



Phytoplankton growth characterization in short term MPN culture assays using 18S metabarcoding and qRT-PCR

Brian Petri ^{a, *}, Subba Rao Chaganti ^b, Po-Shun Chan ^a, Daniel Heath ^{b, c}

^a Trojan Technologies, 3020 Gore Rd, London, ON, N5V 4T7, Canada

^b Great Lakes Institute for Environmental Research, University of Windsor, 401 Sunset Avenue, Windsor, ON, N9B 3P4, Canada

^c Department of Biological Sciences, University of Windsor, 401 Sunset Avenue, Windsor, ON, N9B 3P4, Canada

ARTICLE INFO

Article history:

Received 9 February 2019

Received in revised form

30 July 2019

Accepted 31 July 2019

Available online 31 July 2019

Keywords:

Ballast water

Next-generation sequencing

Viability methods

Invasive species

Most probable number (MPN)

Quantitative PCR

ABSTRACT

The most probable number dilution-culture assay (MPN) is used to enumerate viable phytoplankton in regulatory tests of ballast water treatment systems. However the United States Coast Guard has not yet accepted MPN, in part due to concerns of biased results due to cells being viable but not growing. MPN does not assess the fate of every cell, and thus the bias can only be evaluated by a companion method that assesses the ability of the various taxa to grow. This growth ability (“growability”) is the complement of the bias, and has been evaluated by microscopic taxonomy of before-culture and after-culture samples. However, microscopic taxonomy is extremely laborious and few data have been produced for phytoplankton growability in MPN assays. To address the need for more and more reliable growability data, a method was developed using next-generation sequencing (NGS) and quantitative real time PCR (qRT-PCR) techniques that target the V9 region of the 18S rRNA gene for the taxonomic identification and growth assessment of eukaryotic phytoplankton, respectively. This growability method was applied to MPN samples from a ballast water management system test that were incubated with two different enrichment media at two different temperatures. DNA was extracted from filters of before-culture and after-culture samples, and assessed for taxonomy by NGS and for PCR template DNA concentration by qRT-PCR. Growth ratios based on changes in 18S template concentration over the incubation period were calculated for each taxon, and dead-cell DNA persistence through a 14 day incubation was verified to be <1% and did not influence the growth calculations. In total, 95 of 97 eukaryotic phytoplankton in the before-culture sample demonstrated growth, with definitive growth ratios ranging from 4.0×10^1 – 2.6×10^5 . An additional 13 taxa demonstrated growth from non-detect in before-culture samples. Taxa-based growability values were 87–88% in individual incubation conditions with no statistical differences among conditions, and 98% for all conditions combined. When growability was weighted by the before-culture abundance of each taxa, relevant to regulations based on all organisms regardless of taxa, community-based growability was >99% in each condition and in all conditions combined because the most abundant taxa all exhibited growth. This study verifies that conventional phytoplankton MPN assays produce accurate results with low bias from undetected viable cells, regardless of enrichments and incubation temperatures. This work can provide regulatory confidence for broader acceptance of MPN assays without limitations.

© 2019 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Enumerating viable organisms is necessary in studies concerned with aquatic ecosystem and public health protection, water and wastewater disinfection, microbial risk assessment, and invasive

species control by new treatment applications such as treatment of ballast water on ships. Viability, i.e. the ability to reproduce, is measured via culture-based assays which are the “gold-standard” for evaluating the status of microorganisms (Emerson et al., 2017). Reproductive capacity is critical as it is necessary to, for example, cause infection in a host or successfully colonize a new habitat. Culture methods are appropriate for evaluating all disinfection mechanisms, and are the *de facto* and legal standard for evaluating the safety of treated drinking water in the United States (40 C.F.R. §

* Corresponding author.

E-mail address: bpetri@trojanuv.com (B. Petri).

141.74, 2018) and elsewhere, as well as for other drink, food, and waste water. A relatively new application for culture methods is for the assessment of phytoplankton viability after ballast water treatments mandated to control the spread of potentially harmful and invasive non-native species (Blatchley et al., 2018; Cullen and MacIntyre, 2016).

