An assessment of RNA content in *Prymnesium parvum*, *Prymnesium polylepis*, cf. Chattonella sp. and *Karlodinium veneficum* under varying environmental conditions for calibrating an RNA microarray for species detection

| Title | An assessment of RNA content in *Prymnesium parvum*, *Prymnesium polylepis*, cf. Chattonella sp. and *Karlodinium veneficum* under varying environmental conditions for calibrating an RNA microarray for species detection |
| Author(s) | McCoy, Gary R.; Touzet, Nicolas; Fleming, Gerard T.; Raine, Robin |
| Publication Date | 2014-02-03 |
| Publisher | Wiley |
| Link to publisher's version | http://dx.doi.org/10.1111/1574-6941.12277 |
| Item record | http://hdl.handle.net/10379/4154 |
| DOI | http://dx.doi.org/DOI: 10.1111/1574-6941.12277 |
An assessment of RNA content in *Prymnesium parvum*, *Prymnesium polylepis*, cf. *Chattonella* sp. and *Karlodinium veneficum* under varying environmental conditions for calibrating an RNA microarray for species detection

Gary R McCoy¹, Jessica U. Kegel², Nicolas Touzet¹,³, Gerard TA Fleming¹, Linda K Medlin² and Robin Raine¹

¹ The Ryan Institute, National University of Ireland, Galway, Ireland
² Marine Biological Association of the UK, The Citadel, Plymouth, PL1 2PB UK
³ Current address: School of Science, Centre for Environmental Research, Innovation and Sustainability, Institute of Technology Sligo; Sligo, Ireland

Corresponding author: Gary McCoy

Martin Ryan Building,
National University of Ireland, Galway,
University Road,
Galway

Tel: +353 (0)91 493231
Fax: +353 (091) 525005
Email: g.mccoy3@nuigalway.ie

Keywords: Microflagellates, RNA, calibration curves, microarrays, MIDTAL.

Abstract

Traditional methods of identification and enumeration can be somewhat ambiguous when identifying phytoplankton that requires electron microscopic examination to verify specific

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/1574-6941.12277

This article is protected by copyright. All rights reserved.
morphological features. Members of the genus *Prymnesium* (division Haptophyta), members of the Raphidophyceae and naked dinoflagellates are examples of such phytoplankton whose identification can be difficult. One alternative to traditional microscopy-based methods of identification is to use molecular protocols to detect target species. Methods that measure cellular DNA and RNA content can be used to estimate the number of cells present in a sample. This study investigated the variation of RNA yields in *Prymnesium parvum*, *Prymnesium polylepis*, cf. *Chattonella* sp. and *Karlodinium veneficum* cells grown under different light, temperature, salinity, and inorganic nutrient conditions. This information was used to calibrate the signal intensity of a variety of oligonucleotide probes spotted onto the MIDTAL (Microarrays for the Detection of Toxic Algae) microarray, which is being developed to aid national monitoring agencies and to provide a faster means of identifying and quantifying harmful phytoplankton in water column samples.

**Introduction**

Phytoplankton monitoring programmes usually include the identification and enumeration of harmful species in water samples. Traditional methods, such as light microscopy, are typically used, whereby only a small sub-sample of the phytoplankton assemblage is examined (Karlson *et al.*, 2010). Traditional methods of phytoplankton identification and enumeration can be somewhat ambiguous when attempting to identify naked flagellates. Their cell body can be greatly distorted by preservative reagents, which makes their identification almost impossible; this is particularly true for the raphidophytes (Band-Schmidt *et al.*, 2012). More advanced techniques, such as Transmission Electron Microscopy (TEM), are hence usually needed for accurate identification. Molecular methods offer a way of overcoming these problems together with reducing the time needed for analysis. Some of these molecular methods include fluorescent *in-situ* hybridisation (FISH, Scholin *et al.*).
al., 1997; Groben & Medlin, 2005; Touzet et al., 2010), sandwich hybridisation (Scholin & Anderson, 1998), quantitative PCR (Galluzzi et al., 2008; Touzet et al., 2009; Kavanagh et al., 2010), high-throughput sequencing technologies (Logares et al., 2013; Zhan et al., 2013) and microarrays (Metfies & Medlin, 2005).

The most challenging aspects of molecular methods relate to quantification (Metfies & Medlin, 2008). This is an important aspect relating to the monitoring of harmful algae in particular as threshold levels apply to cell densities as well as toxin levels in certain national monitoring programmes (Karlson et al., 2010). This problem also extends beyond harmful algae monitoring, where other environmental studies have used molecular techniques to identify and estimate abundances of mixed bacterial communities in food, soil and water (Call et al., 2003; Loy et al., 2005; Lee et al., 2008; Kostić et al., 2010).

In the FP7 EU funded project MIDTAL (Microarrays for the Detection of Toxic Algae, http://www.midtal.com), a molecular method based on microarray technology was employed for detecting and quantifying HAB species to help aid national monitoring agencies. Microarray signals from species level probes bound to fluorescently-labelled RNA extracts were used to estimate the number of cells present in culture and field samples. Here we investigated whether or not RNA content in microalgal cells varied when cultures were exposed over a four day period to different environmental stress conditions. The four microflagellates selected for this study were the haptophytes *Prymnesium parvum* and *P. (= Chrysochromulina) polylepis*, an unidentified species closely resembling a raphidophyte, here referred to as cf. Chattonella sp., and the naked dinoflagellate *Karlodinium veneficum*. Results were used to calibrate the microarray signals associated for several oligonucleotide probes targeting these species spotted on the MIDTAL microarray.

This article is protected by copyright. All rights reserved.
An existing hierarchical group of oligonucleotide probes specific for the haptophytes *P. parvum* and *P. polylepis* were lengthened to 25 nucleotides and were spotted onto the microarray chip. New probes, also 25 nucleotides in length, were designed for cf. *Chattonella* sp. and *K. veneficum*. The testing of these species against various stress conditions to determine the RNA content per cell was necessary to calibrate the microarray signal so it can be used as a quantitative tool for estimating cell numbers.

**Material and methods**

**Algal strains**

*Prymnesium parvum* and *Prymnesium polylepis* strains were purchased from CCMP (Provasoli-Guillard National Centre for Culture of Marine Phytoplankton (USA), Bigelow Laboratory for Ocean Sciences), SAG (Culture Collection of Algae (Sammlung von Algenkulturen der Universität Göttingen) or kindly provided by Bente Edvardsen (Department of Biology, University of Oslo, Oslo; Table S1). *Karlodinium veneficum* strains were purchased from CMSTAC (Center for Marine Science Toxic Algal Collection, University of North Carolina), CCCM (Canadian Centre for the Culture of Microorganisms) and PCC (The Pasteur Culture Collection of Cyanobacteria; Table S1). Strains and sequences of cf. *Chattonella* sp. were provided by Carmelo Tomas for probe development and testing of the microarray. Algal cultures were maintained in f/2 (Guillard, 1983, with silicate) or IMR½ algal media (without silicate and with 10 nM selenite) (Eppley et al., 1967) at 15 ± 1°C, under a white fluorescent light with a photon flux of 150 µE m⁻² s⁻¹ and a 14:10 light/dark cycle. The experimental setup was performed as part of an EU consortium; thus, there were slightly different aspects to the execution of the experiments because they were performed in different laboratories with different types of laboratory conditions and equipment. The experiments for *Prymnesium* sp. were conducted at the National University of Ireland,
Galway (NUIG), and cf. *Chattonella sp.* and *K. veneficium* at the Marine Biological Association (MBA) Plymouth, UK.

**Cell Counts**

Sub-samples for cell counts were kept at 4°C and fixed with either 5 µl of Lugol’s iodine solution added to 500 µl of culture or 200 µl of glutaraldehyde 0.2 % to 1 ml of the culture mix. Concentrations of cells were determined either under light microscopy using the Fuchs-Rosenthal chamber method or with a flow cytometer (Becton Dickinson FACSCalibur, BD) after adding a known amount of beads (4.8 µm, caliBRITE Beads BD Biosciences).

