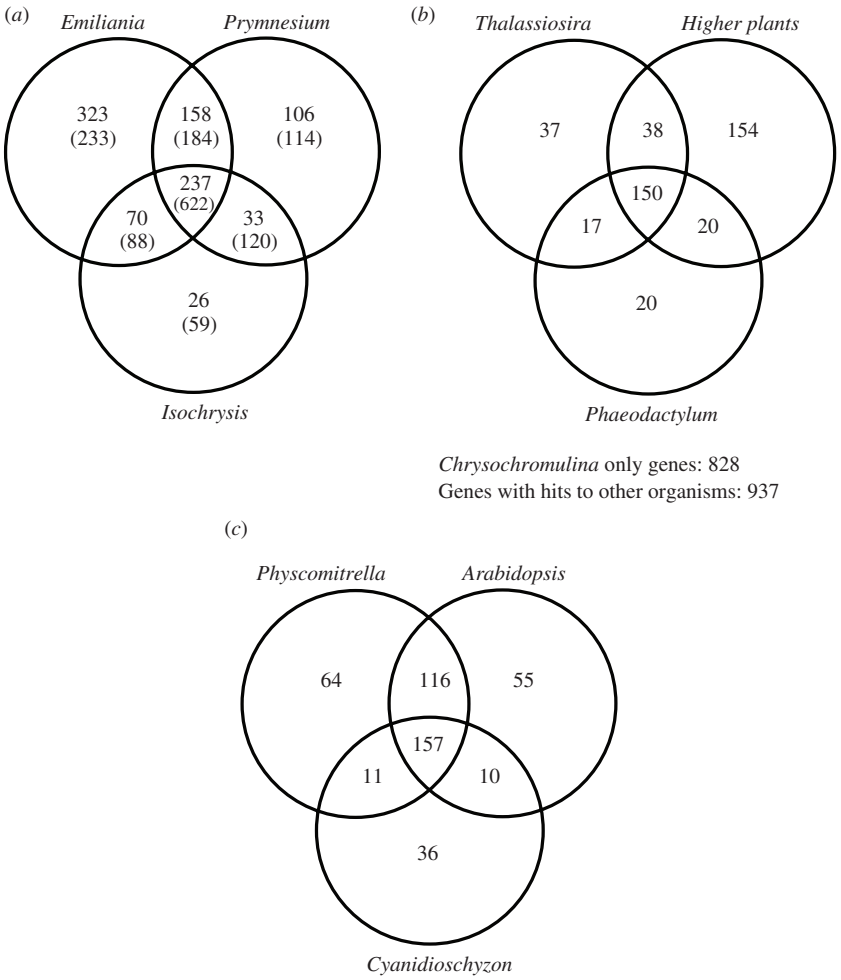

Fig. 3. Venn diagrams depicting matches of *Chrysochromulina polylepis* ESTs to different databases. (*a*) Matches to *Prymnesium parvum*, *Emiliana huxleyi*, and *Isochrysis galbana* with score thresholds of 120 and 100 (in parentheses). (*b*) Matches (score threshold 100) to diatoms *Thalassiosira pseudonana*, *Phaeodactylum tricornutum*, and higher plants. (*c*) Matches (score threshold 100) to *Physcomitrella patens*, *Arabidopsis thaliana*, and *Cyanidioschyzon merolae*.

Downloaded By: [Stiftung] At: 14:20 6 September 2010

were identified as nearly full length KS domains. The phylogeny of these sequences was analysed with the data set of John *et al.* (2008) containing all available eukaryotic (PKS type I) KS sequences. This analysis showed that PKS4 and PKS7 fall into the prymnesiophyte PKS gene clade. This topology is well supported by bootstrap analysis among PKS sequences of the prymnesiophyte *Emiliania huxleyi* (Fig. 6).

We chose PKS7 and the control genes GAPDH and cytochrome f for the gene expression studies, the latter two identified based on similarity among the ESTs (Table 4 panel *a* and *b*). After demonstrating that PKS7 was a PKS fragment of *C. polylepis* origin, its expression profile over 24 h was determined by semi-quantitative PCR (Fig. 7 and Table 4 panel *c*). Whereas GAPDH was mainly expressed during the light period (Fig. 7*b*), on the contrary PKS7 exhibited stronger expression towards the end of the dark period (Fig. 7*a*). The expression of cytochome f remained constant over the light/dark cycle of the synchronized culture (data not shown).

Discussion

Genome characterization

Modern high-throughput sequencing projects have yielded a large number of fully sequenced genomes of organisms ranging from prokaryotes to humans, providing many new exciting insights into phylogenetic relationships and genetic diversity. A few whole genome projects on eukaryotic unicellular organisms (protists) have also been completed, e.g. for the prasinophyte *Ostreococcus tauri* (Derelle *et al.*, 2006) and *Micromonas* (Worden *et al.*, 2009), the marine diatoms *Thalassosira pseudonana* (Armbrust *et al.*, 2004) and *Phaeodactylum tricornutum* (Bowler *et al.*, 2008), *Emiliania huxleyi* (by Betsy Read, conducted by the US Department of Energy Joint Genome Institute (JGI, www.jgi.

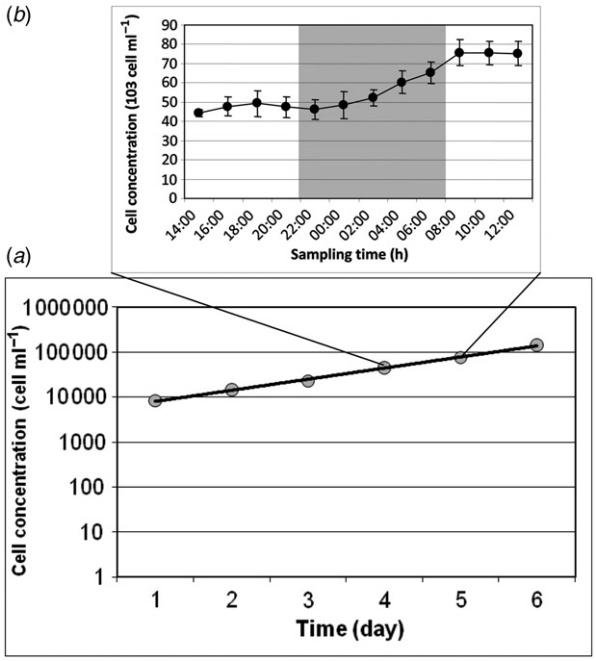

Fig. 4. *Chrysochromulina polylepis* cell concentrations. (*a*) Mean ± SD ($n = 3$) cell concentrations over a 6-day period in synchronized batch cultures. (*b*) Cell concentrations (mean ± SD; $n = 3$) over the sampling period of 24 h. Shaded area indicates dark phase over the 24-h light–dark cycle.

