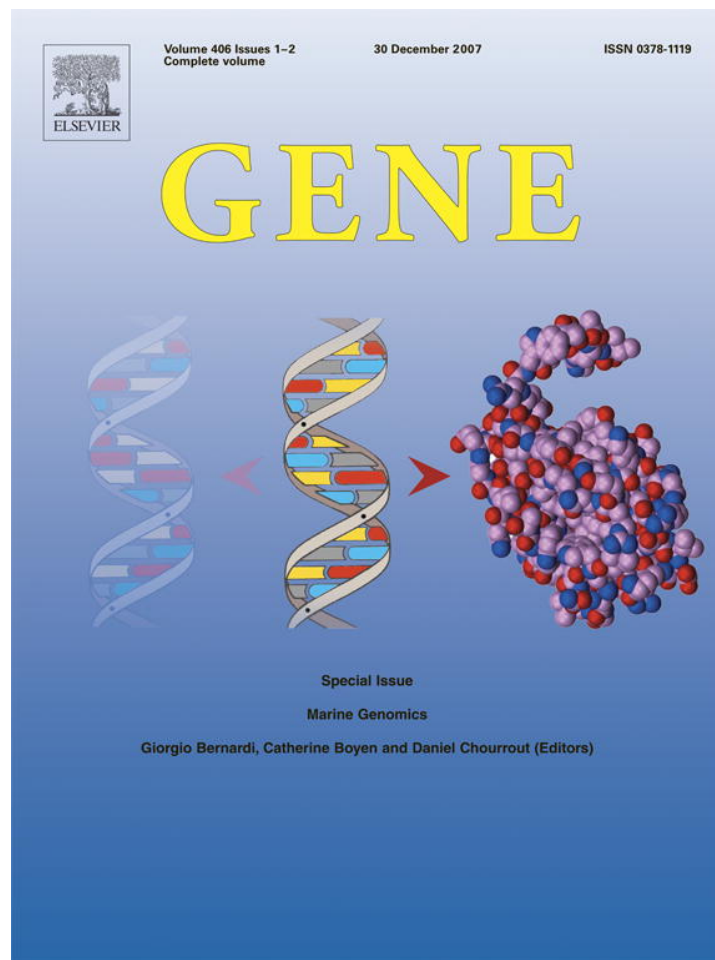


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article was published in an Elsevier journal. The attached copy is furnished to the author for non-commercial research and education use, including for instruction at the author's institution, sharing with colleagues and providing to institution administration.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Pilot study of an EST approach of the coccolithophorid *Emiliana huxleyi* during a virus infection

Jessica Kegel^a, Michael J. Allen^b, Katja Metfies^a, William H. Wilson^c,
Dieter Wolf-Gladrow^a, Klaus Valentin^{a,*}

^a Alfred Wegener Institute for Polar and Marine Research, Am Handelshafen 12, D-27570 Bremerhaven, Germany

^b Plymouth Marine Laboratory, Prospect Place, The Hoe, Plymouth, PL1 3DH, United Kingdom

^c Bigelow Laboratory for Ocean Sciences, 180 McKown Point Road, POB 475, Maine 04575, USA

Received 15 January 2007; received in revised form 20 September 2007; accepted 4 October 2007

Available online 12 October 2007

Abstract

Blooms of the coccolithophorid *Emiliana huxleyi* can be infected by viruses, which can lead to bloom-termination. This pilot study used an expressed sequence tag (EST) approach to get a first view of gene-expression changes that occur during viral infection of *E. huxleyi*. cDNA libraries were constructed from uninfected cultures and 6, 12, and 24 h after infection with *E. huxleyi*-specific virus 86 (EhV-86). From each library 60–90 ESTs were randomly selected and annotated manually with PhyloGena. Viral genes were identified using BLAST-Search of the known viral genome. The data of this study show, that 6 h after viral infection the algal transcriptome changed significantly although few viral transcripts were present. At this point, changes mainly concerned transcripts related to photosynthesis and protein metabolism. However, after 24 h viral transcripts were most abundant. Viral transcripts found at this stage of viral infection encode proteins involved in protein degradation, nucleic acid degradation, transcription and replication.

© 2007 Elsevier B.V. All rights reserved.

Keywords: *Emiliana huxleyi*; Coccolithophore; EhV-86; EST sequencing

1. Introduction

Emiliana huxleyi (Lohmann, 1902; Hay et al., 1967) is the most abundant coccolithophore and an important member of the marine phytoplankton. It is well known for its immense coastal and open ocean blooms ranging from sub-polar to tropical latitudes (Balch et al., 1992; Brown and Yoder 1994) that can cover 10,000 km² or more (Holligan et al., 1993, Winter et al., 1994). *E. huxleyi* is also regarded as a major sink for calcium carbonate in the ocean (Eide 1990, Samtleben and Bickert 1990, Baumann et al., 2004). Due to the reflection of their coccoliths

blooms can be observed by satellites (Holligan et al., 1983, Balch et al., 1991).

The abundance and wide distribution of *E. huxleyi* and its production of calcium carbonate coccoliths and dimethylsulfo-niopropanate (DMSP) make it an important species with respect to sediment formation and to ocean climate and natural acid rain (Charlson et al., 1987, Westbroek et al., 1993, Malin et al., 1994). Furthermore it is a key species for current studies on global biogeochemical cycles (Westbroek et al., 1994).

Viral lysis is thought to be one of the main causes for the termination of *E. huxleyi* blooms. Several studies have investigated the role of viruses in controlling the bloom development of *E. huxleyi* (Bratbak et al., 1993, 1995, 1996, Brussaard et al., 1996, Castberg et al., 2001, Jaquet et al., 2002, Wilson et al., 1998, 2002a, 2002b). It became evident from these investigations that viruses are intrinsically linked to the decline of *E. huxleyi* blooms.