**Experimental setup**

Stock cultures of each strain were grown in 500 ml conical flasks enriched with 300 ml of fresh medium under the previously described conditions. Every day cell counts were performed on each strain to determine when culture growth was in exponential stage. At that stage, a known amount of culture was transferred into 200 ml conical flasks containing 100 ml of fresh medium modified to the stress condition being tested. For each stress condition, the flasks at time zero (T0) were inoculated separately with a set cell concentration, ca. 75,000 cells/ml for the three individual *P. polylepis* strains and ca. 250,000-400,000 cells/ml for the three individual *P. parvum* strains. For cf. *Chattonella sp.* and *K. veneficum* species, 20 ml of each stock culture from the three strains, which represented about 20 cells in the former and 26-30 cells in the latter, were inoculated separately into the 200 ml tissue culture flasks. The control (optimal) cultures were grown as the initial stock culture.

Light intensity, temperature stress, salinity modulation and nutrient depletion were tested, one parameter being changed with each set of cultures (Table S2). For the stress...
conditions, the strains were inoculated in flasks containing f/2 or IMR½ at lower (LS) and higher (HS) stress than the control conditions (Table S1). Nitrogen and phosphorous depletion was carried out using modified f/2 Guillard or IMR½ algal seawater medium with the absence of nitrate (N- condition) or the absence of phosphate (P- condition). Although nitrate or phosphate was not added to the medium, the presence of these compounds could have been present in trace amounts in the seawater stock and carryover from the inoculums. An additional experiment performed at NUIG consisted of inoculating all three strains from *Prymnesium* species together (ALL mix) across the four conditions tested and carried out along with the experiments based on the individual *Prymnesium* strains. This was done to verify whether or not the RNA content of strains grown together (ALL mix) were significantly different from individual strains grown separately under the above stress conditions because it is assumed that natural populations contain a mixture of genotypes.

Exposed cultures were incubated for 24 hours (T1), 48 h (T2) and 72 h (T3) after which a known volume between 10-15 ml was taken from each flask for RNA extraction. Samples were centrifuged and the algal pellet collected for each time point, with the exception of cf. *Chattonella sp.* and *K. veneficum* because they grew very slowly and only a single time point (T2) was taken. For the cf. *Chattonella sp.* and *K. veneficum* cultures, the volumes collected for RNA extraction were pooled together for the three strains of each species so as to optimise extraction efficiency because of the low numbers of cells.

**RNA extraction**

A volume of 5 to 15 ml from each test culture was transferred to 15 ml tubes, which were centrifuged at 6,000 rcf for 10 minutes. The supernatant was then removed to leave around 2 ml of sample. Tubes were centrifuged a second time at 6,000 rcf for 5 minutes and the remaining supernatant completely removed using a micropipette and a vacuum pump,
without disturbing the pellet. TRI-Reagent (1 ml) was immediately added to each pellet, homogenised and transferred to 2 ml screw cap tubes containing 0.1 ml of acid washed glass beads (213-300 µm). If RNA extraction was not performed on the same day, samples were stored at -80°C until further processing.

RNA extractions were performed as described in the MIDTAL RNA extraction protocol (Lewis et al., 2012) with modifications as follows. The 2 ml screw cap tubes were incubated at 60°C for 10 minutes at maximum speed on a thermoshaker and vortexing twice for 20 seconds during incubation. An aliquot (100 µl) of 1-bromo-3-chloro-propane (BCP:Sigma) or 200 µl chloroform was added to each tube, the mixture was vortexed for 15 seconds and left to settle for 5 minutes. The whole content of the tube was transferred to a pre-spun 2 ml heavy phase lock (PL) tubes (5-PRIME; 12,000g for 30 sec), which was then homogenised manually for 15 seconds and allowed to stand for a further 5 minutes at room temperature. The tubes were centrifuged at 4°C for 15 minutes at 12,000g and the upper aqueous phase (~500-550 µl) from the PL tubes was transferred to a new 1.5 ml RNase-free tube. An equal volume of isopropanol was added (500 µl) prior to vortexing for 15 seconds. The tube was then incubated at -20°C for 1 hour, centrifuged again for 15 minutes and the supernatant carefully removed without disturbing the RNA pellet using a micropipette. The RNA pellet was washed with 1 ml of 75% ethanol, centrifuged and the supernatant was completely removed. After the final centrifugation step, the pellet was air dried for 3-5 min while being kept on ice. RNA pellets from each sample were resuspended in 25-50 µl of RNase free water. An aliquot of suspended RNA sample was taken to determine its RNA concentration using a NanoDrop 1000 Spectrophotometer. The total RNA amount (ng) was then related to the amount of cells processed in the sample, providing RNA yields (pg/cell) for all the varying environmental conditions tested. The samples were stored at -80°C until later use.
RNA clean-up and labelling

An ammonium acetate precipitation step was added to improve RNA quality and labelling efficiency. The eluted RNA was defrosted on ice and 0.5 volume of 7.5 M ammonium acetate (NH4Ac) and 2 volumes of ethanol (EtOH absolute, stored at -20°C) were added. RNA precipitation steps are described in detail elsewhere (Kegel et al., 2013). RNA labelling and fragmentation were as described in Lewis et al. (2012). Prior to labelling 10 ng of *Dunaliella tertiolecta* RNA was added as an internal control to each eluted RNA sample being tested for *Prymnesium* spp., cf. *Chattonella* spp and *Karlodinium*.

Microarray calibration

Sequences of *Prymnesium* spp., cf. *Chattonella* sp. and *K. veneficum* were analysed *in silico* using ARB (Ludwig et al., 2004) to design specific probes in those instances where published FISH probes were not available. Probes for *Prymnesium* originally designed for a FISH format and for higher taxonomic levels in the Haptophyta division were lengthened to 25 nucleotides in length, with the exception of probes Clade01old_25_dT and PparvD01_25_dT (Table S3). The probe sequences for all probes designed or modified from FISH probes for the entire project for the MIDTAL microarray are patent pending as a universal microarray for the detection of toxic algae and the entire hybridisation kit including the array and all necessary reagents are commercially available from Microbia Environement (France).

Four different amounts of CY5-labelled (cyanine-5) RNA (1 ng, 5 ng, 25 ng and 100 ng) for cf. *Chattonella* spp. and *Karlodinium*, were hybridised on the third generation MIDTAL microarray slides (SCHOTT nexterion) to create calibration curves normalised to the TATA box protein control (POSITIVE_25_dT), with the exception of *Prymnesium* spp. for which only two RNA amounts were used (25 ng and 100 ng). Calibration curves

This article is protected by copyright. All rights reserved.
normalised to the *Dunaliella* control (DunGS02_25_dT_dT) were also generated with two different amounts of CY5-labelled RNA (25 ng and 100 ng) for *Prymnesium* spp., cf. *Chattonella* spp. and *Karlodinium*. All hybridisation mixtures contained 30 µl of 2x hybridisation buffer, 3 µl Poly-dA (1µM), 5 ng of TBP-control and were adjusted to 60 µl with nuclease-free water.

Hybridisations were carried out as in chapter 9 in Lewis et al., (2012) with some modifications, which included a pre-blocking and washing step of the microarray slide by shaking for 20 minutes at 50°C, denaturing of the hybridisation mixture for 10 minutes at 94°C and hybridisation to the slide for 60 minutes at 65°C. After hybridisation, slides were washed with washing buffers (SSC/EDTA/SDS) at room temperature, followed by a final wash at 50°C (Lewis et al., 2012; Kegel et al., 2013). Finally, the slides were scanned (Perken Elmer Microarray Scanner or GenePix 4000B, Molecular Devices) and total signals were determined as the average of the feature-background ratio of all 8 spots for each probe using the GenePix 6.0 software programme. Further analysis was carried out with the GPR-Analyzer ver. 1.24 (Dimatti & Edvardsen, 2013). Signal intensity were normalised to the internal control probes spotted on the microarray to allow comparison of signal strength between slides.

**Statistical analysis**

For the RNA stress experiments a two-way ANOVA was used followed by Bonferroni post test analysis to detect significant differences between each treatment at each period of time, unless stated otherwise. Linear regression analyses and Pearson’s correlation tests were carried out for 1) amount of RNA and cell number and 2) microarray signal and cell number relationships. The slope of the linear regression analysis was used in the GPR-
Analyzer program to infer cell numbers from field material. All statistical analysis was carried out in GraphPad Prism 5.

**Growth rates**

Growth rate was calculated from the specific growth rate (K') equation $K' = \frac{\ln (N2 / N1)}{(t2 - t1)}$, where N1 and N2 correspond to cell concentration at time1 (t1) and time2 (t2) (Levasseur *et al.*, 1993).