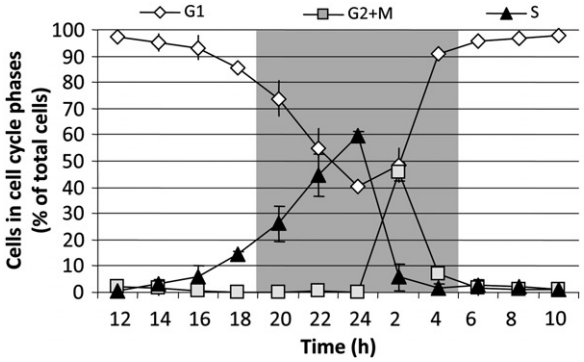

Fig. 5. DNA concentration from three synchronized *Chrysochromulina polylepis* batch cultures. Percentage of cells in Gap 1 (G1), DNA synthesis (S), and Gap 2 and mitosis (G2+M) phases of the cell cycle. Data points are means ± SD ($n = 3$). Shaded area: 10-h dark period.

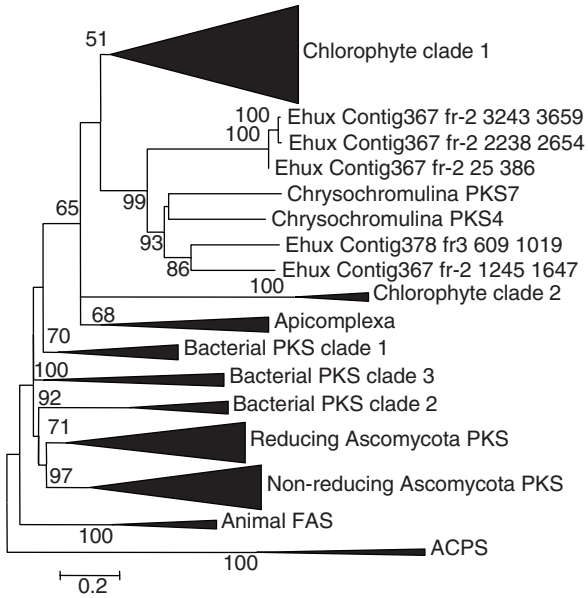

Fig. 6. Maximum likelihood phylogenetic analysis of the ketoacyl synthase (KS) domain from type 1 polyketide synthases. KS domains of *Chrysochromulina polylepis*, PKS4 and PKS7, were analysed with representative KS domains from most described clades. Corresponding taxon names can be taken from Kroken *et al.* (2003) and John *et al.* (2008). Numbers at the branches indicate bootstrap values. Scale bar represents corrected evolutionary divergence.

doe.gov) in collaboration with the user community), and the slime-mould *Dictyostelium discoideum* (Eichinger *et al.*, 2005). Other protist genome sequencing projects are in progress and several are close to completion. However, even with ever increasing sequencing capacities, not all organisms can be easily fully sequenced, because of limitations caused by large genome size and/or high DNA repetitiveness. A further constraint is that the size, the DNA content, and the organisational structure of the genome are unknown for most protists.

EST data are probably the most extensively produced genetic data at present, providing a huge dataset from phylogenetically and evolutionary diverse organisms (e.g. Rudd, 2003). Several EST datasets are available, mostly from diatoms and dinoflagellates, but the generation of many more is in progress, e.g. for *Alexandrium fundyense* (Hackett *et al.*, 2005), *Alexandrium ostenfeldii* (Jaeckisch *et al.*, 2008), *Phaeodactylum tricornutum* (Scala *et al.*, 2002), *Karenia brevis* (Lidie *et al.*, 2005), *Fragilariopsis cylindrus* (Mock *et al.*, 2006), *Emiliania huxleyi* (Wahlund *et al.*, 2004), *Isochrysis* (Pereira *et al.*, 2004), *Galdieria sulphuraria* (Weber *et al.*, 2004), and *Prymnesium parvum* (La Claire, 2006).

Knowledge about the genomic structure and organisation of photoautotroph protists in general is scarce, with the exception of species with completed genome projects [*Thalassiosira pseudonana* (Armbrust *et al.*, 2004), *Phaeodactylum tricornutum* (Bowler *et al.*, 2008), *Ostreococcus tauri* (Derelle *et al.*, 2006), *Micromonas* (Worden *et al.*, 2009), *Cyanidioschyzion merolae* (Matsuzaki *et al.*, 2004)]. The genomic data published for prymnesiophytes comprises the genome of *Emiliania huxleyi* (predicted to be 150–200 Mb) http://bioinfo.csusm.edu/Coccolithophorids/Emiliana-huxleyi/ http://genome.jgi-psf.org/ and EST libraries from

Downloaded By: [Stiftung] At: 14:20 6 September 2010

Table 4. Analysis for orthologues of glyceraldehyde-3-phosphate dehydrogenase (panel *a*), cytochrome f (panel *b*), and polyketide synthase (panel *c*) of *Chrysochromulina polylepis* deduced from amino acid sequences.

Panel *a*				
GAPDH	**Species**	**Score**	**e-value**	**Accession**
glyceraldehyde-3-phosphate dehydrogenase	*Zea mays*	304	3e-89	CAA33620.1
glyceraldehyde-3-phosphate dehydrogenase	*Phaeodactylum tricornutum*	302	1e-87	AAU81889.1
glyceraldehyde-3-phosphate dehydrogenase	*Odontella sinensis*	301	1e-88	AAU81890.1
Panel *b*				
cytochrome f	**Species**	**Score**	**e-value**	**Accession**
cytochrome f	*Emiliania huxleyi*	340	3e-92	YP_277330.1
cytochrome f	*Porphyra yezoensis*	283	5e-75	YP_536946.1
apocytochrome f	*Guillardia theta*	280	4e-74	AAC35685.1
Panel *c*				
PKS type I (PKS7)	**Species**	**Score**	**evalue**	**Accession**
polyketide synthase	*Microcystis aeruginosa*	276	1e-72	BAB12210.1
type I fatty acid synthase	*Cryptosporidium parvum*	275	3e-72	AAC99407.1
polyketide synthase	*Cryptosporidium parvum*	254	5e-66	AAN60755.1