Viruses are the most abundant biological agents in marine aquatic environments (Bergh et al., 1989, Suttle 2000, Wommack and Colwell 2000) and it is likely that most microbial

Abbreviations: bp, base pair; cDNA, complementary to RNA; cfu, colony forming units; EST, expressed sequence tag; *e* value, expectation value; fcp, fucoxanthin/chlorophyll binding protein; HSP70, heat shock protein 70; mRNA, messenger RNA; ORF, open reading frame; pfu, plaque forming unit.

* Corresponding author. Alfred Wegener Institute for Polar and Marine Research, Am Handelshafen 12, D-27570 Bremerhaven, Germany. Tel.: +49 471 4831 1452; fax: +49 471 4831 1149.

E-mail address: Klaus.Valentin@awi.de (K. Valentin).

organisms can be infected with a particular virus. Therefore they play important roles in nutrient (Wilhelm and Suttle 1999) and biogeochemical (Fuhrmann 1999) cycling, and influence structure and diversity of microbial and phytoplankton communities (Fuhrmann 1999, Wommack and Colwell 2000). Viruses have also been observed to infect a wide range of aquatic algae (Van Etten et al., 1991, Van Etten and Meints 1999), including bloom-forming marine phytoplankton (Nagasaki et al., 1994a, 1994b, Jacobsen et al., 1996, Nagasaki and Yamaguchi 1997, Sandaa et al., 2001).

A range of different viruses that infect *E. huxleyi* (EhV) were isolated from the English Channel and off the coast of Bergen, Norway (Castberg et al., 2002, Wilson et al., 2002b) and were analysed for their phylogeny (Schroeder et al., 2002), ecological successions in mesocosm experiment (Schroeder et al., 2003) and genome structure (EhV-86) (Allen et al., 2006, 2007, Wilson et al., 2005). Characterization of their sequences revealed that the *E. huxleyi* viruses are large double-stranded DNA viruses with genomes approximately 410 kbp in size and that they belong to a new virus genus termed Coccolithovirus based on the phylogeny of their DNA-polymerase gene (Schroeder et al., 2002). Coccolithoviruses belong to the *Phycodnaviridae* (Wilson et al., 2005), a diverse family of large icosahedral viruses that infect marine or freshwater eukaryotic algae. They all contain dsDNA genomes ranging from 180–560 kb (Van Etten et al., 2002).

Expressed sequence tag (EST) analysis is a useful tool to study gene expression and to discover novel genes. ESTs are small pieces of DNA sequence that are generated by sequencing and based on the creation of a cDNA library. By statistical evaluation of the frequency of the sequences for specific genes it is possible to develop an expression profile at different environmental conditions for genes of different cDNA libraries. Thereby it is possible to investigate the up- and down regulation of genes (Schmitt et al., 1999) or to compare the gene expression under different conditions (Rafalski et al., 1998). The establishment of ESTs from *E. huxleyi* at different stages of viral infection could thus be an effective means for expression analysis of virus infected cultures for which the viral genome is known. ESTs specify the type and rate of viral and host transcripts at a particular time. As a result of that, it is possible to speculate on mechanisms of host-virus interaction that occur in the host cell during viral infection in both partners.

The aim of this work was to provide a functionally annotated preliminary set of ESTs from *E. huxleyi* expressed before and during a virus infection in order to determine differentially expressed genes. The results of this study made it possible to estimate the proportional abundance of viral transcripts in relation to the whole transcriptome of the host cell during progression of the infection.

2. Materials and methods

2.1. Strains and growth conditions

Cultures of *E. huxleyi* CCMP1516 were grown in *f/2* medium (Guillard 1975) at 15 °C under a 16:8 light–dark cycle (150 $\mu\text{mol photons m}^{-2}\text{s}^{-1}$). Because of the availability of the complete

genome sequence of the virus EhV-86, this species was used for the infection of *E. huxleyi*. Exponentially growing cultures (approx. 1.2×10^6 cells mL^{-1}) were inoculated with EhV-86 lysate (2 mL per litre of culture, approx. 1×10^6 pfu mL^{-1}) in the middle of the dark phase.

2.2. RNA extraction from uninfected cells

Cultures (50 mL) were harvested on 1.2 μm filters (Millipore), transferred into a cryogenic vial (Nalgene), immediately frozen in liquid nitrogen and stored at -80 °C until use for analysis.

RNA of uninfected cultures was isolated at five different time points in series. This approach was chosen because RNA from infected cells (see below) was taken at different times after infection and as such at different phases of the cell cycle. Starting time was at the late exponential phase (approx. 1×10^6 cells mL^{-1}) and the last point was at the beginning of the stationary phase (approx. 3.3×10^6 cells mL^{-1}) (Fig. 1). Total RNA was isolated with the RNeasy Plant Mini Kit (Qiagen) according to the manual including one more washing step with buffer RW1 and buffer RPE and a DNase digestion subsequent to the original protocol. Afterwards mRNA was isolated with the Oligotex mRNA Mini Kit (Qiagen). Before library construction the mRNA from five different time points was pooled (Fig. 1) and precipitated with 0.5 volumes of LiCl overnight at -20 °C. Following centrifugation (20,000 g, 60 min, 4 °C), the supernatant was discarded and the pellet was washed three times with 100 μL ice-cold 70% ethanol. The pellet was air dried, resuspended in 10 μL DEPC-treated water and used for library construction.