**Results**

*Prymnesium parvum* and *P. polylepis*

Comparison of the growth rate for all *Prymnesium* strains under the range of environmental stresses applied (light, temperature, salinity and nutrients) is shown in Figs. 1a and 1b. In general, there was an increase in cell numbers even in sub-optimal conditions taken from T0 to T3 (72 h period). However, this was not the case for *P. polylepis*, strain UiO038, for which cell numbers from the initial inoculation fell by up to 150 fold, and growth decreased by the end of the test period in salinity and nutrient culture conditions. A similar result was observed with the *P. polylepis* culture condition containing a mixture of all the strains (UiO037, UiO038, and CCMP1757) grown together. However, cell numbers did not decrease (up to 8.3 times for salinity and 6.5 times for nutrient depletion) to the extent that occurred when UiO038 strain cultures were grown separately. This had an effect on determining RNA content per cell because the strains that had a declining growth rate showed a higher standard deviation in RNA yield (pg/cell Table 1). Correlation between growth rate (d$^{-1}$) and average RNA yield (pg/cell) are shown in Fig. 2. In general, the culture conditions had a non-significant negative correlating growth rate (d$^{-1}$) when compared with average RNA yields (pg/cell). However, *P. parvum* strains grown under nutrient ($r = 0.7895; p =$

This article is protected by copyright. All rights reserved.
treatment and *P. polylepis* strains grown under salinity (*r* = -0.9277; *p* < 0.0001; *n* = 36) and nutrient (*r* = -0.5790; *p* = 0.0485; *n* = 36) treatments had a significant correlation (Fig. 2d, 2g & 2h). The only treatments to be positively correlated were *P. parvum* and *P. polylepis* strains grown under nutrient and temperature treatments, respectively (Fig. 2d & 2f).

*Prymnesium parvum* showed no significant change in RNA yield (pg/cell) under light, temperature and salinity variation; but, there was a significant change (*P* = 0.006) between the treatments for nutrient stress but not for treatments over time (Fig. 3). The average RNA content for *P. parvum* for the whole data set across three strains (UiO054, CCMP and SAG) and all conditions tested was 0.57 ± 0.16 pg/cell (*n* = 108; Table 1).

*Prymnesium polylepis* showed no significant change in RNA yield (pg/cell) under any of the environmental conditions (Fig. 3), with an average RNA content of 1.33 ± 0.67 pg/cell for the whole data set across the three strains (UiO036, UiO037 and CCMP) for all the conditions tested (*n* = 90; note excludes UiO 038 strain Fig. 3).

The difference in average RNA yield (pg/cell) between the three strains grown individually and ALL mix for stress conditions for *P. parvum* under light and temperature and for *P. polylepis* under light, temperature, salinity and nutrient conditions was not significant (*P* > 0.05; Figs. 4a & 4b). However, there was a significant difference in the average RNA yields (pg/cell) between the temperature and salinity conditions with *P. parvum* (*P* < 0.05; Fig. 4a). The average RNA yield (pg/cell) obtained from the ALL mix for *P. parvum* strains (*P. parvum* ALL mix = 0.64 pg/cell, s.d. = 0.18) or a combination of three *P. polylepis* ALL mix (*P. polylepis* ALL mix = 2.85 pg/cell, s.d. = 2.80) from each stress condition was also determined (Fig. 4a & 4b).
P. polylepis strain UiO038 used in the salinity and nutrient stress experiments was not very adaptive to changing environments and this resulted in cell numbers from initial inoculation decreasing (10 fold decrease) in numbers over the stress period of 72 h. For this reason, P. polylepis strain UiO 038 was not included in the salinity and nutrient results shown in Fig. 3. Therefore, to compare average RNA yields (pg/cell) between the individual strains and the strains grown together (ALL mix), the UiO038 strain data was included (Fig. 4b).

There was a positive linear relationship (P. parvum R^2 = 0.76; P. polylepis R^2 = 0.74) between cell numbers and total RNA amounts (ng) in both Prymnesium species (Figs. 5a & 5c). The correlation was significant between cell numbers and total RNA amounts (ng) (P. parvum r = 0.8717 p< 0.0001, (n = 108); P. polylepis r = 0.8578 p < 0.0001 (n = 108)) for all the separate strains across each stress condition (Figs. 5a & 5c). There was also a positive linear relationship (P. parvum All mix R^2 = 0.72; P. polylepis All mix R^2 = 0.70) between cell numbers and total RNA amounts (ng) in both Prymnesium ALL mix species (Figs. 5b & 5d). The correlation was significant between cell numbers and total RNA amounts (ng) (P. parvum (All mix) r = 0.85 p < 0.0001, (n = 35); P. polylepis (All mix) r = 0.84 p < 0.0001 (n = 35)) for all the strains grown together across each stress condition (Figs. 5b & 5d).

Cf. Chattonella sp.

There was an increase in growth rate (d\(^{-1}\)) in all experimental culture conditions from T0 to T2 (48 h period) for all three strains of cf. Chattonella (Fig. 1c). All culture conditions showed a negative non-significant correlation between growth rate (d\(^{-1}\)) and RNA yield (pg/cell; Fig. 6a to 6d); except for salinity and nutrients which were significant (r = -0.6697; p = 0.0485; n = 9 and r = -0.6977; p = 0.0366; n = 9, respectively; Figs. 6c & 6d).

This article is protected by copyright. All rights reserved.
Cf. *Chattonella* *sp.* showed no significant change (p > 0.05) in RNA yield (pg/cell) between strains across the treatments of light, temperature, salinity, and nutrient experiments taken at the 48 h period (Fig. 7). The average RNA content for cf. *Chattonella* for the entire data set across three strains (CMSTAAC300, CMSTAC305 and CMSTAC307) and conditions was 17.82 ± 8.81 pg/cell (n = 36; Table 1).

There was a positive linear relationship (R² = 0.58) between cell numbers and total RNA amounts (ng) for cf. *Chattonella* species (Fig. 5e). The correlation between cell numbers and total RNA amounts (ng) (r = 0.76 p < 0.0001, (n = 36)) of all the strains across each stress condition was significant (Fig. 5e).

**Karlodinium veneficum**

Growth rates (d⁻¹) for *Karlodinium veneficum* were estimated from only one time period at 48 hours (Fig. 1d). There was an increase in cell numbers in all experimental culture conditions from T0 to T2 (48 h period) except for strains CCCM734 (L1) and PCC517 (L3) at low salinity (Fig. 1d). All tests showed a negative correlation between growth rate and RNA yields (pg/cell; Fig. 6e to 6g), with salinity only being significant (r = -0.8483; p = 0.0078; n = 9; Fig. 6g).

*Karlodinium veneficum* showed no significant change in RNA yield (pg/cell) between strains and various treatments of light, salinity and nutrient stress during the 48 hour period (Fig. 7). However there was a significant difference between treatments for temperature (p = 0.0070), but not between the strains. Due to only one period of 48 h being available, there are no replicate values to provide standard errors between strains. Therefore, to see if differences between the temperature treatments were significant, the RNA yields for the three different strains (CCCM734, PCC709 and PCC517) were averaged and analysed by a 1 way ANOVA followed by a Tukey multiple comparison test. The results showed a significant difference (p
= 0.0169) only between the low (10 °C) temperature treatment with either optimal (15 °C) or high (20 °C) treatments. The average RNA content for *K. veneficum* for the whole data set across three strains (CCCM734, PCC709 and PCC517) and conditions was 44.99 ± 43.85 pg/cell (n = 36; Table 1).

The linear relationship ($R^2 = 0.29$) between cell numbers and RNA yields for *K. veneficum* and correlation ($r = 0.53 \ p < 0.0008$, (n = 36)) of all the strains across each stress condition was significant (Fig. 5f).

**Microarray calibration curves**

Higher taxon probes specific to the hierarchy of a given species had to produce a positive signal for a species level signal to be considered positive. Probe signal intensities on the microarray were recorded as positive when they produced a signal-to-noise ratio $\geq 2$. As the higher taxon probes typically gave a greater intensity than the corresponding species level probe, false positive results were thus eliminated (Table S3).

Calibration curves in this study are represented as cell numbers of target species normalised to the microarray signals obtained from the two control probes POSITIVE_25_dT (TATA-box protein) and DunGS02_25_dT_dT (*D. tertiolecta*). The calibration curves for *P. parvum*, *P. polylepis*, *cf. Chattonella* and *K. veneficum* were generated using increasing amounts of labelled RNA (1 ng, 5 ng, 25 ng and 100 ng) for hybridisation to the third generation microarray (Fig. 8).