Regulations for ballast water management system (BWMS) type-approval testing are not harmonized: for enumerating viable organisms in the regulated 10–50 µm size class (the majority of which are phytoplankton), the most probable number culture assay (MPN) is used in International Maritime Organization (IMO) type-approvals (IMO, 2016; IMO, 2017) but has not been accepted for United States Coast Guard (USCG) type approvals (Coast Guard Maritime Commons, 2016). This is problematic for ultraviolet (UV) – based BWMSs, because the existing USCG enumeration method, a stain-motility (SM) method (ETV, 2010), is not appropriate for evaluating UV-based BWMSs designed to eliminate viable organisms (IMO, 2017). The SM method does not measure the mechanism of UV disinfection, damage of DNA to prevent reproduction (Blatchley et al., 2018), which results in 10-fold higher UV dose requirements than culture methods (Lundgreen et al., 2018).

The USCG has expressed concerns that organisms cannot be cultured reliably (Coast Guard Maritime Commons, 2015), implying that MPNs are inaccurate due to viable cells that do not grow to detection. Indeed, in water treatment applications, viable organisms that do not grow in culture assays represent a measurement bias that is not protective of public health or the environment (Cullen, 2018). MPN assays do not classify each cell in a sample, so any growth bias can only be assessed in a companion method. One approach that has been used to estimate the growth bias for MPN on natural communities of phytoplankton is to make taxonomic assessments by microscopy on before-culture (BC) and after-culture (AC) samples, reporting the percentage of taxa or the abundance weighted percentage of taxa that demonstrated growth. The resulting “growability” values are the complement of the growth bias (Cullen, 2018). The approach is straightforward in principle but laborious in practice, which has limited the production of data. Encouragingly high values have been reported for subsets of MPN assays assessed for quality assurance (Madsen and Petersen, 2015), but rigorous values based on complete MPN arrays have not yet been published.

The approach of documenting BC and AC taxa can be improved using a modern assessment technology such as metabarcoding, that is, the use of Polymerase Chain Reaction (PCR) amplification of a marker DNA fragment from environmental DNA (eDNA) followed by massively parallel (Next Generation) sequencing (NGS) to identify organisms based on comparison to DNA sequence databases (Pawlowski et al., 2016). To address the need for more eukaryotic phytoplankton growability data, a method was developed in this work using metabarcoding to create BC and AC taxonomy lists, and quantitative real-time PCR (qRT-PCR) to classify each taxa as growing or not by determining whether each had increased or decreased in target sequence numbers, AC relative to BC. The focus was on eukaryotic phytoplankton, as prokaryotic phytoplankton (blue-green algae) will effectively be excluded from the regulated 10–50 µm size category.

This method was applied to measure growability in conventional MPN assays for BWMS tests and also to test for differences in growability among different incubation conditions (enrichment media, temperature), to inform on the factors that need to be standardized (or not) in MPN methods. The method showed that growability was high and insensitive to the tested incubation conditions. These results have significant implications not only for standardizing and validating MPN methods used in BWMS testing, but also for the determination of impacts of environmental

stressors on phytoplankton ecology in nature.

2. Materials and methods

2.1. Sample collection and MPN methods

This work used split samples collected from an MPN evaluation experiment performed by staff of the DHI Ballastwater Laboratory in Hørsholm, Denmark. The MPN experiment was to inform a panel formed under the Environmental Technology Verification (ETV) program in the United States to evaluate MPN as a method for measuring treatment performance of BWMSs in type-approval tests. The experiment evaluated various aspects of MPN method performance for different incubation conditions (enrichment media, temperature). The experiment was based on a 10–50 µm size fraction sample of untreated control water from a standard BWMS test at the DHI Maritime Technology Evaluation Facility in Hundested, Denmark, originating from the Kattegat (North Sea) at an ambient water temperature of 10 °C. The sample was processed in general according to the DHI standard MPN method (IMO, 2016), in 4 sets of MPNs prepared with two different enrichment media and incubated at two different temperatures, each in triplicate. All media were prepared using 0.22 µm filter-sterilized sample water, with either Guillard's f/2 (Guillard, 1975; Guillard and Rypther, 1962) or Keller Media (Keller et al., 1987) as enrichments. The two incubation temperatures were 20 °C and 10 °C. These 4 experimental incubation conditions are referred to as G10, G20, K10 and K20, with G and K referring to Guillard f/2 and Keller enrichment media respectively, and 10 and 20 referring to the incubation temperatures.