2.3. RNA extraction from infected cells

After 6, 12 and 24 h of virus-infection cultures (250 mL) were filtered through 0.45 μm filters (Millipore). The filtrate was discarded and the filters transferred to clean petri dishes. Cells from each filter were resuspended in 2 mL of $1 \times$ phosphate buffered saline (PBS), centrifuged (20,000 g, 5 min), resuspended (by vortexing) in 2 mL RNeasy lysis buffer (Qiagen) and stored at -20 °C until ready for processing. RNA extraction was performed using an RNeasy Midi Kit (Qiagen). Samples were centrifuged (20,000 g, 2 min) and the pellet resuspended in 2 mL RLT buffer (+20 μL 2-mercaptoethanol). Following vigorous vortexing (1 min, in 5 s bursts), the samples were spun (20,000 g, 5 min) and the supernatant transferred to a 15 mL Falcon tube containing 2 mL 70% ethanol. Following vigorous mixing the samples were applied to a Qiagen MidiPrep column, centrifuged (3200 g, 5 min) and the flow-through discarded. Columns were washed twice with 2.5 mL RPE buffer (3200 g, 5 min) and transferred to a new Falcon tube. RNase free water (250 μL) was added, the samples incubated (room temperature, 1 min) and the RNA eluted by centrifugation (3200 g, 5 min).

To precipitate RNA solutions, 0.5 volumes of 7.5 M NH_4Ac and 2 volumes of 100% ethanol was added and the samples incubated at -80 °C overnight. Following centrifugation (20,000 g, 30 min), the supernatant was discarded and the pellet washed twice with 0.5 mL 70% ethanol (20,000 g, 30 min). The

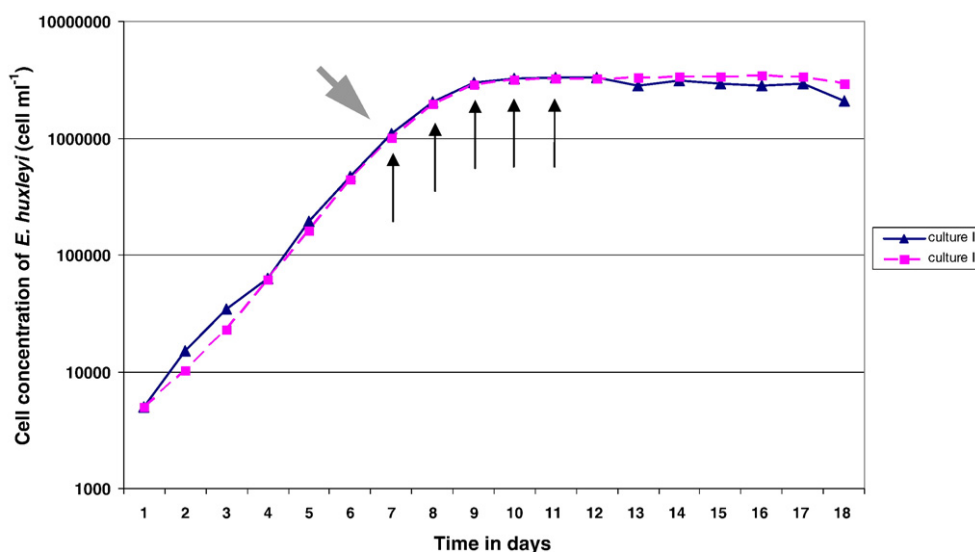


Fig. 1. Growth curve of an uninfected *Emiliania huxleyi* culture and the time points for the pooled mRNA (demonstrated by black arrows). The grey and dashed arrow shows the equivalent time point when the virus was added to the culture (approx. 1.2×10^6 cells mL^{-1}) in the separate infection experiment.

pellet was air dried, resuspended in 50 μL RNase free water and stored at -80°C .

2.4. Library construction

The uninfected library stems from pooled RNA collected throughout the growth phases and from 2 independent cultures (Fig. 1). Less than 1 μg of mRNA was used to establish a cDNA Library with the CloneMiner™ cDNA Library Construction Kit (Invitrogen) according to the manual. First strand synthesis was performed using a Biotin-attB2-Oligo(dT) primer with the following sequence: Biotin-GGCGGCCGCACAACCTTTGTACAAGAAAGTTGGGT(T)₁₉ and SuperScript™ II Reverse Transcriptase. Subsequent to the second-strand synthesis using *Escherichia coli* DNA polymerase, blunt end products were ligated with an attB1 adapter through T4 DNA Ligase. After size fractionation (>500 bp) cDNA-fragments were cloned into the cloning vector pDONR 222 with the BP Clonase enzyme. Plasmids were used to transform ElectroMax DH10B competent cells via electroporation, and random clones were picked for quality control analysis.

2.5. Construction of cDNA libraries after virus infection

Total RNA (4 μg each) was used for the construction of cDNA libraries prepared by vertis Biotechnologie AG (Munich, Germany). From each total RNA poly A⁺ RNA was prepared. With the poly A⁺ RNA first-strand cDNA synthesis was performed using an oligo(dT)-linker primer and M-MLV-RNase H⁻ reverse transcriptase. Synthesis of the second strand was carried out with a random linker primer and Klenow exo-DNA polymerase. The resulting cDNAs were then amplified with 17 (6 and 12 h p.i.) and 16 cycles (24 h p.i.) of LA-PCR (Barnes 1994).

For cloning, the cDNAs were subjected to a limited exonuclease treatment to generate 5' overhangs at both ends of the

cDNAs. After size fractionation on an agarose gel and elution of cDNAs >0.5 kb the cDNAs were directionally ligated into the Eco RI and Bam HI sites of the plasmid vector pBS II sk⁺. Ligations were electroporated into T1 Phage resistant TransforMax™ EC100™-T1^R (Epicentre) electro-competent cells. After transformation, glycerol was added to a final concentration of 15% (v/v).

