Dinoflagellates Come from Outer Space, but Haptophytes and Diatoms Do Not

Uwe John[1], Thomas Mock[1], Klaus Valentin[1], Allan D. Cembella[2], and Linda K. Medlin[1]
[1]*Alfred-Wegener-Institut, Postfach 120161, D-27570 Bremerhaven, Germany, ujohn@awi-bremerhaven.de;*
[2]*Institute for Marine Biosciences, National Research Council (Canada), 1411 Oxford St., Halifax, Nova Scotia, Canada B3H 3Z1*

Abstract

Normalized cDNA libraries were generated for a dinoflagellate (*Alexandrium ostenfeldii*) and a haptophyte (*Chrysochromulina polylepis*), both photosynthetic species capable of forming toxic blooms. Partial sequences were obtained from 2500 clones from each library. After annotation, these represented about 1400 unique sequences for each species. Several genes putatively related to toxin synthesis were detected. Yet only 9% of the total sequences were homologues to known genes for *A. ostenfeldii,* whereas the corresponding number for *C. polylepis* was 32%. A cDNA library for the psychrophilic diatom *Fragilariopsis cylindrus* was also established after cold-shock treatment. From this library, 350 clones were partially sequenced and 40% were identified by BLAST search. In summary, the percentage of identifiable genes in the dinoflagellate was substantially lower than in other protists examined. These data provide preliminary indications that dinoflagellates possess a radically different genome from other "algal protists" and thus their uniqueness makes them interesting candidates for genomics research.

Introduction

As members of the phytoplankton, diatoms and free-living dinoflagellates are often the dominant eukaryotic groups contributing to plankton biomass and primary production in both coastal and oceanic waters. Yet unlike the diatoms and other phytoplankton, dinoflagellates have long been considered as belonging to both the botanical and zoological realm, because this group comprises many heterotrophic and parasitic forms. In fact, the dinoflagellates have more apparent affinities with ciliates and certain amoebae than with classic microalgal groups, posing a problem in the construction of an explicit phylogeny for the lower eukaryotes (Taylor, 1980). Although a few genes (18S and 28S rDNA, in particular) are frequently sequenced for phylogenetic reconstruction, there is little genomic information on the major groups of eukaryotic microalgae. Even less genomic data is available on gene expression related to primary and intermediary metabolism, biosynthesis of secondary metabolites, including toxins, and the molecular ecology of adaptation to ecological niches. One attempt to resolve these unknowns is to apply the technologies for mass sequencing of genomic DNA of key microalgal species. Such species may include important primary producers, *e.g.,* spring-bloom diatoms; those involved in harmful algal blooms (HABs), *e.g.,* certain dinoflagellates, haptophytes, and raphidophytes; or psychrophiles, *e.g.,* sea-ice microalgae. Although at present it is not practical to completely sequence the genomes of all key species, much information can be gained from random sequencing of cDNA libraries, *i.e.,* by generating expressed sequence tag (EST) libraries. Such libraries are constructed by extracting the messenger RNA, which represents the transcripts of all genes being expressed at a given time. These mRNAs are converted to DNA by reverse transcriptase, cloned, and randomly sequenced. Each clone essentially represents a single gene. In this study, we compared the percentage of identifiable genes from cDNA libraries obtained from the diatom *Fragilariopsis cylindrus,* the toxic dinoflagellate *Alexandrium ostenfeldii,* and the toxic haptophyte *Chrysochromulina polylepis.*