**Prymnesium parvum**

Both Eukaryote probes EukS_1209_25_dT and EukS_328_25_dT produced a positive microarray signal when hybridised with 25 ng and 100 ng of labelled *P. parvum* RNA. The only higher group probe for Prymnesiophyta (PrymS01_25_dT and PrymS02_25_dT) to
return a significant linear regression was PrymS01_25_dT normalised to POSITIVE_25_dT ($R^2 = 0.98$; $p = 0.0400$). However, both these higher group probes signal intensities were positively correlated except for PrymS02_25_dT normalised to DunGS02_25_dT_dT, which produced a negative linear regression (Table 2). The Class level probe PrymS03_25_dT for Prymnesiophyceae normalised signal intensity with both control probes had a positive linear relationship $R^2 \geq 0.84$; but, only the normalised POSITIVE_25_dT curve was significant ($p = 0.0276$). Both the clade level probes (Clade01old_25_dT (Prynymesium) and Clade01new25_dT (Prynymesium B1 clade sensu Edvardsen et al., 2000)) normalised signal intensities were positively correlated $R^2 \geq 0.73$ and significant ($p \leq 0.0477$) except for Clade01new25_dT normalised to DunGS02_25_dT_dT (Table 2).

Two species specific probes for Prynymesium parvum were redesigned for the third generation chip from original sequences in the 28S (PparvD01_25_dT; Töbe et al., 2007) and 18S regions (Prymparv01_25_dT; Eller et al., 2007) (Table S3). The first probe PparvD01_25_dT signal showed a significant ($R^2 = 0.99$; $p = 0.0020$) linear regression when normalised to POSITIVE_25_dT, but was not significant when normalised to DunGS02_25_dT_dT ($R^2 = 0.94$; $p = 0.0655$; Table 2). The second probe Prymparv01_25_dT showed a significant linear regression when signal intensities were normalised against both POSITIVE_25_dT ($R^2 = 0.99$; $p = 0.0149$) and DunGS02_25_dT_dT ($R^2 = 0.98$; $p = 0.0401$) probes (Table 2).

Cross reactivity was observed for HeteroS01_25_dT, HeterokontComp (Heterokonts) and DinoB_25_dT (Dinophytes) probes for RNA amounts $\geq 25$ ng, suggesting false positive microarray signals. These false positive signals increased for other probes when higher amounts of RNA were hybridised to the microarray, including probes PschGS01_25_dT (Pseudochattonella sp.), L*Kare0308A25_dT (Karenia sp.), KmGcS06_25_dT (Karenia
mikimoto), PcaserausD03_25_dT (Pseudo-nitzschia spp.), SSHaka0193A25_dT (H. akashiwo), DphyexacutaFS01_25_dT (Dinophyseiaceae (Dinophyslis+Phalacroma)), and KveneD06_25_dT (Karlodinium veneficum).

**Prymnesium polylepis**

All probes that were meant to be highlighted when hybridised with labelled *P. polylepis* RNA successfully passed the hierarchy test. Both PrymS01_25_dT and PrymS02_25_dT probe signal intensities produced a significant linear regression when normalised to POSITIVE_25_dT (R^2 ≥ 0.98; p ≤ 0.0373). However, when normalised to DunGS02_25_dT_dT, the linear regression was not significant with PrymS02_25_dT (Table 2). Class level probe PrymS03_25_dT signal intensity was not significant (R^2 = 0.97; p = 0.0615) when normalised with POSITIVE_25_dT as opposed to DunGS02_25_dT_dT normalisation resulting in a significant linear regression (R^2 = 0.99; p = 0.0251; Table 2). The signal intensities of the two Clade level probes Clade01old_25_dT and Clade01new25_dT had a significant linear regression when normalised to POSITIVE_25_dT (R^2 ≥ 0.985; p ≤ 0.0451), however, they were not significant when normalised to DunGS02_25_dT_dT (Table 2). The only species-specific probe spotted on the third generation chip for *P. polylepis* is Cpoly01_25_dT, its corresponding signal intensity returned a significant liner regression when normalised to POSITIVE_25_dT (R^2 > 0.98; p = 0.0361); this was not significant when normalised to DunGS02_25_dT_dT (Table 2).

Cross reactivity was observed for HeteroS01_25_dT, HeterokontComp (Heterokonts) and DinoB_25_dT (Dinophytes) for RNA amounts ≥25 ng. Several false positive signals were also observed at the same RNA amount ≥25 ng for the probes PschGS01_25_dT (*Pseudochattonella* sp.), L*Kare0308A25_dT (*Karenia* sp.), KbreD03_25_dT (*Karenia* sp.),
PcaserausD03_25_dT (*Pseudo-nitzschia* spp.) and LSGcat0270A24_dT (*Gymnodinium catenatum*).

cf. *Chattonella* sp.

The higher group Eukaryote probe EukS_1209_25_dT produced a positive microarray signal for all RNA amount (1 ng, 5 ng, 25 ng and 100 ng) hybridised to the third generation chip. However, EukS_328_25_dT did not produce a positive microarray signal for cf. *Chattonella* RNA amount <25 ng, which has implications for the hierarchical groups below this probe, which would be recorded as false positives. The species level probes for cf. *Chattonella* (CtoxS05_25_dT, CtoxiS07_25_dT and CtoxiS09_25_dT) all produced positive microarray signal-to-noise ratio values >2 with 1 ng of labelled cf. *Chattonella* RNA being hybridised to the third generation chip, which is equivalent to <200 cells. All three cf. *Chattonella* probe signal intensities returned significant linear regressions when normalised to POSITIVE_25_dT signals and only CtoxiS07_25_dT probe when normalised to DunGS02_25_dT_dT ($R^2 \geq 0.85; p \leq 0.0259$). The linear regression was not significant when CtoxS05_25_dT and CtoxiS09_25_dT probes were normalised to the signal returned by DunGS02_25_dT_dT probe (Table 2). In terms of hierarchy, CtoxiS09_25_dT probe produced a higher signal compared to CtoxS05_25_dT and CtoxiS07_25_dT probes and thus this is reflected in the hierarchy file.

Cross reactivity was observed for KmGcS06_25_dT (*Karenia mikimotoi*) at 1 ng and by PschGS04_25_dT (*Pseudochattonella* spp.) at 5 ng. This was also the case for probes AlexGD01_25_dT (*Alexandrium* sp.) and SSGcat0826A27_dT (*Gymnodinium catenatum*) for ≤25 ng RNA. The number of false positives increased for RNA amounts ≥100 ng for probes Clade01old_25_dT (*Prymnesium* B1 clade), PpungcalS01_25_dT (*Pseudo-nitzschia* spp.), PcalfrauD04_25_dT (*Pseudo-nitzschia* spp.), DacutaD02_25_dT (*Dinophysis* spp.),
PverD01_25_dT (*Pseudochattonella verruculosa*) and ProroFBS01 (*Prorocentrum* benthic clade). All false positives are eradicated by invoking the hierarchy file.

**Karlodinium veneficum**

Higher group Eukaryote probes EukS_1209_25_dT and EukS_328_25_dT produced a positive microarray signal when hybridised with low to high labelled RNA amounts of *K. veneficum*. This was also the case with the Class level probes DinoB_25_dT and DinoE12_25_dT, which returned a significant linear regressions when the signals were normalised to both probe controls POSITIVE_25_dT ($R^2 \geq 0.98; p < 0.0001$) and DunGS02_25_dT_dT ($R^2 \geq 0.99; p \leq 0.042$). The genus level probe KargeD01_25_dT signal produced a significant linear regression when normalised to the signal of POSITIVE_25_dT ($R^2 > 0.99; p < 0.0001$). However, this was not significant when normalised to DunGS02_25_dT_dT (Table 2). All the species specific probes (KveneD01_25_dT, KveneD02_25_dT, KveneD03_25, KveneD03_25_dT, KveneD04_25_dT and KveneD06_25_dT) had significant linear regressions ($R^2 \geq 0.99; p \leq 0.0268$) except KveneD06_25_dT when normalised to DunGS02_25_dT_dT ($R^2 = 0.75; p = 0.1127$; Table 2). Species-specific probes KveneD04_25_dT, KveneD03_25_dT KveneD06_25_dT and KveneD03_25 produced a positive microarray signal-to-noise ratio values >2 with labelled RNA amount of 1 ng, which corresponds to ~250 *K. veneficum* cells. KveneD01_25_dT and KveneD02_25_dT probes required at least 5 ng to produce a positive microarray signal.