2.6. EST sequencing

Plasmid DNA was isolated using a standard alkaline lysis procedure, and unidirectional sequencing was accomplished using the M13 HEDGE forward primer (TGA GCG GAT AAC AAT TTC ACA CAG) (Hedge et al., 2000) for the uninfected library and the M13 forward primer (TGT AAA ACG ACG GCC AGT) for the infected libraries, providing sequence from the 5' end of cDNA clones. Sequencing was performed according to the principle of Sanger et al. (1977) using BigDye terminator chemistry from Applied Biosystems.

2.7. Data analysis

For identifying the function of ESTs, sequences were analysed by the program PhyloGena (Hanekamp et al., 2007) on the basis of the SwissProt database. This is a system for an automated phylogenetic annotation of ESTs, genes and genomes. It automatically constructs phylogenetic trees on a per ORF basis and allows annotation on the basis of the function of the neighbouring sequences in the tree. This method is more reliable than simply assuming the function of the “best hit” in a BLAST search. As a threshold for a significant similarity we used an e value of 10^{-7} .

Viral transcripts were identified by BLAST searches of all ESTs against the EhV-86 genome A corresponding analysis verified the origin of all other transcripts from the alga by similar

Table 1
EST assembly results of *E. huxleyi* before virus infection

Sequence name	Redundancy	<i>e</i> value	Putative function
Unknown protein	39	–	–
Fcp-like	10	7.9e–15	Photosynthesis
Fucoxanthin Chlorophyll a/c binding protein (fcp)	3	7.8e–10	Photosynthesis
Light harvesting complex (LHC)	3	3.0e–10	Photosynthesis
Cyclin dependent kinase regulatory subunit, putative	2	3.6e–30	Signal transduction, cell communication, cell cycle
Alpha-glucosidase	2	2.0e–51	Galactose metabolism
Beta-hydrocyacyl-ACP dehydratase	2	8.5e–54	Fatty acid synthesis
Hypothetical conserved protein,	2	5.6e–18	–
Chromosome condensation regulator-like protein	1	7.9e–10	Regulation of the condensation of the chromosomes
Eukaryotic translation initiation factor-like protein	2	2.0e–24	Initiation of translation
Trehalose-6-phosphate-synthase	2	1.0e–29	Starch and saccharose metabolism
Ankyrin related protein	1	5.4e–10	Involved in binding of spectrin at the plasma membrane
Ferredoxin-NADP ⁺ -reductase	1	2.7e–78	Photosynthesis
Glutaredoxin, mitochondrial	1	1.8e–30	Glutathione metabolism
Glyceraldehyde-3-phosphate dehydrogenase (GAPDH)	1	1.6e–117	Glycolysis, gluconeogenesis
N5'-nucleotidase, cytosolic	1	2.6e–20	Transcription
NADH-dehydrogenase	1	2.3e–12	Respiration, fatty acid oxidation
Protein kinase	1	2.3e–18	Signal transduction, phosphorylation
Putative membrane protein	1	5.7e–12	–
Ribosomal protein L3	1	9.5e–37	Protein metabolism
RNA helicase	1	5.7e–48	Transcription

searches against the *E. huxleyi* draft genome, i.e. the trace files of the sequencing runs (<http://www.jgi.doe.gov/>).

3. Results

The aim of this work was to create an initial dataset of the gene expression that occurs during infection of the *E. huxleyi* strain CCMP1516 with the virus EhV-86. Therefore we constructed an EST library from non infected cells and compared it with three EST libraries after virus infection (6, 12 and 24 h).

3.1. Identification and annotation of the ESTs

After elimination of vector- and other problematic sequences, high-quality ESTs with an average length of 520

nucleotides were used for the identification and annotation of the sequences. All ESTs were tested for their origin from the alga and from the virus by BLAST searches against the EhV-86 genome and the *E. huxleyi* draft genome. There was no EST present which did not match either database. All genes were manually annotated. The annotated ESTs had a BLAST *e* value smaller than or equal to 10^{-7} . Identification was based on phylogenetic analysis of all ORFs using PhyloGena. In all libraries the largest fraction of ESTs, both host and virus, were those of unknown function or those not producing any significant hit in BLAST searches.

It was possible to annotate 78 sequences from the uninfected library. A proportion of 17% (13 ESTs) of the sequences encoded for fcp or fcp-like proteins involved in light harvesting and 50% (39 ESTs) were proteins with unknown function.

Table 2
EST assembly results of *E. huxleyi* after 6 h virus infection

Protein	Frequency	<i>e</i> value	Putative function
40S ribosomal protein S11	2	2.0e–42	Protein metabolism
40S ribosomal protein S13	2	9.0e–51	Protein metabolism
60S ribosomal protein L7	2	5.0e–31	Protein metabolism
60S ribosomal protein L25	1	2.0e–43	Protein metabolism
60S ribosomal protein L27	1	2.0e–42	Protein metabolism
Adenylate kinase	1	1.0e–42	Glycolysis; phosphorylation
Elongation factor 1-alpha-like protein (EF1 α)	1	1.0e–66	Protein metabolism
EF-Tu like protein	1	8.0e–67	Protein metabolism
Caltractin-like protein	1	1.0e–21	Calcium binding protein
Eukaryotic translation initiation factor 3 subunit 7	1	1.0e–47	Protein metabolism
Chloroplast ferredoxin-NADP ⁺ -reductase	1	9.0e–93	Photosynthesis
GAPDH cytosolic	1	1.0e–34	Glycolysis, gluconeogenesis
Weakly similar to ubiquitin	1	7.0e–23	Amino acid catabolism
Weakly similar to phosphoribulokinase	1	4.0e–30	Calcin cycle
HSP70-like protein	1	8.0e–38	Stress-induced protein
Phosphoribosylaminoimidazole-succinocarboxamide synthase	1	7.0e–32	Transcription
S-adenosylhomocysteinase	2	8.0e–51	Amino acid catabolism
Unknown protein	43	–	–

Furthermore, proteins were identified involved in photosynthesis, the cell cycle, transcription and protein metabolism (Table 1).