Each MPN array consisted of 5 replicates at 3 dilutions (10^{-2} , 10^{-3} , 10^{-4}), chosen to be able to assess viable phytoplankton concentrations at the expected target concentration of 1,000 cells/mL, and was incubated in constant light for 14 days. Growth was scored based on increases in chlorophyll autofluorescence from day 0 to day 14, monitored using a Turner TD-700 Laboratory Fluorometer. At the end of the incubation, the 5 replicates of each dilution in each of the 12 MPNs were pooled for further post-processing and archiving. The chlorophyll scoring was used to calculate MPN values as per Jarvis et al. (2010), and both are given in Table S1.

For this work to evaluate MPN growability, an 80 mL aliquot of the BC sample and 20 mL aliquots (of 30 mL total) of each of the 36 AC subsamples were used. For the AC samples, the within-MPN aliquot subsamples were re-combined to give 12 composite AC samples, one for each of the 12 MPNs. When generated, the 1 BC and 12 AC samples were filtered using 0.7 µm glass fiber filters and used for DNA extractions.

2.2. DNA isolation

DNA extractions were performed according to the protocol described in Chaganti et al. (2012). In brief, the filters were placed in 2-mL screw cap tubes with 400 µL of ddH₂O, 400 µg of 1.0 mm glass beads, 400 µL Cetyltrimethylammonium Bromide (CTAB) digestion buffer and 400 µL of phenol-chloroform isoamyl-alcohol (25:24:1). The samples were homogenized to break down cell structure using a Thermo Savant Bio 101 Fast Prep homogenizer at the 4.5-speed setting for 2 min. Phase separation was achieved by centrifugation at 13,000 g for 15 min at 4 °C. The supernatant was transferred to another 1.5-mL microcentrifuge tube, and an equal volume of isopropanol was added and the sample was mixed by inversion and held at 21 °C for 30 min. DNA was precipitated by centrifugation (12,000 g) at 4 °C for 15 min. The precipitated DNA pellet was washed twice with 70% ice cold ethanol and air dried before re-suspension in 30 µL sterile milliQ water.

2.3. Metabarcoding

A two-step PCR was performed on all DNA samples targeting the V9 region of the 18S rRNA gene (Forward 5'acctgctgccg GTACACACCGCCCGT3' and reverse 5'acgccaccgagcTGATCCTT CTGCAGGTTCACTAC3'), where a short (12 base) 5' extension (lower case base codes on the primer sequence) was added as template for the second stage PCR. The first step PCR was to amplify the targeted DNA and the second step (a short-cycle PCR) was designed to ligate identifying sequences ("barcodes") and the required adaptor sequences for the ION Torrent NGS library.

The thermocycling protocol for the first PCR of the V9 region of the 18S rRNA gene consisted of an initial denaturing stage at 94 °C for 2 min, followed by 27 cycles of: denaturing at 94 °C for 15 s, annealing at 55 °C for 15 s, elongation at 72 °C for 30 s, and a final elongation step at 72 °C for 7 min, followed by a hold at 4 °C. The second short-cycle PCR to create the barcoded library consisted of an initial denaturing stage at 94 °C for 2 min, followed by 8 cycles of: denaturing at 94 °C for 15 s, annealing at 60 °C for 15 s, elongation at 72 °C for 30 s, and a final elongation step at 72 °C for 7 min, followed by a hold at 4 °C. All barcoded PCR amplicons were pooled, visualized and excised from an agarose gel then purified using a Qiagen MinElute gel extraction kit. Pooled PCR product was assessed for DNA concentration and fragment size distribution using an Agilent 2100 Bioanalyzer with a High Sensitivity DNA chip (Agilent Technologies, Mississauga, Canada). The samples were sequenced on a "318" microchip on an Ion Torrent Personal Genome Machine ("PGM"; Life Technologies, USA) using 400 bp chemistry. After sequencing, sequence reads were filtered within the PGM software to remove polyclonal and low-quality sequences. Sequences that matched the PGM 3' adaptor were trimmed to remove non-informative sequence data.