Cross reactivity leading to false positives was observed at the lowest levels of 1 ng with probes HeterokontCOMP, PrymS02_25_dT (Prymnesiophyta), KbreD05_25_dT (*Karenia brevis*) and PmulacalD02_25_dT (*P. multistriata*+*P. calliantha*+*P. australis*). The
number of false positives increased with higher amounts of *K. veneficum* RNA, including PrymS01_25_dT (*Prymnesiophyta*), AlexGD01_25_dT (*Alexandrium* sp.), L*Kare0308A25_dT (*Karenia* sp.), DphyFS02_25_dT (*Dinophysiaceae* (*Dinophysis*+*Phalacroma*)) and DphyexacutaFS01_25_dT (*Dinophysiaceae* (*Dinophysis*+*Phalacroma*)) at hybridisations of 5 ng. Multiple false positives were recorded for $\geq 25$ ng RNA amounts; however all false positives are eliminated by the hierarchy file.

**Discussion**

The determination of the relationship between RNA amount, microalgal cell numbers and microarray signals was a key objective of the MIDTAL project. A reliable estimate of toxic algal cell concentrations in environmental samples could then be obtained. In order to achieve this, however, it was necessary to evaluate how RNA yields varied in microalgal strains subjected to different stress conditions.

*Prymnesium parvum* and *P. polylepis*.

*Prymnesium spp.*, which belong to the division Haptophyta, have been well documented for their ability to form golden brown blooms often associated with large fish mortality and extensive economic loss (Edvarsden & Paasche, 1998). This, in turn, has prompted widespread studies into the ecology of these harmful algae and the factors driving their growth, especially that of *Prymnesium parvum*, which is one of the most toxic and well-studied species (Guo *et al.*, 1996; Edvarsden & Paasche, 1998; Landsberg, 2002; Granéli *et al.*, 2012). Another prymnesiophyte is *Chrysochromulina polylepis*, a very toxic species that caused a devastating bloom in south-western Sweden and parts of the Norwegian Sea in 1988, killing over 800 tons of farmed fish with losses estimated at 10 million (Edvarsden & Paasche, 1998). Recently, morphological and ribosomal DNA sequence data have revealed

This article is protected by copyright. All rights reserved.
that *C. polylepis* is more closely related to *Prymnesium* spp. than to other *Chrysochromulina* spp., and it has since been reclassified into the genus *Prymnesium* as *Prymnesium polylepis*. (Edvarsden *et al.*, 2011).

For the two *Prymnesium* spp. strains tested in this study, salinity modulation caused the lowest variation in cellular RNA yield (pg/cell) over the period tested. This is not surprising because it has been reported that *P. parvum* can grow over a wide range of salinities (Edvarsden & Paasche, 1998) whereas *P. polylepis* is not usually found in estuarine waters where salinity varies. Several nutrient studies have been carried out on *P. parvum* and *P. polylepis* strains in relation to nitrogen : phosphorus ratio modulation to assess intracellular carbon, nitrogen and phosphorus dynamics together with the potential production of haemolytic substances (Johannsson & Granéli, 1999; Granéli *et al.*, 2012). Toxic effect potential has been reported highly variable among the haptophytes under nutrient limiting conditions, which can display allopathy related inhibitions toward other phytoplankton species and other marine organisms (Johannsson & Granéli, 1999; Granéli *et al.*, 2012). Potentially impacting on the intracellular RNA amount caused by growth suppression, which may explain the highly variable *P. polylepis* RNA yields (pg/cell) for UiO037, UiO038, and CCMP1757 strains grown separately and all the strains grown together (ALL mix) for nutrient stress conditions. This also may suggest that strains within the same species may also try to outcompete one another during nutrient limitation. This possibly had an effect on finding statistical differences for *P. polylepis* between the three individual strains and the ALL mix cultures under varying salinity and nutrient conditions because of the substantial standard deviations, which made the statistical analysis less conservative.

cf. *Chattonella* spp.

An unknown microflagellate was isolated from a fish kill in Torquay Canal, Rehoboth Bay Delaware, USA (Bowers *et al.*, 2004; Tomas *et al.*, unpublished). It was initially
identified as *Chattonella* cf. *verruculosa* based on its multiple plastids and a flagellar structure similar to that of other raphidophytes. A total of about 7 strains were isolated from this first fish kill and subsequent blooms from Delaware, Maryland and North Carolina. A phylogenetic analysis using 18S rDNA data revealed that the strains belonged to a new algal class, sister to a clade containing raphidophytes, xanthophytes and phaeophytes. A description of a new algal class has been submitted for publication (Tomas et al., unpublished) and to not invalidate the old name yet, we referred to the strains used in this study as cf. *Chattonella*. To date, all strains have been tested toxic for a breve-like toxin (Bourdelais et al., 2002).

Cf. *Chattonella* spp. was most affected by salinity, with lower salinity producing the highest RNA yield (pg/cell). Raphidophyte species have been known to be salinity tolerant, producing increases in the level of toxins in low salinity environments, which are thought to have evolved from predation pressures (Strom et al., 2013). This could be a possible explanation for cf. *Chattonella* increased ribosomal activity under low salinity conditions. Nutrient depletion only affected the RNA yield (pg/cell) of one strain (CMSTA 307). Reactions to light and temperature were not significantly different among the three strains. Since this is a new algal class, very little is known of the biodiversity and distribution of its species. It is hence not feasible to speculate about how representative the results obtained in this study are for the strains tested.

*Karodinium veneficum*

The genus *Gymnodinium* until recently comprised a diverse assemblage of naked (unarmored) dinoflagellates. With molecular techniques and enhanced SEM techniques, several new genera such as *Karenia*, *Karodinium*, and *Takayama* have been delineated from *Gymnodinium* and the corresponding species numbers have steadily increased (Daughberg et
Karlodinium veneficum, formerly Karlodinium micrum, is a common member of temperate, coastal phytoplankton assemblages, occasionally forming blooms and whose toxins, karlotoxins, cause membrane permeabilisation associated with fish kills (Bachvaroff et al., 2009, Van Wagoner et al., 2008).

Among the strains of Karlodinium veneficum tested, lower temperatures and lower salinities affected the growth rate and RNA yield (pg/cell) more than the other stress conditions. The type locality of its synonym, K. micrum, is cold temperate, and K. micrum has been reported from a broad geographic range in cold temperate waters in both hemispheres and in river habitats presumably exhibiting estuarine conditions (Bergholtz et al., 2005). It is unclear how many of these reports could represent a cryptic species because K. micrum is now a later synonym of K. veneficum based on identical morphology and near identical rDNA LSU sequences (Bergholtz et al., 2005).

**Microarray calibration curves**

The *P. parvum* species specific probe PparvD01_25_dT originally called PRYM694 was not extended by 25 nucleotides as the name suggests, and is still the original sequence plus the addition of a 15 nucleotide poly-T tail from the second generation chip (McCoy et al., 2013). This may explain the poor performance of the PparvD01_25_dT probe compared to the other *P. parvum* specific probe Prymparv01_25_dT, which was re-designed from the genus level probe PrymGS01_25 from the second generation chip (Eller et al., 2007). It was established that the PrymGS01_25 probe had more affinity towards *P. parvum* species when extended by 25 nucleotides plus poly-T tail on the third generation chip (McCoy et al., 2013). The Clade level probes Clade01old_25_dT for Prymnesium target species contains the poly-T tail but was not extended by 25 nucleotides and is the same original sequence from Simon et al. (1997) and was renamed Clade01 in the second generation chip (McCoy et al., 2013).
However, given that the probe still produces a good signal for Prymnesium targets, it was included on the third generation chip.

*Prymnesium spp.* detection limits were determined from the second generation microarrays at levels >5 ng of RNA, which approximately corresponds to 8,800 cells of *P. parvum* and 3,800 cells of *P. polylepis*. The only difference between the second and third generation microarray was that probes that showed non-specificity in the former were omitted from the latter (McCoy *et al.*, 2013). The slopes of the calibration curves generated using varying amounts of labelled RNA were then used to infer cell numbers from the microarray signal intensities via a GPR-Analyzer program (Dittami & Edvardsen, 2013).