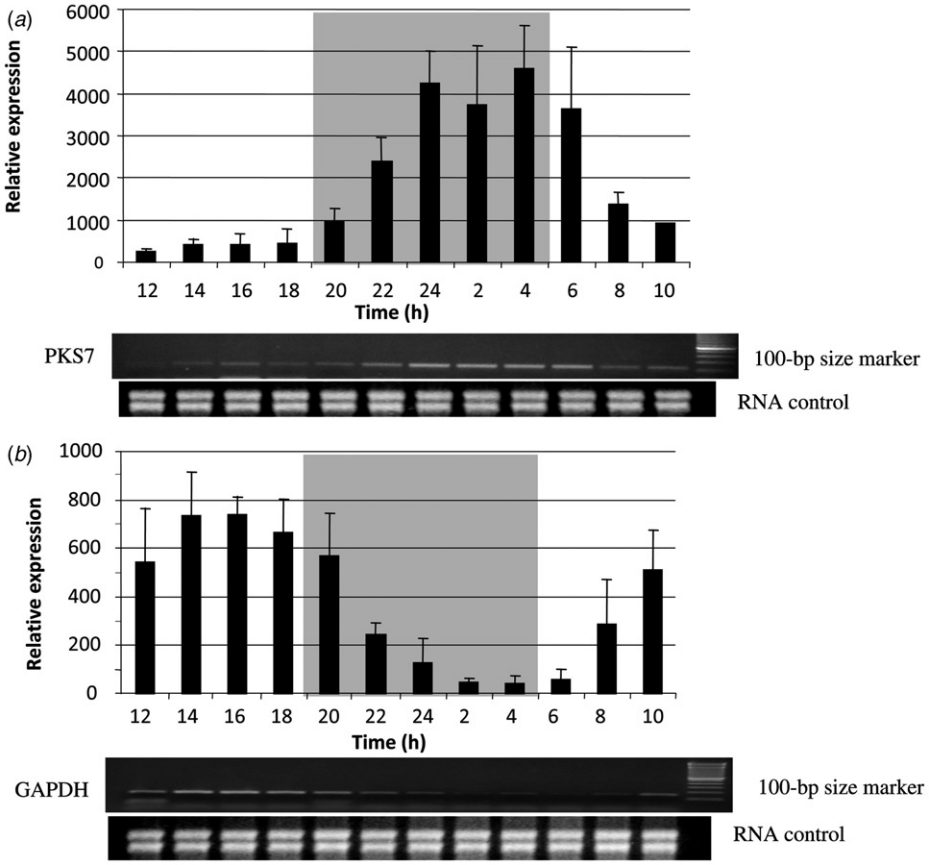

Fig. 7. *Chrysochromulina polylepis*. Mean ± SD ($n = 3$) of relative gene expression of *C. polylepis* in synchronized batch cultures for (*a*) PKS7 and (*b*) GAPDH. 18S and 28S rRNA bands used as control to demonstrate equal RNA concentration and quality. Relative gene expression was deduced from band intensities of semi-quantitative RT-PCR amplicons.

Isochrysis (Pereira *et al.*, 2004) and from *Prymnesium parvum* (La Claire, 2006). In this study the predicted genome size of about 230 Mb for *Chrysochromulina polylepis* is much larger than that of *Arabidopsis thaliana* (125 Mb), *Paramecium tetraurelia* (150 Mb), *Chlamydomonas reinhardii* (120 Mb), *Phaeodactylum tricornutum* (27 Mb), *Thalassiosira pseudonana* (32 Mb), *Ostreococcus tauri* (12 Mb), and *Cryptosporidium parvum* (9 Mb). Yet, it is many times smaller than genomes of *Spirogyra* (1969 Mb), *Euglena* (1300 Mb) and *Fucus* (529 Mb), making it a possible candidate for whole genome sequencing. Moreover, since the closest relative so far characterized (*E. huxleyi*)

Downloaded By: [Stiftung] At: 14:20 6 September 2010

Downloaded By: [Stiftung] At: 14:20 6 September 2010

has a genome in the same size range, a comparative genomics approach seems to be not only feasible but desirable. One possible reason for its relatively large genome size might be the heteromorphic haploid-diploid life cycle in the class Prymnesiophyceae, whereas such as life cycle has not been documented in its sister class Pavlovaphyceae, characterized by smaller genome size (Nosenko *et al*., 2007). Differences in genome size between the two classes are also mirrored in their plastid size (Sáez *et al*., 2001) and in the structure of the mitochondrial genome in the two classes (Sánchez-Puerta *et al*., 2004).

The cell size of microalgae has significant impacts on their ecological success (summarized in von Dassow *et al*., 2008). There is current debate about the correlation between cell size and genome size (Cavalier-Smith, 2005). This positive correlation seems to fit dinoflagellates (LaJeunesse *et al*., 2005), diatoms (von Dassow *et al*., 2008) and cryptomonads (Beaton & Cavalier-Smith, 1999). With a cell diameter of approximately 7 μm *C. polylepis* is smaller than or in the same size range as the two diatoms (*Thalassiosira pseudonana*, centric, 2.3–5.5 μm; *Phaeodactylum tricornutum*, Pennales, 3–25 μm) and *Chlamydomonas* (~10 μm) (Merchant *et al*., 2007). Thus, a positive correlation between genome size and cell size is not always observed. Here we compared the EST data sets from *Isochrysis*, *Prymnesium*, *Emiliania* and *Chrysochromulina* to identify potential ESTs common to these three Prymnesiophyceae. Potential applications of ESTs include reconstruction of phylogenetic relationships. In this approach, gene fragments are obtained via sequencing of randomly selected cDNA clones. Overlapping sequences are clustered, and orthologues to known genes are determined by means of sequence similarity. These sequences are then employed to construct phylogenetic trees. For instance Bapteste *et al*. (2002) used ESTs to study the phylogenetic affinities of amoebans and related taxa. However, we used this comparative approach to identify additional candidate genes that might be involved in the biosynthesis and/or regulation of toxins and their respective toxicity. We found that 303 ESTs were common to *C. polylepis* and *P. parvum*. Many of the genes involved had unknown function but included the PKS genes of both species. Comparative genomic approaches can help to elucidate candidate genes among a large set of ESTs, particularly in cases where the majority of genes are of unknown function but the biochemical pathways of interest are at least partially described. We compared only a limited EST dataset of three species, in future more ESTs per species and more species will increase significantly the potential success of our approach. Nevertheless, this is a good starting point for exploring non-model organisms and unknown biochemical pathways and their regulation.