Materials and Methods

Culture and Harvesting of Algae Unialgal cultures were grown in 10-L glass flasks on 14:10 h light:dark cycle in a controlled growth chamber. *Chrysochromulina polylepis* B1511 from Oslo Fjord, Norway, was grown in IMR/2 growth medium (Eppley *et al.,* 1967), supplemented with 10 nM selenite; *Alexandrium ostenfeldii* AOSH1 from Ship Harbour, Nova Scotia, Canada in K medium (Keller *et al.,* 1987); and *Fragilariopsis cylindrus,* isolated from Antarctic sea-ice in the eastern Weddell Sea, was grown in f/2 medium (Guillard and Ryther, 1962). *Fragilariopsis cylindrus* was grown at 5°C at a photon flux density of 35 µmol m^{-2} s^{-1}. The other species were maintained at 15°C, at a photon flux density of 45 µmol m^{-2} s^{-1} for *C. polylepis,* and 90 µmol m^{-2} s^{-1} for *A. ostenfeldii.* All cultures were harvested in the middle of exponential growth by continuous flow centrifugation.

Isolation of mRNA Total RNA of *C. polylepis* and *F. cylindrus* was isolated with an RNeasy Plant Mini Kit (Qiagen, Hilden, Germany). For *F. cylindrus,* mRNA was isolated from ca. 100 µg total RNA with an Oligotex mRNA Midi-Kit (Qiagen, Hilden, Germany.) Total RNA of *A. ostenfeldii* was extracted with peqGOLD RNApure (PEQLAB Biotechnology, Erlangen, Germany) and cleaned with peqGOLD Optipure following manufacturer's instructions.

cDNA Library Construction and Sequence Analysis Normalized cDNA libraries of *C. polylepis* and *A. ostenfeldii* were constructed in cooperation with BASF Plant Science (Ludwigshafen, Germany) and prepared by Vertis Biotechnologie AG (Freising, Germany). The cDNAs were synthesised from 2.4 µg total RNA and directionally cloned into Not I/Asc I-sites of a plasmid vector pFDX3840 (sup-

Steidinger, K. A., J. H. Landsberg, C. R. Tomas, and G. A. Vargo (Eds.). 2004. Harmful Algae 2002. Florida Fish and Wildlife Conservation Commission, Florida Institute of Oceanography, and Intergovernmental Oceanographic Commission of UNESCO.

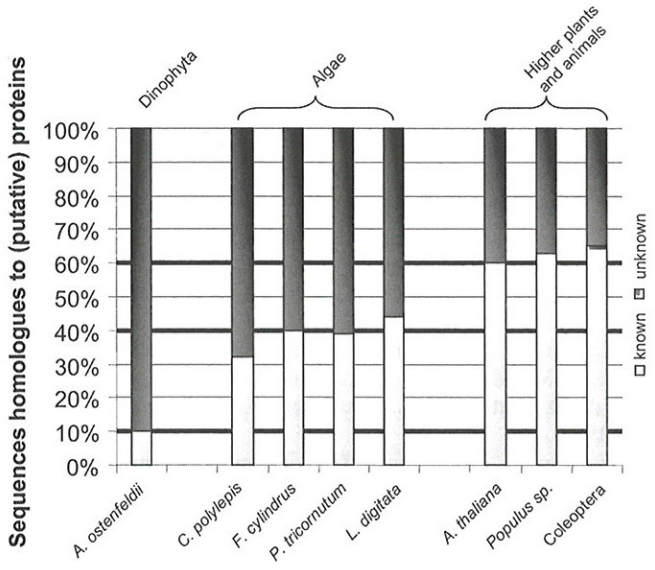

Figure 1 Comparison of ESTs of several species of different systematic groups. Dark bar shows percentage of unknown sequences; light bar indicates percentage of ESTs that show homology to proteins or putative proteins.

plied by Prof. Dr. Ralf Reski, Freiburg). The cDNA libraries were normalised to reduce redundant DNA, ideally leaving only a single copy of each expressed gene to be cloned and sequenced. Normalisation of the cDNA was performed according to Ko (1990) with several modifications. The cDNA library of *F. cylindrus* was synthesised with approximately 800 ng mRNA using a SMART™ cDNA Library Construction Kit (Clontech) following the manufacturer's instructions. White plaques were used for large-scale PCR analysis with 5′and 3′ λTripEx LD-Insert Screening Amplimers. Approximately 350 clones larger than 564 bp were chosen for sequencing analysis. From each normalised cDNA library of *C. polylepis* and *A. ostenfeldii,* 2500 clones were sequenced from the 5' end. The inserts from *F. cylindrus* were also sequenced from the 5' end using λTripEx2 Sequencing Primer. On average, 500 bp were determined. Sequences were compared to detect overlapping clones. All contiguous segments were compared to gene libraries (*e.g.,* GenBank) by BLAST searches and all possible reading frames were analysed.