Detection limits were as low as 1 ng RNA for cf. *Chattonella* and *K. veneficum*, respectively, when normalised to the control probe POSITIVE_25_dT. Low detection limits of 1 ng and 5 ng RNA obtained from pure cultures may not necessarily be achievable with field samples as there may be potential interferences from non-target RNA. For example, McCoy *et al.* (submitted) showed that during a bloom of *Alexandrium minutum* in the North Channel of Cork Harbour (Ireland) a minimum of 3,900 cells were needed to produce a positive signal for the species-specific probe AminuS01_25_dT in field samples, whereas in Taylor *et al.* (submitted), 1 ng of RNA extracted from a culture of *A. minutum* returned a suitable microarray signal intensity corresponding to 270 *A. minutum* cell equivalents.

In the environment, cells will be present in a variety of growth stages. Whereas the experiments described were all carried out during the exponential growth phase, some cells in environmental samples will be present at late exponential or stationary phase, and some may even be dead or senescent. It might be expected that RNA will be present at varying levels. The presence of cells in, for example, stationary phase will likely underestimate cell abundances in environmental samples based on RNA content per cell. In the environment, the

This article is protected by copyright. All rights reserved.
life time of micro-algal blooms is approximately 10 days (Parsons et al., 1984; Mann and Lazier, 1996), during which most of the time will be spent in exponential growth. One must also bear in mind that the MIDTAL microarray was designed for use as part of an early warning system, i.e. in a situation where cells would be actively growing. Therefore, it would be detecting cells as they are starting to bloom and thus would be in exponential growth phase and the inference of cell numbers would be more accurate then. Currently all monitoring programs are based on cell numbers and as cell numbers increase above a critical threshold, the fisheries are closed. Thus, the optimal use of the microarray is fully in keeping with current monitoring practices. We therefore consider the technique to be robust for monitoring purposes. In bacterial studies it has been observed that in times of nutrient deprivation together with other stresses, microorganisms survive by down regulating rRNA biosynthesis, ribosomal proteins and DNA replication, which is dictated by up regulating the levels of regulatory gene RpoS. This leads not only to physiological changes but also to stress resistance in the form of secondary metabolites, antibiotics and toxins and the stability of rRNA (Navarro Llorens et al., 2010). Hence, it is concluded that an adequate correlation between cell counts and microarray signals will be obtained when developed with exponentially growing cells.

The differences between hybridisations with pure cultures or field samples have proved statistically significant in terms of detection limits (McCoy et al., submitted). Experiments consisting of spiking field samples with known amounts of RNA extracted from pure cultures should be considered to ascertain this aspect further. The discrepancy observed with A. minutum containing culture and field samples is somewhat harder to observe when dealing with microflagellates. Indeed identification of the flagellates by light microscopy alone is difficult and the cell counts obtained for preserved water samples may be
underestimated. Light microscopy analysis is often inadequate to carry out specimen identification at the species level.

The cross reactivity issues observed in this study were somewhat consistent across all of the four species used during the experiments described above. This would suggest that in order to minimise false positives, all that is required is to remove the probes that are causing problems. However, it is not a simple matter because some of the cross-reacting probes are part of hierarchies used for other species and would then require the development of a new hierarchy file. The principal of the hierarchy file has provided the best means to eliminate false positives which are likely to occur when field samples with an unknown composition of species are taken for RNA extraction and analysis.

The linear regressions carried out on the data were more consistent when normalised to the positive TATA box protein control POSITIVE_25_dT than those normalised to the Dunaliella specific control DunGS02_25_dT, which may lead to microarray calibration curves returning erroneous cell count estimates. One reason for this observation may be that the same standardised stock solution of TATA box protein was used by all the MIDTAL partners and hence a more consistent control for data normalisation. In addition, each partner cultivated and harvested their own Dunaliella sp. control strain, which could have led to inconsistencies across different laboratories. This error could be reduced if the extracted RNA of Dunaliella sp. spiked to each sample for normalisation was standardised and included in the MIDTAL kit produced by Microbia Environement (France) and not made up by each individual user. However, the inclusion of Dunaliella cells is meant to be a control on the extraction efficiency of the sample. Also if the signal from Dunaliella is too high, then this will tend to affect marginal but positive signals, and render them negative. In a different series of experiments carried out with the Dunaliella controls, the probe was spotted in three
different concentrations to minimise the possibility of observing a saturated high signal (data not shown). A deterioration of the stock solution of TATA box protein was observed over time in some occasions, which caused the corresponding signal to be absent on the microarray chip. There is no warning prior to hybridisation as to the deteriorating state of that the stock solution of TATA box protein, hence having a back-up positive control consisting of known amounts of spotted *Dunaliella* sp. on the chip still remains a necessity despite the issues mentioned above. To test the accuracy of the microarray, actual cell counts taken from natural seawater field samples were correlated with those inferred from the microarray with a good correlation (Medlin, 2013). The toxicity of field samples was also used as an indicator of species presence and compared against the microarray signal and cell count results.

Although there are number of molecular techniques being developed that can accurately detect and quantify low abundances of harmful phytoplankton species, many of these methods only target one particular group or species present in a field sample. The main advantage of the MIDTAL microarray over other quantitative tools is its capacity to detect and quantify multiple species in a single analysis. A further advantage is that since there is no PCR step, it is not susceptible to any unknown inhibitors in the field sample. The results obtained from this study will aid with the further development and improvement of the MIDTAL microarray.

**Conclusion**

The series of experiments reported here showed a positive linear response of increasing RNA yield with increasing microalgal cell numbers. The RNA content per cell was not affected by the environmental stress caused by modulations in light, temperature, salinity and nutrients and over time for the species *Prymnesium spp.* cf. *Chattonella* and *K.*

This article is protected by copyright. All rights reserved.
veneficum. The calibration curves showing the relationship between microarray signal intensities and RNA amounts for each species provide a confident indication of the presence of a given species in an environmental water sample as well as its abundance.

Acknowledgements

This work was funded through the EU 7th Framework Programme (FP7-ENV-2007-1-MIDTAL-201724). The authors would like to thank Bente Edvardsen and Carmelo Tomas for provided algal strains. The scientific and technical assistance of Jixin Chen and co-workers of the MIDTAL project team is also greatly appreciated.

References


This article is protected by copyright. All rights reserved.


This article is protected by copyright. All rights reserved.


This article is protected by copyright. All rights reserved.


This article is protected by copyright. All rights reserved.

Mann KH & Lazier RJN (1996) Dynamics of Marine Ecosystems: *Biological Physical Interactions in the Oceans, 2nd* (eds), Blackwell Science.


This article is protected by copyright. All rights reserved.


This article is protected by copyright. All rights reserved.


**Figure Legends**

*Fig 1.* Growth rates of individual *P. parvum, P. polylepis, cf. Chattonella* and *K. veneficum* strains for low (L), optimal (O) and high (H) light, temperature, salinity and nutrient (N- and P- indicate nitrate or phosphate depletion and are represented as L (P-), O (N+P+) and H (N-)) conditions. Calculated from day 0 (T1) and day 3 (T2 = 72 h) for *Prymnesium spp.* (a and b) and day 2 (T2 = 48 h) for *cf. Chattonella* and *K. veneficum* (c and d) time points. The numbers represent the different strains of *P. parvum* (“1” (UIO054), “2” (SAG127.79) “3” (CCMP709)), *P. polylepis* (“1” (UIO036), “2” (UIO037) “3” (UIO038) “4” (CCMP1757)), cf. *Chattonella* (“1” (CMSTAC300), “2” (CMSTAC305) “3” (CMSTAC307)) and *K. veneficum* (“1” (CCCM734), “2” (PCC709) “3” (PCC517)). The word “mix” represents the strains that were grown together in the same culture flasks for the *Prymnesium* species. The test conditions for each treatment follow that in Table S2.

*Fig 2.* Correlation between average RNA yield (pg/cell) and growth rates (d⁻¹) over different treatments and four different culture conditions (light, temperature, salinity and nutrients) for *Prymnesium parvum* (a to d) and *Prymnesium polylepis* (e to h). Each data point is indicated by a label: the first letters indicate low (L), optimal (O) and high (H) and the second letter indicates light (L), temperature (T), salinity (S) and nutrients (N- and P- indicate nitrate or phosphate depletion) experiments, the numbers represent the different strains of *P.*

This article is protected by copyright. All rights reserved.
*parvum* ("1" (UIO054), "2" (SAG127.79) "3" (CCMP709)) and *P. polylepis* ("1" (UIO036), "2" (UiO037) "3" (UiO038) "4" (CCMP1757)). n.s. = not significant, r = Pearson correlation coefficient, p = significance of the correlation and * = level of significance. The test conditions for each treatment follow that in Table S2.