2.4. Bioinformatic analyses

Raw sequences obtained from the NGS were quality filtered (maxEE 1.0) and truncated to 140 bp by excluding the primers, and sequences exhibiting barcode mismatches, primer mismatches and short or low quality reads were removed. Further data processing was performed using the UPARSE algorithm (Edgar, 2013) following the default parameters. Chimeric sequences were removed using the UCHIME v4.2 method (Edgar et al., 2011). Operational taxonomic units (OTUs) were assigned with a 97% similarity threshold and OTUs represented only by either a single (singleton) or two (doubleton) sequences across all samples (i.e., all BC and AC samples) were excluded from the data prior to metabarcoding species assignment. Further, QIIME software v 1.8 (Caporaso et al., 2012) was used to assign taxonomy to the representative sequence for each OTU using blast at 80% sequence match at the lowest possible taxonomic level with the SILVA small subunit (SSU) reference database (Quast et al., 2013; Yilmaz et al., 2014).

The 18S small subunit primers selected for this rRNA gene study were designed to amplify across a wide diversity of eukaryotes (Machida and Knowlton, 2012); however, because the performance of MPN to grow phytoplankton was being measured, which as applied is based on scoring of changes in chlorophyll auto-fluorescence, the identified taxa were screened for phytoplankton taxa (i.e., presence or absence of chlorophyll; Adl et al., 2012). For a small number of taxa, not enough taxonomic information was available to make a definitive assessment; these were categorized as having chlorophyll but were sub-categorized as having chlorophyll with high probability (some stramenopile and one cryptophyte taxa, both groups for which the majority of taxa are photosynthetic, and some alveolate and dinoflagellate taxa, of which approximately half the taxa are photosynthetic). Following

taxonomic assignment, OTUs assigned to taxa other than phytoplankton (taxa with no chlorophyll, unassigned reads, and taxa of chlorophyll-containing land plants (Embryophyta within Archaeplastida) that are probably present only as debris in near shore waters) were removed from further analyses.

To assess whether the sequence read numbers approached full taxonomic characterization of the samples, rarefaction curves were created using PAST software showing the relationship between the number of sequence reads and the number of recovered taxa for the BC and all replicates of the AC samples. The goal of rarefaction is to determine if the relationship approaches an asymptote to ensure sufficient read-depth has been achieved.

2.5. Quantitative real-time PCR (qRT-PCR)

Sequence read number data obtained from NGS are only semi-quantitative and cannot be reliably compared across samples. To make later growth determinations, NGS sequence read numbers from BC and AC samples need to be normalized via qRT-PCR. The PCR template concentration was determined for all DNA samples by amplifying the V9 region of the 18S rRNA gene (using the same primers as for the NGS library) using SYBR Green qRT-PCR. The QuantStudio® 12K Flex Real-Time PCR System (Applied Biosystems USA) was used for qRT-PCR, with a volume of 20 µL and the Power SYBR Green master mix (Warrington, UK). The reaction mixture consisted of 10 µL of 2 x SYBR master mix solution, 0.5 µL of 10 mM concentration each forward and reverse primer, 1 µL of DNA template and 8 µL milliQ water. The reaction protocol consisted of 95 °C for 10 min followed by 40 cycles of 95 °C for 15 s and annealing at 60 °C for 1 min.

The PCR efficiency of the V9 region of the 18S rRNA qRT-PCR was estimated using the program LinRegPCR (Ramakers et al., 2003), and amplicon efficiency, threshold and Cq values were obtained and used to calculate theoretical starting DNA concentrations (N_0) per technical replicate in LinRegPCR (Ramakers et al., 2003) using the unbiased method of Tuomi et al. (2010). Technical replicates were averaged for each sample.

2.6. Growth calculations

To determine whether a taxon demonstrated growth during the MPN incubation, sequence read concentrations were determined and compared for BC and AC samples. For each sample j , the sequence read number for each taxa i ($NS_{i,j}$) was normalized using the total number of 18S V9 PCR template copies from qRT-PCR ($N_{0,j}$), and the volume of original sample processed for DNA extraction (V_j) to calculate the concentration of 18S V9 PCR template copies in each sample ($C_{i,j}$):

$$C_{i,j} = \frac{1}{V_j} \times \frac{NS_{i,j}}{\sum_{i=1}^m NS_{i,j}} \times (N_{0,j})$$

The relevant volumes were 80 mL for the BC sample and 0.037 mL for each AC sample ($(5 \times 0.01 + 5 \times 0.001 + 5 \times 0.0001) \times 20 / 30 = 0.037$).