In our genomic analysis we placed particular emphasis on the identification and characterization of putative polyketide synthase (PKS) in *Chrysochromulina* because of the relatively common occurrence of polyether toxins among toxigenic protists, particularly in dinoflagellates, which are likely derived *via* polyketide biosynthetic pathways. Recent studies have shown that apicomplexans, close relatives of the dinoflagellates, have the modular type I PKS (Zhu *et al*., 2002) and phylogenomic analysis indicates that this modular type is typical of protists (John *et al*., 2008). However, with non-axenic cultures, it is difficult to prove whether or not sequences generated are indeed of protist origin and not from potential bacterial contamination (Snyder *et al*., 2003, 2005; Cembella & John, 2006). This is true in particular for PKS genes from protists because they have been shown to belong exclusively to the modular PKS 1 type (John *et al*., 2008).

Working with degenerate primers, Snyder *et al*. (2003) managed to amplify fragments from non-axenically grown *Karenia brevis* cultures, showing homology to putative type I or II PKS genes. Yet, whether this gene was derived from the dinoflagellates or associated bacteria was unclear. Nevertheless, the presence of at least one PKS gene in the toxic dinoflagellate *Karenia brevis* was proven by the same author (Snyder *et al*., 2003). To circumvent the problems associated with the use of degenerate primers for PCR approaches, Monroe & van Dolah (2008) managed to isolate full-length sequences of single catalytic domains from PKS genes from the same organism, by screening cDNA libraries. They claim that the sequence is most similar to type I modular PKS, but that the structure is most similar to type II (for review see Hertweck *et al*., 2007). Fragments isolated from the dinoflagellate *Amphidinium* sp., which produces the polyketide amphidinolide, also showed similarity to β-ketoacyl synthase (KS), acyl transferase (AT), dehydratase (DH), ketoreductase (KR), and acyl carrier protein (ACP) and thioesterase (TE) in known type I PKS (Kubota & Kobayashi, 2006).

Dinoflagellate genes are known to contain a unique trans-splicing leader sequence, which can now be used to discriminate a dinoflagellate-specific gene from a mixed culture (Lidie & van Dolah, 2007; Zhang *et al*., 2007; Monroe & van Dolah, 2008; Zhang & Lin, 2008). Whether a similar mechanism is present in prymnesiophytes such as *Chrysochromulina* is unknown. In any case, phylogenetic analysis of the ketoacyl (KS) domain based upon the sequences of PKS4 and PKS7 from *Chrysochromulina* showed definitive phylogenetic

associations. Both sequences clustered among those of *E. huxleyi* (John *et al.*, 2008) and therefore formed a monophyletic haptophyte clade (Fig. 6). The PKS genes of *C. polylepis* belong to the type I group, as is the case for *Karenia brevis*, but whereas the structure of the dinoflagellate PKS, with its discrete catalytic domains, suggests a novel type I-like PKS gene, *C. polylepis* seems to exhibit the conventional modular type I structure. We cannot totally exclude the possibility of recent horizontal gene transfer of the PKS genes into *Chrysochromulina*, but given their phylogenetic clustering within the haptophyte clade this is unlikely.

Gene expression and regulation in synchronous cultures

Knowledge of the genetic regulation of toxin synthesis may help us to understand the environmental conditions favouring toxicity and the ecological relevance of many of those substances. We studied gene expression in synchronized cultures, which are useful in order to enhance the signal from induced toxin production (Pan *et al.*, 1999; John *et al.*, 2001; Eschbach *et al.*, 2005) and provide insights into regulation because the induction of a given biosynthetic pathway is phased to the cell division cycle and essentially occurs at the same time point for all cells in a culture. A high degree of synchronization is necessary to allow for clear identification of cell cycle stages when following temporal changes in gene expression throughout a 24-h sampling period. In practice, the induction of synchronous (as opposed to merely phased) cell division by manipulation of the photoperiod is only possible when the length of the cell division cycle closely approximates the length of the photoperiod (1 division per day). Partial synchronization in *C. polylepis* was obtained by the sequential inoculation of increasingly larger culture volumes of cells from early exponential growth phase preconditioned to a 14-h:10-h light–dark regime.

The cell cycle of *C. polylepis* has previously been analysed by flow cytometry (Edvardsen, 1996; Eschbach *et al.*, 2005). In our experiments, cell-cycle analysis was accomplished by quantifying the amount of DNA in fixed cells (Grey *et al.*, 1990), but we used an improved method based on glutaraldehyde fixation (Eschbach *et al.*, 2001). At the beginning of our experiments, *C. polylepis* cultures were in quasi-steady state (roughly balanced growth) and the cells passed through S and G2+M phases during the dark period (Fig. 6). Mitotic division in the dark is typical for many microalgae (Taroncher-Oldenburg *et al.*, 1997, 1999; John *et al.*, 2001; Farinas *et al.*, 2006), although there are exceptions to this rule, particularly for shade-adapted benthic species (Pan *et al.*, 1999). In a previous study it was shown that during the light period, all *C. polylepis* cells accumulated in G1 phase during which growth and other metabolic functions, such as chlorophyll and toxin biosynthesis, were carried out (Eschbach *et al.*, 2005).

The vast majority of phycotoxins are polyethers that are most likely derived *via* polyketide synthases (Wright & Cembella, 1998). As described above, ichthyotoxicity and cytolytic effects on cells and tissues caused by exposure to *P. parvum* led to speculation that they may be caused by similar toxins (John *et al.*, 2002). Among the prymnesiophytes, the prymnesins (PRM1 1 and PRM 2) isolated from *P. parvum* are the only structurally described toxins (Igarashi *et al.*, 1998).

In most cases thus far the heterologous expression of PKS genes has been studied for drug discovery or production (Schümann & Hertweck, 2006), or for investigation of differential expression patterns in different tissues or species (Karppinen & Hohtola, 2007; Lopez-Erraquin *et al.*, 2007). We expect that there are studies currently underway to examine the expression of PKS genes under various environmental stimuli in both lower eukaryotes and bacteria, but nothing is published to our knowledge. The only related work we are aware of is on the biosynthesis of the cyclic heptapeptide regulated by a peptide-PKS system *via* the *mcy* gene cluster in the cyanobacterium *Microcystis aeruginosa* (Kaebernick *et al.*, 2000). In the cyanobacterium the *mcy* mRNA levels were shown to increase during early and mid-exponential growth phase in a light-dependent manner.