From the library 6 h after virus infection 67 sequences were annotated, from which 64% (43 ESTs) were encoded for unknown proteins. Most of the identified proteins are ribosomal proteins and elongation factors, which are responsible for protein synthesis (Table 2). In this context, expression of HSP70 and *s*-adenosylhomocysteinase are conspicuous because they are an indication for stress. In contrast to the uninfected library, no fcp-like proteins were identified in the EST library generated 6 h post viral infection. However, three genes of unknown function from the virus EhV-86 were identified.

Twelve hours post infection 82 sequences could be annotated. In this EST library, 58.5% (48 ESTs) of the ESTs had no significant match against the SwissProt database in BLAST searches and were classified as proteins of unknown function. However, 12% of the 82 ESTs were ribosomal proteins which are involved in protein metabolism (Table 3). We could identify 5% (4 ESTs) of viral genes of unknown function. The expression of stress proteins like rotamase, RAS-like protein and HSP70 is conspicuous. Fcp-like proteins were again missing.

After 24 h the host viral assemblage transcriptome is dominated by the virus. 80 ESTs of 91 annotated sequences were viral genes. Two of the 11 host genes could be identified as 60S ribosomal protein L8 and GDP-D-mannose 4, 6-dehydratase. Only 10% of the viral genes had significant matches in SwissProt (Table 4). All of the viral proteins of unknown function had a length between 800 and 1100 bp.

By comparing the ESTs divided into different functional categories, we found that before virus infection photosynthesis-related genes dominate in the host, but after 6 h post infection their abundance decreases rapidly in the libraries (Fig. 2). Furthermore, after 6 h virus infection stress-induced host genes were identified. 24 h post infection the viral genes clearly dominate the library. In all libraries a high number of genes of unknown function were found.

4. Discussion

Our data show that it is possible to determine expression profiles throughout a viral infection process using EST libraries. In our case we were lucky to have available genome sequences for both partners, i.e. the virus and the host. Thus we were able to determine the exact number of ESTs from both partners in the libraries. Our data show, that a large proportion of genes active in both partners are of unknown function. These genes would have escaped the analysis if we had focussed on known genes and using classical approaches. It is one of the strengths of the EST approach that also unknown genes are found.

4.1. Possible infection mechanism of the virus EhV-86 in *E. huxleyi*

Our conclusions are based on approx. 320 annotated EST-sequences, but there is already a trend recognizable. Less than 6 h post infection the virus seems to change the expression pattern of *E. huxleyi* significantly. Our data indicate down

Table 3
EST assembly results of *E. huxleyi* after 12 h virus infection

Protein	Frequency	<i>e</i> value	Putative function
40S ribosomal protein S19	2	4.0e–21	Protein metabolism
40S ribosomal protein S19-like	1	6.0e–18	Protein metabolism
60S ribosomal protein L10	2	1.0e–51	Protein metabolism
60S ribosomal protein L8	1	2.0e–66	Protein metabolism
78kDa glucose regulating protein	1	1.0e–38	Stress-induced protein
Adenine phosphoribosyltransferase-like protein	1	1.0e–15	Transcription
ADP-ribosylation factor (Arf)	1	1.0e–40	Stress-induced protein
Elongation factor 1-alpha-like protein (EF1α)	1	6.0e–64	Protein metabolism
Calmodulin	2	1.0e–74	Signaltransduction; calcium sensor
Calmodulin-like protein	2	8.0e–41	Signaltransduction; calcium sensor
Similar to cytosolic GAPDH	1	4.0e–20	Glycolysis, gluconeogenesis
Similar to nuclear GAPDH	1	2.0e–42	Glycolysis, gluconeogenesis
Glycin cleavage protein, mitochondrial	1	2.0e–37	Amino acid catabolism
HSP70-like protein	1	4.0e–29	Stress-induced protein
Nuclear protein with unknown function, transcription regulator?	1	3.0e–32	–
Methionin-aminopeptidase	1	3.0e–44	Protein metabolism
Peptidyl-prolyl <i>cis</i> -trans isomerise (rotamase)	1	1.0e–70	Stress-induced protein
Similar to phosphoribulokinase	1	1.0e–49	Calvin cycle
Mitochondrial processing protease α subunit	1	6.0e–19	Protein metabolism; protein catabolism
Ras-like protein	1	2.0e–21	Stress-induced protein
Ras-related protein, GTPase	1	3.0e–21	Stress-induced protein
<i>S</i> -adenosyl-L homocysteine hydrolase-like protein	1	4.0e–51	Amino acid catabolism
Weakly conserved protein with ATP binding site, similar to elongation factor	1	2.0e–26	Protein metabolism
Stress-activated protein kinase	1	1.0e–47	Stress-induced protein
Ubiquitin	1	2.0e–71	Amino acid catabolism
UDP- <i>N</i> -acetylglucosamine pyrophosphatase	1	5.0e–13	UDP- <i>N</i> -acetylglucosamine biosynthetic process
Unknown protein	48	–	–