The taxon-specific growth ratio in each AC sample ($R_{i,ACj}$) was calculated by dividing the AC copy concentration by the BC copy concentration:

$$R_{i,ACj} = \frac{C_{i,ACj}}{C_{i,BC}}$$

The mean R_i for each condition was calculated by averaging non-zero $R_{i,ACj}$ values for replicate samples of each condition. Taxa were scored as growers for $R_i > 1$ and as non-growers for $R_i \leq 1$. In some

cases, a taxon was detected AC but not BC ($C_{i,BC} = 0$), these were also scored as growers but sub-classified as new-growers. Finally, in some cases, a taxon was detected BC but not AC ($C_{i,AC} = 0$), these were scored as non-growers but sub-classified as potential non-growers, because a lack of AC detection provides no numerical proof of non-grower status and could be the result of detection limitations. The growth scoring for each taxa at each condition was expressed as a dimensionless binary factor g_i , set to 1 for growers and new growers, and to 0 for non-growers and potential non-growers. When considering all conditions combined, g_i was set to 1 if a taxon had demonstrated growth in any of the conditions.

For each condition and all conditions combined, taxa-based growability (GT) was calculated as the percentage of taxa observed to grow:

$$GT = \frac{\sum_{i=1}^m g_i(i)}{m} \times 100$$

where m is the total number of taxa observed in all AC and BC samples. The community-based growability (GC) was calculated as the BC-abundance weighted percentage of species observed to grow:

$$GC = \frac{\sum_{i=1}^m g_i(i) \times C_{i,BC}}{\sum_{i=1}^m C_{i,BC}} \times 100$$

2.7. Time series qRT-PCR of dead cells

It is known that eDNA can include DNA from dead cells (e.g., Cangelosi and Meschke, 2014; Carini et al., 2016), raising the possibility that AC detection may be of dead cell DNA present in the original sample. A time series qRT-PCR test was conducted on heat-killed phytoplankton to determine the persistence of dead cell DNA. A one liter water sample was collected from the Detroit River (Windsor, Ontario, Canada) and incubated for one week to amplify the number of phytoplankton for this test, after which 10 aliquots of 50 mL were prepared. All the aliquots were heat treated at 75 °C for 15 min to kill the native phytoplankton. Two 50-mL aliquots were collected on each of days 0, 1, 5, 11 and 14, and were filtered and eDNA extracted for qRT-PCR of the V9 region of the 18S rRNA, following the same protocol as described for qRT-PCR above. To confirm the heat-treated cells were dead, 9 replicates of 0.2 mL from the untreated sample and 9 replicates of 0.2 mL from each 50-mL heat-treated aliquot were placed in 96-well plates and incubated for 13 days at 20 °C. The plates were monitored for phytoplankton growth via chlorophyll auto-fluorescence using a PerkinElmer Victor 3V plate reader with a 430/8 nm optical excitation filter and a 680/10 nm optical emission filter.

3. Results

3.1. Metabarcoding

After excluding the poor quality sequences, 913,754 high-quality sequences were obtained. After assigning taxonomy to OTU representative sequences using the SILVA small subunit database (Yilmaz et al., 2014), 299 unique accession numbers were assigned as taxa. Of these, 86 taxa (29% of taxa, 68% of sequences) were classified as having chlorophyll for certain, 24 taxa (8% of taxa, negligible % of sequences) with high probability, 8 taxa (3% of taxa, negligible % of sequences) as certain but land plants, 180 taxa (60% of taxa, 24% of sequences) as not having chlorophyll, and one taxon (5 sequences) was unassigned (Table 1a). For the purpose of

measuring the growth performance of chlorophyll-scored MPNs for phytoplankton, the analysis was restricted to the 110 taxa with certain or high probability of containing chlorophyll, but not the land plants.

Among samples, the numbers of sequences were fairly evenly distributed, with the exception of the K10 samples (Table 2). The combined total sequence counts for the three replicate AC samples for each incubation condition (except K10) were similar and approximately twice the numbers in the BC sample. Although qRT-PCR normalization had not yet been applied, this signified significant growth occurred in these samples as each AC sample contained 1/2,162 the volume of the BC sample. Within each sample, sequences associated with chlorophyll-containing phytoplankton represented a high proportion of the total sample sequences (Table 2). The rarefaction plots showed moderately convergent curves for all but the replicates of the K10 sample (Fig. 1). Those convergent curves indicate that sufficient sequence coverage was achieved to characterize community diversity, and that further sequencing would not likely increase the numbers of taxa substantially.