We previously showed that the toxicity of *C. polylepis* increased at the transition from dark to the light phase (Eschbach *et al.*, 2005). Here we demonstrated that the PKS genes in *C. polylepis* expressed increased transcript levels in the dark phase. These two observations correlated nicely and suggested that PKS genes may indeed be linked to toxicity in this species. We caution, however, that this cannot be causally demonstrated because the chemical structures of the *C. polylepis* toxins are unknown and toxicity was not measured in this study. Based upon bioassay responses and conjecture regarding their mode of action, these toxins may be analogous or homologous to the mixed polyether prymnesins found in the related haptophyte *Prymnesium parvum*, but this remains to be established. Furthermore, we found several copies of different putative PKS genes, of which at least two are encoded in the *C. polylepis* genome. Thus, it is not clear which (if any) particular PKS gene products are responsible for toxicity. Moreover, the toxin cell quota could be regulated at several steps, involving transcription, mRNA

Downloaded By: [Stiftung] At: 14:20 6 September 2010

stability, translation, and protein activity, therefore the increase in transcript levels is not necessarily directly linked to an increase in toxin levels (Cembella & John, 2006). In future studies, detailed analyses combining toxin/toxicity measurements (analytical or *via* bioassay) and molecular genetic approaches will further elucidate insights into the expression and regulation of *C. polylepis* toxin production and its underlying processes.

Compared to the well-studied Opisthokonta and higher plant clade, molecular analysis in protists is still limited. Particularly for prymnesiophytes, molecular and physiological data are scarce. Therefore, the EST approach we present here is a first step which gives insights into genes involved in toxicity and growth control and is a starting point for elucidating genome properties and the complex life cycle of *C. polylepis* and probably other prymnesiophytes. This is the first study of PKS gene expression in microalgae and we generated interesting insights into the characteristics of the *Chrysochromulina polylepis* genome. With its approximately 230 MB genome size and its evolutionary relationship to *Emiliania huxleyi*, it is a perfect candidate for a future comparative genomics approach.

Downloaded By: [Stiftung] At: 14:20 6 September 2010

Acknowledgements

This research was supported as part of the EU projects EUKETIDES (QLK3-CT-2002-01940) and ESTTAL (GOCE-CT2004-511154). We thank B. Edvardsen for providing the *C. polylepis* strains (University of Oslo), M. Reckermann for helping us with the flow cytometry measurements and B. Beszteri and H. Vogel for fruitful discussions.

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucl. Acids Res.*, **25**: 3389–3402.

Ansari, M.Z., Yadav, G., Gokhale, R.S. & Mohanty, D. (2004). NRPS-PKS: a knowledge-based resource for analysis of NRPS/PKS megasynthases. *Nucl. Acids Res.*, **32**(Suppl. 2): W405–413.

Armbrust, E.V., Berges, J.A., Bowler, C., Green, B.R., Martinez, D., Putnam, N.H., Zhou, S., Allen, A.E., Apt, K.E., Bechner, M., *et al.* (2004). The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. *Science*, **306**: 79–86.

Ausubel, F., Kingston, R., Moore, D., Seidman, J., Smith, J. & Struhl, K. (1987). *Current Protocols in Molecular Biology*. John Wiley & Sons Inc., New York, NY, USA.

Ayuso-Sacido, A. & Genilloud, O. (2005). New PCR primers for the screening of NRPS and PKS-I systems in Actinomycetes: detection and distribution of these biosynthetic gene sequences in major taxonomic groups. *Microb. Ecol.*, **49**: 10–24.

Bairoch, A., Apweiler, R., Wu, C.H., Barker, W.C., Boeckmann, B., Ferro, S., Gasteiger, E., Huang, H., Lopez, R., Magrane, M., *et al.* (2005). The Universal Protein Resource (UniProt). *Nucl. Acids Res.*, **33**: D154–159.

Bapteste, E., Brinkmann, H., Lee, J.A., Moore, D.V., Sensen, C.W., Gordon, P., Durufle, L., Gaasterland, T., Lopez, P., Muller, M., *et al.* (2002). The analysis of 100 genes supports the grouping of three highly divergent amoebae *Dictyostelium*, *Entamoeba*, and *Mastigamoeba*. *Proc. Nalt. Acad. Sci.* USA **99**: 1414–1419.

Baskaran, N., Kandpal, R.P., Bhargava, A.K., Glynn, M.W., Bale, A. & Weissman, S.M. (1996). Uniform amplification of a mixture of deoxyribonucleic acids with varying GC content. *Genome Res.*, **6**: 633–638.

Beaton, M.J. & Cavalier-Smith, T. (1999). Eukaryotic non-coding DNA is functional: evidence from the differential scaling of cryptomonad genomes. *Proc. R. Soc. Lond. B*, **266**: 2053–2059.

Beck, H.P. (1978). A new analytical method for determining duration of phases, rate of DNA synthesis and degree of synchronization from flow-cytometric data on synchronized cell population. *Cell Tissue Kinet.*, **11**: 139–148.

Borejsza-Wysocki, W. & Hrazdina, G. (1996). Purification, characterisation, and antibody development to benzalaceton synthases from raspberry fruits. *Plant Physiol.*, **110**: 791–799.

Börner, T. & Dittmann, E. (2005). Molecular biology of cyanobacterial toxins. In *Harmful Cyanobacteria* (Huisman, J., Matthijs, H., and Visser, P., editors), 25–40. Springer Verlag, Berlin.

Bowler, C., Allen, A.E., Badger, J.H., Grimwood, J., Jabbari, K., Kuo, A., Maheswari, U., Martens, C., Maumus, F., Otillar, R.P., *et al.* (2008). The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature*, **456**: 239–244.

Cavalier-Smith, T. (2005). Economy, speed and size matter: Evolutionary forces driving nuclear genome miniaturization and expansion. *Annals Bot.*, **95**: 147–175.

Cembella, A.D. & John, U. (2006). Molecular physiology of toxin production and growth regulation in harmful algae. In *Ecology of Harmful Algae* (Granéli, E. and Turner, J.T., editors). Springer Verlag, Berlin.

Dahl, E., Bagoien, E., Edvardsen, B. & Stenseth, N.C. (2005). The dynamics of *Chrysochromulina* species in the Skagerrak in relation to environmental conditions. *J. Sea Res.*, **54**: 15–24.