**Fig 3.** Average RNA yield (pg/cell) for *Prymnesium parvum* and *Prymnesium polylepis* under various treatments of light, temperature, salinity and nutrient culture conditions. Note that results of salinity and nutrients conditions for *P. polylepis* do not contain UiO 038 strain data because of cell absence or death during stress period. A 2way ANOVA followed by a Bonferroni post test was used to analysis if there was any significant difference between treatments and over time. (n.s. = not significant; p > 0.05).

**Fig 4.** Comparisons of the average RNA amount (pg/cell) between individual strain grown separately and all (mix) strains of each *Prymnesium* species grown together. (a) *P. parvum* comparison of average RNA yields for UiO 054, CCMP 709, and SAG 127.79 strains grown separately and all (mix) the strains grown together over four conditions. (b) *P. polylepis* comparison of average RNA yields for UiO 036, UiO 037, and CCMP 1757 strains grown separately and all (mix) the strains grown together over light and temperature conditions. For *P. polylepis* salinity and nutrient conditions it is the comparison of average RNA yields for UiO 037, UiO 038, and CCMP 1757 strains grown separately and all (mix) the strains grown together.

**Fig 5.** Correlation between total RNA amount (ng) and cell numbers over four different culture conditions (light, temperature, salinity and nutrient) for *P. parvum*, *P. polylepis*, cf. *Chattonella* and *K. veneficum*. (a) Each *P. parvum* strain (UiO 054, CCMP 709, and SAG 127.79) and (c) *P. polylepis* strain (UiO036, UiO037, UiO038 and CCMP1757) grown separately (n = 108) and (b) all *P. parvum* strains and (d) *P. polylepis* strains grown together (n = 35) from T1 to T3. (e) Each cf. *Chattonella* strain (CMSTAC300, CMSTAC305 and CMSTAC307) and (f) *Karlodinium veneficum* strain (CCCM734, PCC709 and PCC517) grown separately from T2 (n = 36, respectively). The black line represents the linear regression, r = Pearson correlation coefficient, p = significance of the correlation (alpha=0.05).

**Fig 6.** Correlation between RNA yield (pg/cell) and growth rates (d⁻¹) over different treatments and four different culture conditions (light, temperature, salinity and nutrients) for cf. *Chattonella* (a to d) and *Karlodinium veneficum* (e to h). Each data point is indicated by a
The first letters indicate low (L), optimal (O) and high (H) and the second letter indicates light (L), temperature (T), salinity (S) and nutrients (N- and P- indicate nitrate or phosphate depletion) experiments, the numbers represent the different strains for cf. *Chattonella* (“1” (CMSTAC300), “2” (CMSTAC305) “3” (CMSTAC307)) and *K. veneficum* (“1” (CCCM734), “2” (PCC709) “3” (PCC517)). n.s. = not significant, r = Pearson correlation coefficient, p = significance of the correlation and * = level of significance. The test conditions for each treatment follow that in Table S2.

**Fig 7.** RNA yield (pg/cell) for cf. *Chattonella* strains (CMSTAC300, CMSTAC305 and CMSTAC307) and *Karlodinium veneficum* strains (CCCM734, PCC709 and PCC517) under various treatments of light, temperature, salinity and nutrient culture conditions. Because replicate values are not included, it is necessary to assume that there is no interaction. In other words, this analysis assumes that the strains have the same effect (if any) at all levels of conditions using 2way ANOVA. (n.s. = not significant; p > 0.05).

**Fig 8.** Calibrations curves between normalised microarray signals and corresponding cell numbers. (a and b) Calibration curves for *Prymnesium parvum* showing higher group probes (PrymS01_25_dT, PrymS02_25_dT and PrymS03_25_dT), clade level probes (Clade01old_25_dT and Clade01new25_dT) and species level probes (PparvD01_25_dT and Prymparv01_25_dT) normalised to control probes POSITIVE_25_dT and DunGS02_25_dT_dT, respectively. (c and d) Calibration curves for *Prymnesium polylepis* showing higher group probes (PrymS01_25_dT, PrymS02_25_dT and PrymS03_25_dT); clade level probes (Clade01old_25_dT and Clade01new25_dT) and species level probes (CpolyS01_25_dT) normalised to control probes POSITIVE_25_dT and DunGS02_25_dT_dT, respectively. (e and f) Calibration curves for cf. *Chattonella* spp. showing species level probes (CtoxiS09_25_dT, CtoxiS05_25_dT and CtoxiS07_25_dT) normalised to control probes POSITIVE_25_dT and DunGS02_25_dT_dT, respectively. (g and h) Calibration curves for *Karlodinium veneficum* showing class level probes (DinoB_25_dT and DinoE12_25_dT); genus level probe (KargeD01_25_dT) and species level probes (KveneD01_25_dT, KveneD02_25_dT, KveneD03_25_dT, KveneD03_25, KveneD04_25_dT, and KveneD06_25_dT) normalised to control probes POSITIVE_25_dT and DunGS02_25_dT_dT, respectively (Note: trendline forced intercept of zero).

**Table 1.** Average RNA content per cell (pg/cell) of individual species strains for *P. parvum, P. polylepis, cf. Chattonella* and *K. veneficum*.
<table>
<thead>
<tr>
<th>Species</th>
<th>P. parvum</th>
<th>P. polylepis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strain</td>
<td>UIO054</td>
<td>CCMP709</td>
</tr>
<tr>
<td>Average RNA content per cell (pg/cell)</td>
<td>0.60</td>
<td>0.62</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.14</td>
<td>0.15</td>
</tr>
<tr>
<td>Species</td>
<td>cf. Chattonella</td>
<td>K. veneicum</td>
</tr>
<tr>
<td>Strain</td>
<td>CMSTAC300</td>
<td>CMSTAC305</td>
</tr>
<tr>
<td>Average RNA content per cell (pg/cell)</td>
<td>21.43</td>
<td>15.92</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>11.93</td>
<td>6.34</td>
</tr>
</tbody>
</table>

Table 2. The $R^2$ values, corresponding slope and level of significance of linear regression for normalised microarray signals against the controls POSITIVE\_25\_dT and DunGS02\_25\_dT\_dT are shown below. Each probe is represented that should be highlighting when pure culture labelled RNA from *P. parvum*, *P. polylepis*, cf. *Chattonella* and *K. veneicum* is hybridised to the third generation microarray (Note. trendline forced intercept of zero).