For K10 samples, two replicates failed to sequence and the third was below the expected threshold for sequence read depth (Table 2), which was consistent with the observed poor qRT-PCR amplification for the same samples and a lack of phytoplankton growth by chlorophyll autofluorescence in the MPN (Table S1). The K10 samples were reported to have fungal contamination during incubation, and fungal taxa comprised 61.5% of the sequences in K10 samples, compared to $2.0 \pm 0.8\%$ in all other samples. It was concluded that the K10 samples were atypical and were excluded from further analysis.

3.2. BC community composition

A total of 259 taxa were identified in the BC sample (Table 1b), of which 97 were classified as putative phytoplankton taxa. While the phytoplankton represented a minority of the detected taxa (37%), they represented a majority of the sequences (84%). Among the phytoplankton taxa, the most abundant groups were Stramenopiles (59% of phytoplankton taxa, 67% of phytoplankton sequences), Alveolates (21% of taxa, 33% of sequences) and Archaeplastida (18% of taxa, 1% of sequences). The major phytoplankton taxa within the Stramenopiles were diatoms and secondarily chrysophytes, within the Alveolates were dinoflagellates, and within the Archaeplastida were green algae. Although this analysis was based on phytoplankton, the community composition of the non-phytoplankton taxa is also reported here. The most abundant non-phytoplankton groups were Opisthokonts (36% of non-phytoplankton taxa, 36% of non-phytoplankton sequences), Alveolates (19% of taxa, 30% of sequences), Rhizaria (15% of taxa, 12% of sequences) and Stramenopiles (14% of taxa, 19% of sequences). The major non-phytoplankton taxa within the Opisthokonts were fungi and various flagellated protists, within the Rhizaria were cercozoans (amoeboids), within the Alveolates were heterotrophic dinoflagellates, and within the Stramenopiles were filamentous protists (Labyrinthulomycetes) and non-photosynthetic chrysophytes. Among all the taxa, the most abundant were a diatom (*Thalassiosira* spp., 45% of sequences) and a photosynthetic dinoflagellate (*Karlodinium* spp., 26% of sequences), and 5 diatoms, 2 heterotrophic dinoflagellates, 1 metazoan, 1 Labyrinthulomycetes and 1 fungus that each comprised 1–4% of sequences. Together these 12 taxa comprised 90% of the total sequences in the BC sample.

3.3. Phytoplankton growth assessments

The mean growth ratio R_i was calculated for each taxa in each

Table 1

Numbers of taxa and numbers of sequences, classified by taxonomic group and by presence or absence of chlorophyll (CHL). A) all samples; B) before-culture sample only. Some OTU sequences did not have enough taxonomic discrimination to make a definitive assessment for the photosynthetic status of the taxon (i.e., presence of chlorophyll), so were categorised as having chlorophyll with high probability (CHL-HP, stramenopiles, chryptophytes, alveolates and dinoflagellates). Phytoplankton were defined as taxa with any probability of having chlorophyll, but excluding land plants (within Archaeplastida).