Dahl, E., Lindahl, O., Paasche, E. & Throndsen, J. (1989). The *Chrysochromulina polylepis* bloom in Scandinavian waters during spring 1988. In *Novel Phytoplankton Blooms: Causes and Impacts of Recurrent Brown Tides and Other Unusual Blooms* (Cosper, E.M., Bricelj, V.M., and Carpenter, E.J., editors), 383–405. Springer Verlag, Berlin.

Derelle, E., Ferraz, C., Rombauts, S., Rouzé, P., Worden, A.Z., Robbens, S., Partensky, F., Degroeve, S., Echeynié, S., Cooke, R., *et al.* (2006). From the cover: Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *Proc. Natl. Acad Sci. USA*, **103**: 11647–11652.

Dittmann, E. & Wiegand, C. (2006). Cyanobacterial toxins – occurrence, biosynthesis and impact on human affairs. *Mol. Nutr. Food. Res.*, **50**: 7–17.

Edvardsen, B. (1996). *Toxicity, autecology and life history of* Chrysochromulina polylepis (*Haptophyta*). PhD thesis, University of Oslo, Norway.

Edvardsen, B., Eikrem, W., Green, J.C., Andersen, R.A., Moon-Van Der Staay, S.Y. & Medlin, L.K. (2000). Phylogenetic reconstructions of the Haptophyta inferred from rRNA sequences and available morphological data. *Phycologia*, **39**: 19–35.

Eichinger, L., Pachebat, J.A., Glöckner, G., Rajandream, M.A., Sucgang, R., Berriman, M., Song, J., Olsen, R., Szafranski, K., Xu, Q., *et al.* (2005). The genome of the social amoeba *Dictyostelium discoideum*. *Nature*, **435**: 43–57.

Eppley, R.W., Holmes, R.W. & Strickland, J.D.H. (1967). Sinking rates of the marine phytoplankton measured with a fluorochrometer. *J. Exp. Mar. Biol. Ecol.*, **1**: 191–208.

Downloaded By: [Stiftung] At: 14:20 6 September 2010

Eschbach, E., John, U., Cembella, A.D., Medlin, L.K., Reckermann, M. & Edvardsen, B. (2005). Cell cycle dependent expression of toxicity by the ichthyotoxic prymnesiophyte *Chrysochromulina polylepis*. *Aquatic Microb. Ecol.*, **39**: 85–95.

Eschbach, E., Reckermann, M., John, U. & Medlin, L.K. (2001). A simple and highly efficient fixation method for *Chrysochromulina polylepis* (Prymnesiophytes) for analytical flow cytometry. *Cytometry*, **44**: 126–132.

Ewing, B. & Green, P. (1998). Base-calling of automated sequencer traces using Phred. *II. Error probabilities*. *Genome Res.*, **8**: 186–194.

Farinas, B., Mary, C., O Manes, CL., Bhaud, Y., Peaucellier, G. & Moreau, H. (2006). Natural synchronisation for the study of cell division in the green unicellular alga *Ostreococcus tauri*. *Plant Mol. Biol.*, **66**: 277–292.

Gjøsæter, J., Lekve, K., Stenseth, N.C., Leinaas, H.P., Christie, H., Dahl, E., Danielssen, D.S., Edvardsen, B., Olsgard, F., Oug, E., *et al.* (2000). A long-term perspective on the *Chrysochromulina* bloom on the Norwegian Skagerrak coast 1988: A catastrophe or an innocent incident? *Mar. Ecol. Prog. Ser.*, **207**: 201–218.

Glöckner, G., Lehmann, R., Romualdi, A., Pradella, S., Schulte-Spechtel, U., Schilhabel, M., Wilske, B., Suhnel, J. & Platzer, M. (2004). Comparative analysis of the *Borrelia garinii* genome. *Nucl. Acids Res.*, **32**: 6038–6046.

Granéli, E., Paasche, E. & Maestrini, S.Y. (1993). Three years after the *Chrysochromulina polylepis* bloom in Scandinavian waters in 1988: Some conclusions of recent research and monitoring. In *Toxic Phytoplankton Bloom in the Sea* (Smayda, T.J. and Shimizu, Y., editors), 23–32. Elsevier Science, Amsterdam.

Grey, J.W., Dolbeare, F. & Pallavicini, M.G. (1990). Quantitative cell cycle analysis. In *Flow Cytometry and Sorting* 2nd ed (Melamed, M.R., Lindmo, T.L., and Mendelsohn, M.L., editors), 445–467. Wile-Lis, New York.

Guindon, S. & Gascuel, O. (2003). A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.*, **52**: 696–704.

Hackett, J., Scheetz, T., Yoon, H., Soares, M., Bonaldo, M., Casavant, T. & Bhattacharya, D. (2005). Insights into a dinoflagellate genome through expressed sequence tag analysis. *BMC Genomics*, **6**: 80.

Hallegraeff, G. (2003). Harmful algal blooms: A global overview. In *Manual on Harmful Marine Microalgae* (Hallegraeff, G., Anderson, D.M., and Cembella, A.D., editors), 25–49. UNESCO, Paris.

Hanekamp, K., Bohnebeck, U., Beszteri, B. & Valentin, K. (2007). PhyloGena – a user-friendly system for automated phylogenetic annotation of unknown sequences. *Bioinformatics*, **23**: 793–801.

Hertweck, C., Luzhetskyy, A., Rebets, Y. & Bechthold, A. (2007). Type II polyketide synthases: gaining a deeper insight into enzymatic teamwork. *Nat. Prod. Rep.*, **24**: 162–190.

Igarashi, T., Aritake, S. & Yasumoto, T. (1998). Biological activity of Prymnesin-2 isolated from a red tide alga *Prymnesium parvum*. *Nat. Toxins*, **6**: 35–41.

Ikeda, H. & Omura, S. (1997). Avermectin Biosynthesis. *Chem. Rev.*, **97**: 2591–2609.

Jaeckisch, N., Singh, R., Curtis, B., Cembella, A. & John, U. (2008). Genomic characterization of the spirolide-producing dinoflagellate *Alexandrium ostenfeldii* with special emphasis on PKS genes. Proceedings of the 12th International Conference on Harmful Algae (Moestrup, O., *et al.*, editors), 65–67. ISSHA and IOC-UNESCO, Copenhagen, Denmark.

Johansson, N. & Granéli, E. (1999). Cell density, chemical composition and toxicity of *Chrysochromulina polylepis* (Haptophyta) in regulation to different N:P supply ratios. *Mar. Biol.*, **135**: 209–217.