Results

Expressed Sequence Tag (EST) Libraries (Fig. 1) After annotation, the partial sequences of 2,500 clones from normalised cDNA libraries represented 1,443 unique sequences for *C. polylepis* and 1,416 for *A. ostenfeldii.* In the case of *C. polylepis,* 466 (32%) sequences were homologous to known proteins and 617 (46%) to sequences in EST databases. However, only 120 (9%) of the sequences of *A. ostenfeldii* were homologous to known proteins, whereas 541 (38%) showed similarity to EST sequences. Eight putative polyketide synthase (PKS) genes were identified for *C. polylepis,* and two putative PKS genes were found for *A. ostenfeldii.* The phage cDNA library of the diatom *F. cylindrus* comprised approximately 1.5 Mio recombinant clones per milliliter, and thus likely covered the expressed genome at this time point. Of the 186 EST contigs, 72 (40%) showed homology to known proteins or hypothetical conserved open reading frames (ORFs).

Discussion

The percentage of identifiable genes in the dinoflagellate *Alexandrium ostenfeldii* was dramatically lower than for other protistan algae, indicating that dinoflagellates possess a highly unusual genome. The unique characteristics of dinoflagellates with respect to ultrastructure (extra-nuclear spindle, permanently condensed chromatin, etc.), ploidy (typically *n* in the vegative stage), high amount of nuclear DNA, general absence of histones and nucleosomes, and unusual mitosis, meiosis and life history, support the conclusion that much of the DNA may be non-coding and structural, perhaps subjected to a high degree of redundancy. The low degree of similarity of cDNAs in the dinoflagellate to known sequences further suggests that metabolic pathways may also be rather divergent. A comparative study of ESTs in the toxigeneic dinoflagellate, *Karenia brevis,* did show a higher degree of homology to known genes (32%) in other organisms than we found for *A. ostenfeldii* (Lidie *et al.,* 2003). The abundancy expressed genes were primarily related to general metabolism, signal transduction and transcription/translation. The phylogenetic relationship between dinoflagellates, particularly parasitic and heterotrophic species, and apicomplexans has been recognized, although recent molecular studies on the fish-killing dinoflagellate *Pfiesteria piscicida* and other *Pfiesteria*-like dinoflagellates showed that these dinoflagellates form a distinct group and are not that clearly related to the apicomplexa (Litaker *et al.,* 1999). While we do not literally mean to imply that dinoflagellates are extraterrestrial in origin (from "outer space"), our data suggest that a genomic project on dinoflagellates would likely identify interesting and new, perhaps even highly abnormal, genes.

The percentage of identifiable genes in the two chromophytic microalgae examined in this study was similar and approximately two-thirds of that of higher plants, such as *Arabidopsis thaliana* (Meinke *et al.,* 1998) and two woody *Populus* species (Sterky *et al.,* 1998), or animals such as coleopteran insects (Theodorides *et al.,* 2002). Degrees of similarity were in the same range as for two other chromophytes, *e.g., Laminaria digitata* (Phaeophyta), Crepineau *et al.,* 2000; and *Phaeodactylum tricornutum* (Bacillariophyta), Scala *et al.,* 2002. Chromophytic algae contain, however, a higher percentage of unknown genes than higher plants (Meinke *et al.,* 1998). This is not surprising because heterokonts and haptophytes represent evolutionary lineages that separated from other eukaryotic lineages around 1200 million years ago (Medlin *et al.,* 1997), and molecular data on chromophytes are much less studied than for higher plants.