| A | All Samples, Number of Taxa | | | | All Samples, Number of Sequences | | | |
|----------------|--------------------------------|-----------|------------|----------|-------------------------------------|--------------|----------------|-----------|
| | CHL | CHL-HP | No CHL | Exclude | CHL | CHL-HP | No CHL | Exclude |
| Archaeplastida | 19 | 0 | 0 | 8 | 14,595 | 0 | 0 | 77 |
| Haptophyta | 1 | 0 | 0 | 0 | 107 | 0 | 0 | 0 |
| Stramenopile | 59 | 5 | 27 | 0 | 603,358 | 140 | 72,876 | 0 |
| Alveolate | 4 | 18 | 41 | 0 | 71,570 | 4,330 | 83,608 | 0 |
| Cryptophyceae | 3 | 1 | 1 | 0 | 95 | 7 | 106 | 0 |
| Excavata | 0 | 0 | 11 | 0 | 0 | 0 | 565 | 0 |
| Rhizaria | 0 | 0 | 25 | 0 | 0 | 0 | 4,398 | 0 |
| Amoebozoa | 0 | 0 | 8 | 0 | 0 | 0 | 196 | 0 |
| Centrohelida | 0 | 0 | 2 | 0 | 0 | 0 | 271 | 0 |
| Opisthokonta | 0 | 0 | 65 | 0 | 0 | 0 | 57,450 | 0 |
| Unassigned | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 5 |
| Total | 86 | 24 | 180 | 9 | 689,725 | 4,477 | 219,470 | 82 |
| B | Before Culture, Number of Taxa | | | | Before Culture, Number of Sequences | | | |
| | CHL | CHL-HP | No CHL | Exclude | CHL | CHL-HP | No CHL | Exclude |
| Archaeplastida | 17 | 0 | 0 | 8 | 673 | 0 | 0 | 36 |
| Haptophyta | 1 | 0 | 0 | 0 | 11 | 0 | 0 | 0 |
| Stramenopile | 53 | 4 | 22 | 0 | 72,240 | 66 | 3,826 | 0 |
| Alveolate | 4 | 16 | 30 | 0 | 33,358 | 2,292 | 6,208 | 0 |
| Cryptophyceae | 2 | 0 | 1 | 0 | 21 | 0 | 16 | 0 |
| Excavata | 0 | 0 | 10 | 0 | 0 | 0 | 473 | 0 |
| Rhizaria | 0 | 0 | 24 | 0 | 0 | 0 | 2,544 | 0 |
| Amoebozoa | 0 | 0 | 7 | 0 | 0 | 0 | 55 | 0 |
| Centrohelida | 0 | 0 | 2 | 0 | 0 | 0 | 50 | 0 |
| Opisthokonta | 0 | 0 | 57 | 0 | 0 | 0 | 7,407 | 0 |
| Unassigned | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 5 |
| Total | 77 | 20 | 153 | 9 | 106,303 | 2,358 | 20,579 | 41 |

Table 2

Number of sequences obtained for before-culture and after-culture samples with different incubation conditions, for all taxa and for the subset of phytoplankton taxa (only those with chlorophyll excluding land plants).

| Incubation Conditions | | Number of Sequences | |
|-----------------------|------------------|---------------------|---------|
| Enrichment Media | Temperature (°C) | With Chlorophyll | Total |
| Before Culture | | 108,661 | 129,281 |
| Keller's | 20 | 85,966 | 123,066 |
| | | 55,309 | 61,621 |
| | | 63,189 | 78,169 |
| Keller's | 10 | 4 | 4 |
| | | 0 | 2 |
| | | 1,869 | 8,032 |
| Guillard's f/2 | 20 | 46,369 | 76,078 |
| | | 51,075 | 80,479 |
| | | 45,450 | 85,824 |
| Guillard's f/2 | 10 | 111,791 | 127,445 |
| | | 80,145 | 93,040 |
| | | 44,374 | 50,713 |

condition, to determine which phytoplankton grew from BC to AC ($R_i > 1$). Each taxa was classified as either a potential non-grower ($R_i = 0$), a non-grower ($0 < R_i \leq 1$), a grower ($R_i > 1$), or a new grower (R_i undefined). Non-growers were distinguished from potential non-growers because the former had numerical confirmation of decline while the latter had zero AC detections, which could alternatively be an artefact of method resolution. For individual conditions, there were 10–12 potential non-growers, of which only 2 were potential non-growers when results from all conditions were combined (Fig. 2). The two taxa not detected in the AC samples both had only 3 sequences in the BC sample, so there is a high probability that, being very rare, those taxa were not transferred for

MPN incubation. No non-growers were detected in any sample. The remainder of the 97 taxa detected BC were growers. In addition to these growers, there were 13 taxa determined to be new-growers, taxa for which R_i could not be calculated because BC detections were zero (Fig. 2).

Growth was definitive, with mean R_i values (within a condition) ranging from 4.0×10^1 – 2.6×10^5 across taxa (Fig. 3a). For each condition, mean R_i values for growers had a log-normal distribution with $10^3 < R_i < 10^4$ for most taxa (Fig. 3a). In addition, the mean R_i values showed internal consistency in the ability to detect growth (Fig. 3b), as for each taxon mean R_i values were well correlated between conditions (K20-G20: $r = 0.75$; G20-G10: $r = 0.74$; G10-