John, U., Beszteri, B., Derelle, E., Van de peer, Y., Read, B.A., Moreau, H. & Cembella, A.D. (2008). Novel insights into evolution of protistan polyketide synthases through phylogenomic analysis. *Protist*, **159**: 21–30.

John, U., Quilliam, M.A., Medlin, L.K. & Cembella, A.D. (2001). Spirolide production and photoperiod-dependent growth of the marine dinoflagellate *Alexandrium ostenfeldii*. In *Harmful Algal Blooms 2000* (Hallegraeff, G.M., Blackburn, S.I., Bolch, C.J., and Lewis, R.J., editors), 299–302. Intergovernmental Oceanographic Commission (UNESCO), Paris.

John, U., Tillmann, U. & Medlin, L. (2002). A comparative approach to study inhibition of grazing and lipid composition of a toxic and non-toxic clone of *Chrysochromulina polylepis* (Prymnesiophyceae). *Harmful Algae*, **1**: 45–57.

Jordan, R.W. & Green, J.C. (1994). A check-list of the extant Haptophyta of the world. *J. Mar. Biol. Assoc. UK*, **74**: 149–174.

Kaebernick, M., Neilan, B.A., Borner, T. & Dittmann, E. (2000). Light and the transcriptional response of the microcystin biosynthesis gene cluster. *Appl. Environ. Microbiol.*, **66**: 3387–3392.

Karpinnen, K. & Hohtola, A. (2007). Molecular cloning and tissue-specific expression of two cDNAs encoding polyketide synthases from *Hypericum perforatum*. *J. Plant Physiol.*, **165**: 1079–1086.

Ko, M.S.H. (1990). An 'equalized cDNA library' by the reassociation of short double-stranded cDNAs. *Nucleic Acids Res.*, **18**: 5705–5711.

Kroken, S., Glass, N.L., Taylor, J.W., Yoder, O.C. & Turgeon, B.G. (2003). Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes. *Proc. Natl. Acad. Sci. USA*, **100**: 15670–15675.

Kubota, T. & Kobayashi, Y. (2006). Cloning of polyketide synthase genes from amphidinolide-producing dinoflagellate *Amphidinium* sp. *Biol. Pharm. Bull.*, **29**: 1314–1318.

Kusebauch, B., Busch, B., Scherlach, K., Roth, M. & Hertweck, C. (2009). polyketide-chain branching by an enzymatic Michael addition. *Angew. Chem. Int. Ed.*, **48**: 5001–5004.

La Claire, J. (2006). Analysis of expressed sequence tags from the harmful alga, *Prymnesium parvum* (Prymnesiophyceae, Haptophyta). *Mar. Biotechnol.*, **8**: 534–546.

Lajeunesse, T.C., Lambert, G., Andersen, R.A., Coffroth, M.A. & Galbraith, D.W. (2005). *Symbiodinium* (Pyrrhophyta) genome sizes (DNA content) are smallest among dinoflagellates. *J. Phycol.*, **41**: 880–886.

Lassmann, T. & Sonnhammer, E. (2005). Kalign – an accurate and fast multiple sequence alignment algorithm. *BMC Bioinformatics*, **6**: 298.

Lee, T., Yun, S.H., Hodge, K.T., Humber, R.A., Krasnoff, S.B., Turgeon, G.B., Yoder, O.C. & Gibson, D.M. (2001). Polyketide synthase genes in insect- and nematode-associated fungi. *Appl. Microbiol. Biotech.*, **56**: 181–187.

Lidie, K.B. & Van Dolah, F.M. (2007). Spliced leader RNA-mediated *trans*-splicing in a dinoflagellate, *Karenia brevis*. *J. Eukaryot. Microbiol.*, **54**: 427–435.

Lidie, K., Ryan, J., Barbier, M. & Dolah, F. (2005). Gene expression in florida red tide dinoflagellate *Karenia brevis*: analysis of an expressed sequence tag library and development of DNA microarray. *Mar. Biotechnol.*, **7**: 481–493.

Lopez-Errasquin, E., Vazquez, C., Jimènez, M. & Gonzàlez-Jaén, M.T. (2007). Real-time RT-PCR assay to quantify the expression of *fum*1 and *fum*19 genes from the Fumonisin-producing *Fusarium verticillioides*. *J. Microbiol. Methods*, **68**: 312–317.

Mackinnon, S.L., Cembella, A.D., Burton, I.W., Lewis, N., Leblanc, P. & Walter, J.A. (2006). Biosynthesis of 13-desmethly spirolide C by the dinoflagellate *Alexandrium ostenfeldii*. *J. Org. Chem.*, **71**: 8724–8731.

Marsden, R.L., Lee, D., Maibaum, M., Yeats, C. & Orengo, C.A. (2006). Comprehensive genome analysis of 203 genomes provides structural genomics with new insights into protein family space. *Nucl. Acids Res.*, **34**: 1066–1080.

Matsuzaki, M., Misumi, O., Shin-i, T., Maruyama, S., Takahara, M., Miyagishima, S., Mori, T., Nishida, K., Yagisawa, F., Nishida, K., *et al.* (2004). Genome sequence of

the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. *Nature*, **428**: 653–657.

Merchant, S.S., Prochnik, S.E., Vallon, O., Harris, E.H., Karpowicz, S.J., Witman, G.B., Terry, A., Salamov, A., Fritz-Laylin, L.K., Maréchal-Drouard, L., *et al*. (2007). The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science*, **318**: 245–251.

Mock, T., Krell, A., Gloeckner, G., Kolukisaoglu, U. & Valentin, K. (2006). Analysis of expressed sequence tags (ESTs) from the polar diatom *Fragilariopsis cylindrus*. *J. Phycol*., **42**: 75–85.

Monroe, E.A. & Van Dolah, F.M. (2008). The toxic dinoflagellate *Karenia brevis* encodes novel type I-like polyketide synthases containing discrete catalytic domains. *Protist*, **159**: 471–482.

Moore, B.S. & Hertweck, C. (2002). Biosynthesis and attachment of novel bacterial polyketide synthase starter units. *Nat. Prod. Rep*., **19**: 70–99.

Myklestad, S.M. (1995). Release of extracellular products by phytoplankton with special emphasis on polysaccharides. *Science Total Env*., **165**: 155–164.

Nicholson, T.P., Rudd, B.A.M., Dawson, M., Lazarus, C.M., Simpson, T.J. & Coxa, R.J. (2001). Design and utility of oligonucleotide gene probes for fungal polyketide synthases. *Chem. Biol*., **8**: 157–178.