Despite the fact that only a limited number of ESTs were sequenced for the toxigenic species, several significant genes were detected. In both toxic species, *A. ostenfeldii* and *C. polylepis,* a number of genes encoding polyketide synthases (PKS) was found. These genes are known to be involved in the biosynthesis of certain potent polyketide phycotoxins, including spirolides, diarrhetic shellfish poisoning toxins, and brevetoxins (Wright and Cembella, 1998). Their detection enables study of PKS gene expression and the relationship to occurrence and biosynthesis of toxins. In future, the detection of toxin-related PKS transcripts in environmental samples, *e.g.,* by real-time reverse transcriptase polymerase chain reaction (Real-Time RT-PCR), may allow discrimination of toxin-producing from non-toxigenic algal strains.

The present study not only highlights the unique characteristics of dinoflagellate sequences, but also emphasises the potential for detecting a host of important functional genes from marine phytoplankton. We have demonstrated that random sequencing of cDNA libraries is an efficient tool to do so. In view of the enormous global ecological importance of dinoflagellates, diatoms and haptophytes in marine ecosystems, further molecular studies on these eukaryotic protists are definitely warranted. We have no doubt that the most interesting discoveries on the functional significance of these genes still lie ahead.

References

F. Crépineau, T. Roscoe, R. Kaas, B. Kloareg and C. Boyen, Plant Mol. Biol. 43, 503–513 (2000).

R.W. Eppley, Holmes and J.D.H. Strickland, Exp. Mar. Biol. Ecol. 1, 191–208 (1967).

R.R. Guillard and J.H. Ryther, Can. J. Microbiol. 8, 229–239 (1962).

M.D. Keller, R.C. Selvin, W. Claus and R.R.L. Guillard, J. Phycol. 23, 633–638 (1987).

M.S. Ko, Nuc. Acids Res. 19, 5705–5711 (1990).

K.L. Lidie, J.C. Ryan, M. Barbier, and F.M. Van Dolah. 2nd Symposium on Harmful Algae in the U.S., Woods Hole, MA Dec 8–15 (2003).

R. W. Litaker, P.A. Tester, A. Colorni, M.G. Levy and E.J. Noga, J. Phycol. 35, 1379–1389 (1999).

L. K. Medlin, W. H. C. F. Kooistra, D. Potter, G. Saanders and R. A. Wandersen, in: The origin of the algae and their plastids, D. Bhattacharya. Plant Systm. Evol. (Suppl.) 11,187–219 (1997).

D.W. Meinke, J.M. Cherry, C.D. Dean, S. Rounsley and M. Koornneef, Science 282, 662–682 (1998).

S. Scala, N. Carels, A. Falciatore, M.L. Chiusano and C. Bowler, Plant Physiol. 129, 993–1002 (2002).

F. Sterky, S. Regan, J. Karlson, M. Hertzberg, A. Rohde, A. Holmberg, B. Amini, R. Bhalerao, M. Larson, R. Villarroel, M. van Monagu, G. Sandberg, O. Olsson, T.T. Teeri, W. Boerjan, P. Gustafsson, M. Uhlén, B. Sundberg and J. Lundberg, Proc. Natl. Acad. Sci. USA 95, 13330–13335 (1998).

F.J.R. Taylor, Biosystems 13: 65–108 (1980).

K. Theodorides, A. DeRiva, J. Gómez-Zurita, P.G. Foster and A.P. Vogler, Insect Mol. Biol. 11, 467–475 (2002).

J.L.C. Wright and A.D. Cembella, in: Physiological Ecology of Harmful Algal Blooms, D. M. Anderson, G.M. Hallegraeff and A.D. Cembella, eds. (Springer-Verlag, Heidelberg), pp. 427–451 (1998).