Table 4
Identified viral genes after 24 h infection and their putative function

Protein	<i>e</i> value	Putative function
Clp protease (Casein lytic protein)	1.0e–12	Degradation of host proteins
Clp-like protein	4.0e–11	Degradation of host proteins
DNA topoisomerase II	2.0e–43	Packaging of viral DNA; expression of viral genes
DNA dependent RNA Polymerase I	3.0e–06	Priority of expression of viral genes
DNA dependent RNA Polymerase II	2.0e–62	Priority of expression of viral genes
Non histone chromosomal protein Flap endonuclease	7.0e–09 1.0e–48	Packaging of viral DNA Degradation of host nucleic acid
Deoxyuridin 5'-triphosphate nucleotidhydrolase	6.4e–42	Nucleic acid metabolism
Protein kinase	1.0e–12	Regulation

regulation of photosynthesis genes, which is also known in the infection cycle of *Paramecium bursaria* Chlorella Virus-1 (Seaton et al., 1995). Viral infection took place in the middle of the dark phase and the first sample was taken after 6 h, i.e. early in the light phase. During this phase one would expect photosynthesis genes to be up-regulated. However, down-regulation of photosynthesis genes occurred on the benefit of up-regulation of genes related to gene expression and protein synthesis possibly to enhance the expression of viral proteins. The up-regulation of transcription and translation genes of the host could be induced by the virus to facilitate transcription and translation of its genes. Only a few viral transcripts seem to be required for that, because in the 6 h EST library only a minority of 4.5% of the sequences were found to have a viral origin.

It is remarkable that after 6 and 12 h virus infections only a few viral transcripts appear, but that, nevertheless, drastic changes in host gene-expression patterns were induced. After 24 h the transcription of *E. huxleyi* seemingly comes to a standstill. Now the share of viral transcripts in the EST library has reached 90%.

Viral RNA polymerase genes are activated, which may preferably transcribe viral genes or host genes required by the virus. Furthermore, endonuclease and clp-protease appear which could be responsible for the degradation of host DNA and host proteins. The virus has taken over transcription at the latest 24 h after infection, approximately the doubling time of the alga in culture; this could mean that the infection is correlated to the cell cycle of *E. huxleyi*. The reason could be the need of particular host proteins, e.g. for transcription, replication or translation, all of which occur in an organised form during host cell cycle.

4.2. Outlook

More sequences are required to be able to draw more reliable conclusions about changes of the host expression pattern during viral infection. In the ongoing project we have established EST libraries comprising several thousand sequences for *E. huxleyi*. The next step currently under way is to use the sequence information to establish genome arrays for alga and virus and to screen genomes of various *E. huxleyi* and virus strains for genomic differences, and subsequently for transcriptome differences.

Twelve hours after infection, the virus affects the host expression pattern of *E. huxleyi* significantly, but after 24 h the effect was drastic. During this 12 h period there seems to occur a lot of change in the expression of the host and the virus. To get further insight into the infection cycle, it would be of particular interest to construct and analyse a cDNA library after 14–18 h

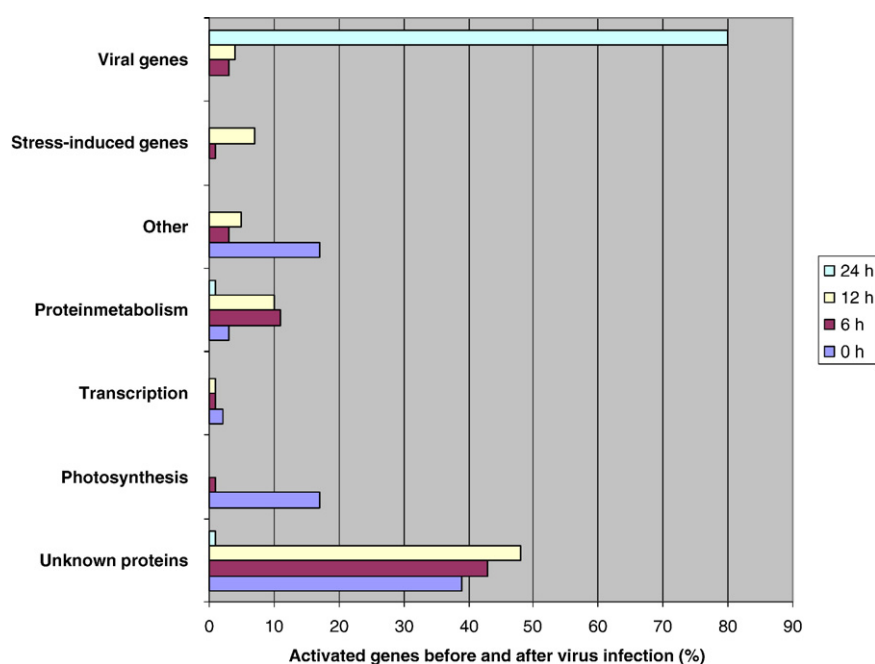


Fig. 2. EST expression profile of *Emiliania huxleyi* before and during virus infection divided into different functional categories based on their putative function (%).

of virus infection, because the virus, described by Castberg et al. (2002) has a latency of 12–14 h.

Acknowledgements

This work was supported by Marine Genomics Europe (EU contract no. 505403). We would like to thank Matthew Hall for technical assistance and Linda Medlin for support.