Nielsen, T.G., Kiorboe, T. & Bjornsen, P.K. (1990). Effect of a *Chrysochromulina polylepis* subsurface bloom on the planktonic community. *Mar. Ecol.Prog. Ser*., **62**: 21–35.

Nosenko, T., Boese, B. & Bhattacharya, D. (2007). Pulsed-Field Gel Electrophoresis analysis of genome size and structure in *Pavlova gyrans* and *Diacronema* sp. *(Haptophyta)*. *J. Phycol*., **43**: 763–767.

Pan, Y., Cembella, A.D. & Quilliam, M.A. (1999). Cell cycle and toxin production in the benthic dinoflagellate *Prorocentrum lima*. *Mar. Biol*., **134**: 541–549.

Pereira, S.L., Leonard, A.E., Huang, Y.-S., Chuang, L.-T. & Mukerji, P. (2004). Identification of two novel microalgal enzymes involved in the conversion of the omega3-fatty acid, eicosapentaenoic acid, into docosahexaenoic acid. *Biochem. J*., **384**: 357–366.

Proksch, P., Edrada, R. & Ebel, R. (2002). Drugs from the seas - current status and microbiological implications. *Appl. Microbiol. Biotech*., **59**: 125–134.

Rudd, S. (2003). Expressed sequence tags: Alternative or complement to whole genome sequences? *Trends Plant Sci*., **8**: 321–329.

Saez, A., Medlin, L.K. & Huss, V. (2001). Determination of the plastid genome size in Haptophytes. *Phycologia*, **40**: 142–157.

Sànchez Puerta, M.V., Bachvaroff, T.R. & Delwiche, C.F. (2004). The complete mitochondrial genome sequence of the haptophyte *Emiliania huxleyi* and its relation to heterokonts. *DNA Research*, **11**: 1–10.

Scala, S., Carels, N., Falciatore, A., Chiusano, M.L. & Bowler, C. (2002). Genome properties of the diatom *Phaeodactylum tricornutum*. *Plant Physiol*., **129**: 993–1002.

Schmidt, L.E. & Hansen, P.J. (2001). Allelopathy in the prymnesiophyte *Chrysochromulina polylepis*: effect of cell concentration, growth phase and pH. *Mar. Ecol. Prog. Ser*., **216**: 67–81.

Schümann, J. & Hertweck, C. (2006). Advances in cloning, functional analysis and heterologous expression of fungal polyketide synthase genes. *J. Biotechnol*., **124**: 690–703.

Skjoldal, H.R. & Dundas, I. (1991). The *Chrysochromulina polylepis* bloom in the Skagerrak and the Kattegat in May–June 1988. *ICES Cooperative Research Report* 175.

Snyder, R.V., Gibs, P.D.L., Palacios, A., Abiy, L., Dickey, R., Lopez, J.V. & Rein, K.S. (2003). Polyketide synthase genes from marine dinoflagellates. *Mar. Biotechnol*., **5**: 1–12.

Snyder, R.V., Guerrero, M.A., Sinigalliano, C.D., Winshell, J., Perez, R., Lopez, J.V. & Rein, K.S. (2005). Localization of polyketide synthase encoding genes to the toxic dinoflagellate *Karenia brevis*. *Phytochem*., **66**: 1767–1780.

Staunton, J. & Weissmann, K.J. (2001). Polyketide biosynthesis: A millenium review. *Nat. Prod. Rep*., **18**: 380–416.

Taroncher-Oldenburg, G., Kulis, D.M. & Anderson, D.M. (1997). Toxin variability during the cell cycle of the dinoflagellate *Alexandrium fundyense*. *Limnol. Oceanogr*., **42**: 1178–1188.

Taroncher-Oldenburg, G., Kulis, D.M. & Anderson, D.M. (1999). Coupling of saxitoxin biosynthesis to the G sub(1) phase of the cell cycle in the dinoflagellate *Alexandrium fundyense*: Temperature and nutrient effects. *Nat. Toxins*, **7**: 207–219.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. & Higgins, D.G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucl. Acids Res*., **25**: 4876–4882.

Von Dassow, P., Petersen, T.W., Chepurnov, V.A. & Armbrust, E.V. (2008). Inter- and intraspecific relationships between nuclear DNA content and cell size in selected members of the centric diatom genus *Thalassiosira* (Bacillariophyceae). *J Phycol*., **44**: 335–349.

Wahlund, T.M., Hadaegh, A.R., Clark, R., Nguyen, B., Fanelli, M. & Read, B.A. (2004). Analysis of expressed sequence tags from calcifying cells of marine coccolithophorid (*Emiliania huxleyi*). *Mar. Biotechnol*., **6**: 278–290.

Weber, A., Oesterhelt, C., Gross, W., Bräutigam, A., Imboden, L., Krassovskaya, I., Linka, N., Truchina, J., Schneidereit, J., Voll, H., Voll, L., *et al*. (2004). EST-analysis of the thermo-acidophilic red microalga *Galdieria sulphuraria* reveals potential for lipid A biosynthesis and unveils the pathway of carbon export from rhodoplasts. *Plant Mol. Biol*., **55**: 17–32.

Worden, A.Z., Lee, J.H., Mock, T., Rouzé, P., Simmons, M.P., Aerts, A.L., Allen, A.E., Cuvelier, M.L., Derelle, E., Everett, M.V., Foulon, E., *et al*. (2009). Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes. *Micromonas*. *Science*, **324**: 268–272.

Wright, J.L.C. & Cembella, A.D. (1998). Ecophysiology and biosynthesis of polyether marine biotoxins. In *Physiological Ecology of Harmful Algal Blooms* (Andreson, D.M., Cembella, A.D., and Hallegraeff, G.M., editors), 427–451. Springer-Verlag, Berlin.

Zhang, H., Hou, Y., Miranda, L., Campbell, D.A., Sturm, N.R., Gaasterland, T. & Lin, S. (2007). Spliced leader RNA trans-splicing in dinoflagellates. *Proc. Natl. Acad. Sci. USA*, **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. *J. Phycol*., **44**: 703–711.

Zhu, G., Lagier, M.J., Stejskal, F., Millership, J.J., Cai, X. & Keithly, J.S. (2002). *Cryptosporidium parvum*: The first protist known to encode a putative polyketide synthase. *Gene*, **298**: 79–89.

Downloaded By: [Stiftung] At: 14:20 6 September 2010