<table>
<thead>
<tr>
<th>Species</th>
<th>Probe name</th>
<th>Normalised probe</th>
<th>$R^2$ value</th>
<th>Slope</th>
<th>P values</th>
</tr>
</thead>
<tbody>
<tr>
<td>PrymS01_25_dT</td>
<td>POSITIVE_25_dT</td>
<td>0.9814</td>
<td>0.000061 ± 0.000004</td>
<td>0.0400*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DunGS02_25_dT_dT</td>
<td>0.7536</td>
<td>0.00024 ± 0.00004</td>
<td>0.1129</td>
<td></td>
</tr>
<tr>
<td>PrymS02_25_dT</td>
<td>POSITIVE_25_dT</td>
<td>0.8657</td>
<td>0.000063 ± 0.000009</td>
<td>0.0910</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DunGS02_25_dT_dT</td>
<td>-0.122</td>
<td>0.00025 ± 0.00007</td>
<td>0.1772</td>
<td></td>
</tr>
<tr>
<td>PrymS03_25_dT</td>
<td>POSITIVE_25_dT</td>
<td>0.9917</td>
<td>0.000033 ± 0.000001</td>
<td>0.0276*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DunGS02_25_dT_dT</td>
<td>0.8401</td>
<td>0.00013 ± 0.00002</td>
<td>0.0969</td>
<td></td>
</tr>
<tr>
<td>Clade01old_25_dT</td>
<td>POSITIVE_25_dT</td>
<td>0.9992</td>
<td>0.0000417 ± 0.0000006</td>
<td>0.0096**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DunGS02_25_dT_dT</td>
<td>0.9722</td>
<td>0.00016 ± 0.00001</td>
<td>0.0477*</td>
<td></td>
</tr>
<tr>
<td>Clade01new_25_dT</td>
<td>POSITIVE_25_dT</td>
<td>0.9789</td>
<td>0.000033 ± 0.000002</td>
<td>0.0426*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DunGS02_25_dT_dT</td>
<td>0.7345</td>
<td>0.00013 ± 0.00002</td>
<td>0.1157</td>
<td></td>
</tr>
<tr>
<td>PparvD01_25_dT</td>
<td>POSITIVE_25_dT</td>
<td>0.9999</td>
<td>0.00000428 ± 0.00000001</td>
<td>0.0020**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DunGS02_25_dT_dT</td>
<td>0.9435</td>
<td>0.000017 ± 0.000002</td>
<td>0.0655</td>
<td></td>
</tr>
<tr>
<td>Species</td>
<td>Trait 1</td>
<td>Trait 2</td>
<td>Trait 3</td>
<td>Trait 4</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>Prymparv01_25_dT</td>
<td>POSITIVE_25_dT</td>
<td>0.9888</td>
<td>0.000051 ± 0.000001</td>
<td>0.0149*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DunGS02_25_dT_dT</td>
<td>0.9813</td>
<td>0.00020 ± 0.00001</td>
<td>0.0401*</td>
<td></td>
</tr>
<tr>
<td>PrymS01_25_dT</td>
<td>POSITIVE_25_dT</td>
<td>0.9993</td>
<td>0.000119 ± 0.000002</td>
<td>0.009**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DunGS02_25_dT_dT</td>
<td>0.7118</td>
<td>0.00026 ± 0.00005</td>
<td>0.1191</td>
<td></td>
</tr>
<tr>
<td>PrymS02_25_dT</td>
<td>POSITIVE_25_dT</td>
<td>0.9841</td>
<td>0.000156 ± 0.000009</td>
<td>0.0373*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DunGS02_25_dT_dT</td>
<td>-0.667</td>
<td>0.00032 ± 0.00010</td>
<td>0.1946</td>
<td></td>
</tr>
<tr>
<td>PrymS03_25_dT</td>
<td>POSITIVE_25_dT</td>
<td>0.9735</td>
<td>0.000099 ± 0.000001</td>
<td>0.0615</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DunGS02_25_dT_dT</td>
<td>0.9934</td>
<td>0.000213 ± 0.000008</td>
<td>0.0251*</td>
<td></td>
</tr>
<tr>
<td>Clade01old_25_dT</td>
<td>POSITIVE_25_dT</td>
<td>0.985</td>
<td>0.000095 ± 0.000007</td>
<td>0.0451*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DunGS02_25_dT_dT</td>
<td>0.9597</td>
<td>0.000201 ± 0.000002</td>
<td>0.0559</td>
<td></td>
</tr>
<tr>
<td>Clade01new25_dT</td>
<td>POSITIVE_25_dT</td>
<td>0.9992</td>
<td>0.000071 ± 0.000001</td>
<td>0.0095**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DunGS02_25_dT_dT</td>
<td>0.7204</td>
<td>0.00016 ± 0.00003</td>
<td>0.1177</td>
<td></td>
</tr>
<tr>
<td>CpolyS01_25_dT</td>
<td>POSITIVE_25_dT</td>
<td>0.9897</td>
<td>0.000083 ± 0.000005</td>
<td>0.0361*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DunGS02_25_dT_dT</td>
<td>0.9209</td>
<td>0.00018 ± 0.00002</td>
<td>0.0712</td>
<td></td>
</tr>
<tr>
<td>CtoxiS09_25_dT</td>
<td>POSITIVE_25_dT</td>
<td>0.9557</td>
<td>0.00025 ± 0.000002</td>
<td>0.0016**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DunGS02_25_dT_dT</td>
<td>0.0101</td>
<td>0.000072 ± 0.000020</td>
<td>0.1716</td>
<td></td>
</tr>
<tr>
<td>CtoxiS05_25_dT</td>
<td>POSITIVE_25_dT</td>
<td>0.855</td>
<td>0.00019 ± 0.00003</td>
<td>0.0086**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DunGS02_25_dT_dT</td>
<td>-1.2066</td>
<td>0.00049 ± 0.00016</td>
<td>0.2064</td>
<td></td>
</tr>
<tr>
<td>CtoxiS07_25_dT</td>
<td>POSITIVE_25_dT</td>
<td>0.9593</td>
<td>0.00018 ± 0.00002</td>
<td>0.0014**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DunGS02_25_dT_dT</td>
<td>0.993</td>
<td>0.00060 ± 0.00002</td>
<td>0.0259**</td>
<td></td>
</tr>
<tr>
<td>DinoB_25_dT</td>
<td>POSITIVE_25_dT</td>
<td>0.9756</td>
<td>0.00043 ± 0.00003</td>
<td>&lt; 0.0001***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DunGS02_25_dT_dT</td>
<td>0.9997</td>
<td>0.00152 ± 0.00001</td>
<td>0.0054**</td>
<td></td>
</tr>
<tr>
<td>DinoE12_25_dT</td>
<td>POSITIVE_25_dT</td>
<td>0.9969</td>
<td>0.000261 ± 0.000007</td>
<td>&lt; 0.0001***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DunGS02_25_dT_dT</td>
<td>0.9866</td>
<td>0.000096 ± 0.00006</td>
<td>0.042*</td>
<td></td>
</tr>
<tr>
<td>KargeD01_25_dT</td>
<td>POSITIVE_25_dT</td>
<td>0.9991</td>
<td>0.000147 ± 0.000002</td>
<td>&lt; 0.0001***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DunGS02_25_dT_dT</td>
<td>0.9686</td>
<td>0.00043 ± 0.00005</td>
<td>0.0681</td>
<td></td>
</tr>
<tr>
<td>KveneD01_25_dT</td>
<td>POSITIVE_25_dT</td>
<td>0.9966</td>
<td>0.000132 ± 0.000003</td>
<td>&lt; 0.0001***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DunGS02_25_dT_dT</td>
<td>0.9995</td>
<td>0.000139 ± 0.000002</td>
<td>0.0074**</td>
<td></td>
</tr>
<tr>
<td>KveneD02_25_dT</td>
<td>POSITIVE_25_dT</td>
<td>0.9993</td>
<td>0.0000630 ± 0.0000008</td>
<td>&lt; 0.0001***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DunGS02_25_dT_dT</td>
<td>0.9999</td>
<td>0.0000917 ± 0.0000004</td>
<td>0.0027**</td>
<td></td>
</tr>
<tr>
<td>KveneD03_25</td>
<td>POSITIVE_25_dT</td>
<td>0.9961</td>
<td>0.000339 ± 0.000001</td>
<td>&lt; 0.0001***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DunGS02_25_dT_dT</td>
<td>0.9999</td>
<td>0.000341 ± 0.0000002</td>
<td>0.0037**</td>
<td></td>
</tr>
<tr>
<td>KveneD03_25_dT</td>
<td>POSITIVE_25_dT</td>
<td>0.9988</td>
<td>0.000440 ± 0.000007</td>
<td>&lt; 0.0001***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DunGS02_25_dT_dT</td>
<td>0.9941</td>
<td>0.00052 ± 0.00002</td>
<td>0.0268*</td>
<td></td>
</tr>
<tr>
<td>KveneD04_25_dT</td>
<td>POSITIVE_25_dT</td>
<td>0.9977</td>
<td>0.00062 ± 0.00001</td>
<td>&lt; 0.0001***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DunGS02_25_dT_dT</td>
<td>0.9872</td>
<td>0.00112 ± 0.00007</td>
<td>0.041*</td>
<td></td>
</tr>
<tr>
<td>KveneD05_25_dT</td>
<td>POSITIVE_25_dT</td>
<td>0.9988</td>
<td>0.000261 ± 0.000004</td>
<td>&lt; 0.0001***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DunGS02_25_dT_dT</td>
<td>0.7542</td>
<td>0.00022 ± 0.00004</td>
<td>0.1127</td>
<td></td>
</tr>
</tbody>
</table>

P values (>0.05 = not significant; 0.01 to 0.05 = significant*; 0.001 to 0.01 = very significant** and <0.001 = extremely significant***)

---

This article is protected by copyright. All rights reserved.
This article is protected by copyright. All rights reserved.