References

- Allen, M.J., Schroeder, D.C., Holden, M.T.G., Wilson, W.H., 2006. Evolutionary history of the Coccolithoviridae. *Mol. Biol. Evol.* 23, 86–92.
- Allen, M.J., Martinez-Martinez, J., Schroeder, D.C., Somerfield, P.J., Wilson, W.H., 2007. Use of microarrays to assess viral diversity: from genotype to phenotype. *Environ. Microbiol.* 9, 971–982.
- Balch, W.M., Holligan, P.M., Ackleson, S.G., Voss, K.J., 1991. Biological and optical properties of mesoscale coccolithophore blooms in the Gulf of Maine. *Limnol. Oceanogr.* 36, 629–643.
- Balch, W.M., Holligan, P.M., Kilpatrick, A.K., 1992. Calcification, photosynthesis and growth of the bloom-forming coccolithophore, *Emiliania huxleyi*. *Cont. Shelf Res.* 12, 1353–1374.
- Barnes, W.M., 1994. PCR amplification of up to 35-kb DNA with high fidelity and high yield from λ bacteriophage templates. *Proc. Natl. Acad. Sci. USA* 91, 2216–2220.
- Baumann, K.H., Böckel, B., Frenz, M., 2004. Coccolith contribution to South Atlantic carbonate sedimentation. In: Thierstein, H.R., Young, J.R. (Eds.), *Coccolithophores*. Springer-Verlag, Berlin Heidelberg, pp. 367–402.
- Bergh, O., Borsheim, K.Y., Bratbak, G., Heldal, M., 1989. High abundance of viruses found in aquatic environments. *Nature* 340, 467–468.
- Bratbak, G., Egge, J.K., Heldal, M., 1993. Viral mortality of the marine alga *Emiliania huxleyi* (Haptophyceae) and termination of algal blooms. *Mar. Ecol. Prog. Ser.* 93, 39–48.
- Bratbak, G., et al., 1995. Viral activity in relation to *Emiliania huxleyi* blooms: a possible mechanism of DMSP release? *Mar. Ecol. Prog. Ser.* 128, 133–142.
- Bratbak, G., Wilson, W., Heldal, M., 1996. Viral control of *Emiliania huxleyi* blooms? *J. Mar. Syst.* 9, 75–81.
- Brown, C.W., Yoder, J.A., 1994. Blooms of *Emiliania huxleyi* (Prymnesiophyceae) in surface waters of the Nova Scotian shelf and the Grand Bank. *J. Plankton. Res.* 15, 1429–1438.
- Brussaard, C.P.D., Kempers, R.S., Kop, A.J., Riegman, R., Heldal, M., 1996. Virus like particles in a summer bloom of *Emiliania huxleyi* in the North Sea. *Aquat. Microb. Ecol.* 10, 105–113.
- Castberg, T., et al., 2001. Microbial population dynamics and diversity during a bloom of the marine coccolithophorid *Emiliania huxleyi* (Haptophyta). *Mar. Ecol. Prog. Ser.* 221, 39–46.
- Castberg, T., et al., 2002. Isolation and characterization of a virus that infects *Emiliania huxleyi* (Haptophyta). *J. Phycol.* 38, 767–774.
- Charlson, R.J., Lovelock, J.E., Andreae, M.O., Warren, S.G., 1987. Oceanic phytoplankton, atmospheric sulphur, cloud albedo and climate. *Nature (Lond.)* 326, 655–661.
- Eide, L.D., 1990. Distribution of coccolithophorids in surface sediments in the Norwegian-Greenland Sea. *Mar. Micropaleontol.* 16, 65–75.
- Fuhrmann, J.A., 1999. Marine viruses and their biogeochemical and ecological effects. *Nature (Lond.)* 399, 541–548.
- Guillard, R.R.L., 1975. Culture of phytoplankton for feeding marine invertebrates. In: Smith, W.L., Chanley, M.H. (Eds.), *Culture of Marine Invertebrate Animals*. Plenum Press, New York, pp. 29–60.
- Hanekamp, K., Bohnbeck, U., Beszteri, B., Valentin, K., 2007. PhyloGena — a system for automated phylogenetic annotation of (algal) expressed sequence tags (EST), genes and genomes. *Bioinformatics* 23 (7), 793–801. doi:10.1093/bioinformatics/btm016.
- Hay, W.W., Mohler, H.P., Roth, P.H., Schmidt, R.R., Boudreaux, J.E., 1967. Calcareous nannoplankton zonation of the Gulf Coast and Caribbean–Antillean area and transatlantic correlation. *Trans. Gulf Coastal Assoc. Geol. Soc.* 17, 428–480.
- Hedge, P., et al., 2000. A concise guide to cDNA microarray analysis. *BioTechniques* 29, 548–562.
- Holligan, P.M., Viollier, M., Harbour, D.S., Camus, P., Champagne-Philippe, M., 1983. Satellite and ship studies of coccolithophore production along a continental shelf edge. *Nature (Lond.)* 304, 339–342.
- Holligan, P.M., et al., 1993. A biogeochemical study of the coccolithophore *Emiliania huxleyi* in the north Atlantic. *Glob. Biogeochem. Cycles* 7, 879–900.
- Jacobsen, A., Bratbak, G., Heldal, M., 1996. Isolation and characterization of a virus infecting *Phaeocystis pouchetii* (Prymnesiophyceae). *J. Phycol.* 32, 923–927.
- Jaquet, S., Heldal, M., Iglesias-Rodriguez, D., Larsen, A., Wilson, W., Bratbak, G., 2002. Flow cytometric analysis of an *Emiliania huxleyi* bloom terminated by viral infection. *Aquat. Microb. Ecol.* 27, 111–124.
- Lohmann, H., 1902. Die Coccolithophoridae, eine Monographie der Coccolithen bildenden Flagellaten. Zugleich ein Beitrag zur Kenntnis des Mittelmeerauftriebs. *Arch. Protistenk.* 1, 89–165.
- Malin, G., Liss, P.S., Turner, S.M., 1994. Dimethyl sulphide: production and atmospheric consequences. In: Green, J.C., Leadbeater, B.S.C. (Eds.), *The haptophyte algae*. Syst. Assoc. Spec. Vol., vol. 51. Clarendon Press, Oxford, pp. 303–320.
- Nagasaki, K., Yamaguchi, M., 1997. Isolation of a virus infectious to the harmful bloom causing microalga *Heterosigma akashiwo* (Raphidophyceae). *Aquat. Microb. Ecol.* 13, 135–140.
- Nagasaki, K., Ando, M., Imai, I., Itakura, S., Ishida, Y., 1994a. Virus-like particles in *Heterosigma akashiwo* (Raphidophyceae): a possible red tide disintegration mechanism. *Mar. Biol.* 119, 307–312.
- Nagasaki, K., Ando, M., Itakura, S., Imai, I., Ishida, Y., 1994b. Viral mortality in the final stage of *Heterosigma akashiwo* (Raphidophyceae) red tide. *J. Plankton Res.* 16, 1595–1599.
- Rafalski, J.A., et al., 1998. New experimental and computational approaches to the analysis of gene expression. *Acta Biochim. Pol.* 45, 929–934.
- Samtleben, C., Bickert, T., 1990. Coccoliths in sediment traps from the Norwegian Sea. *Mar. Micropaleontol.* 16, 39–64.
- Sandaa, R.A., Heldal, M., Castber, T., Thyraug, R., Bratbak, G., 2001. Isolation and characterization of two viruses with large genome size infecting *Chrysochromulina ericina* (Prymnesiophyceae) and *Pyramimonas orientalis* (Prasinophyceae). *Virology* 290, 272–280.
- Sanger, F., Nicklen, S., Coulson, A.R., 1977. DNA-sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74, 5463–5467.
- Schmitt, A.O., Specht, T., Beckmann, G., Dahl, E., Pilarsky, C.P., Hinzmann, B., Rosenthal, A., 1999. Exhaustive mining of EST libraries for genes differentially expressed in normal and tumour tissues. *Nucl. Acids Res.* 27, 4251–4260.
- Schroeder, D.C., Oke, J., Malin, G., Wilson, W.H., 2002. Coccolithovirus (Phycodnaviridae): characterization of a new large dsDNA algal virus that infects *Emiliania huxleyi*. *Arch. Virol.* 147, 1685–1698.
- Schroeder, D.C., Oke, J., Hall, M., Malin, G., Wilson, W.H., 2003. Virus Succession Observed during an *Emiliania huxleyi* Bloom. *Appl. Environ. Microbiol.* 69, 2484–2490.
- Seaton, G.G.R., Lee, K., Rohozinski, J., 1995. Photosynthetic shutdown in *Chlorella* NC64A associated with the infection cycle of *Paramecium bursaria* Chlorella Virus-1. *Plant Physiol.* 108, 1431–1438.
- Suttle, C.A., 2000. Ecological, evolutionary, and geochemical consequences of viral infection of cyanobacteria and eukaryotic algae. In: Hurst, C.J. (Ed.), *Viral Ecology*. Academic Press, New York, pp. 247–296.
- Van Etten, J.L., Meints, R.H., 1999. Giant viruses infecting algae. *Annu. Rev. Microbiol.* 53, 447–494.
- Van Etten, J.L., Lane, L.C., Meints, R.H., 1991. Viruses and virus-like particles of eukaryotic algae. *Microbiol. Rev.* 55, 586–620.
- Van Etten, J.L., Graves, M.V., Müller, D.G., Boland, W., Delaroque, N., 2002. Phycodnaviridae — large DNA algal viruses. *Arch. Virol.* 147, 1479–1516.
- Westbroek, P., et al., 1993. A model system approach to climate forcing. The example of *Emiliania huxleyi*. *Glob. Planet Change* 8, 27–46.
- Westbroek, P., et al., 1994. *Emiliania huxleyi* as a key to biosphere–geosphere interaction. In: Green, J.C., Leadbeater, B. (Eds.), *The Haptophyte Algae*. Clarendon Press, Oxford, pp. 321–334.
- Wilhelm, S.W., Suttle, C.A., 1999. Viruses and nutrients cycles in the sea — viruses play critical roles in the structure and function of aquatic food webs. *Bioscience* 49, 781–788.

- Wilson, W.H., Turner, S., Mann, N.H., 1998. Population dynamics of phytoplankton and viruses in a phosphate-limited mesocosm and their effect on DMSP and DMS production. *Estuar. Coast. Shelf Sci.* 46, 49–59.
- Wilson, W.H., Tarran, G.A., Zubkov, M.V., 2002a. Virus dynamics in a coccolithophore dominated bloom in the north Sea. *Deep Sea Res.* 49, 2951–2963.
- Wilson, W.H., Tarran, G.A., Schroeder, D.C., Cox, M., Oke, J., Malin, G., 2002b. Isolation of viruses responsible for the demise of an *Emiliania huxleyi* bloom in the English Channel. *J. Mar. Biol. Ass. U.K.* 82, 369–377.
- Wilson, W.H., et al., 2005. Complete genome sequence and lytic phase transcription profile of a Coccolithovirus. *Science* 309, 1090–1092.
- Winter, A., Jordan, R.W., Roth, P.H., 1994. Biogeography of living coccolithophores in ocean waters. In: Winter, A., Siesser, W.G. (Eds.), *Coccolithophores*. Cambridge University Press, Cambridge, UK, pp. 161–177.
- Wommack, K.E., Colwell, R.R., 2000. Virioplankton: viruses in aquatic ecosystems. *Microbiol. Mol. Biol. Rev.* 64, 69–114.