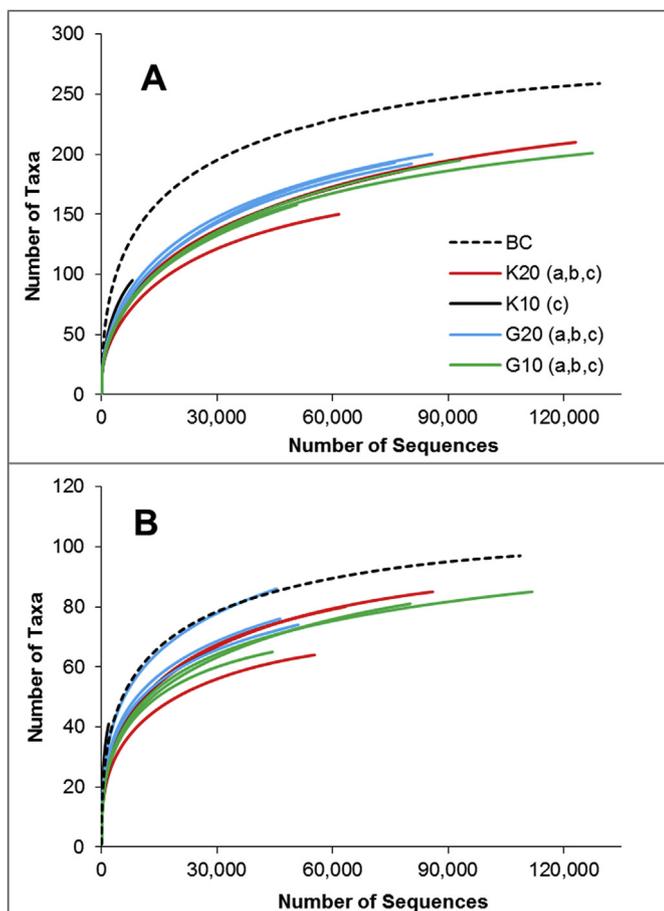


Fig. 1. Rarefaction curves for sequence read counts generated using next generation sequencing data for before-culture (BC) and after-culture (AC) samples. A) All taxa (with and without chlorophyll). B) Phytoplankton taxa with chlorophyll. K20 and K10 refer to incubations with Keller's enrichment media at 20 °C and 10 °C respectively, and G20 and G10 refer to incubations with Guillard's *f/2* enrichment media at 20 °C and 10 °C respectively. Replicates (a, b, c) for each incubation condition are shown as separate curves with the same colour. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

K20: $r = 0.81$).

3.4. Effect of incubation conditions on phytoplankton taxa detections and growth

There were no statistical differences in numbers of taxa that grew among the incubation conditions ($\chi^2 = 0.04$; $P > 0.01$), however there were differences in which taxa grew among incubation conditions (Fig. 4). For the AC taxa, 84 were common to all three incubation conditions, 15 were common to two conditions, and 9 were detected in only one condition. However, all taxa detected in only one AC condition had 3 or fewer sequences (note: all singletons and doubletons across all samples were excluded, but rare sequences within samples were retained in the analysis). This may indicate that differences in detected taxa among incubation conditions may relate to sequence read depth differences, rather than differences in growth.

3.5. eDNA degradation and detection

The use of eDNA and metabarcoding to create BC and AC detection lists can lead to artefacts in growth calculations due to the measurement of DNA from dead cells persisting in the AC samples.

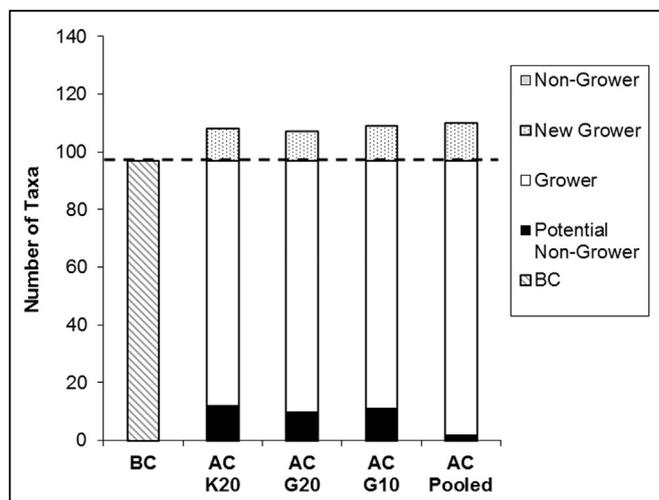


Fig. 2. Numbers of phytoplankton taxa in before-culture (BC) and after-culture (AC) samples, according to growth scoring categories. Growth ratios were calculated as the AC/BC 18S template concentrations for a taxon. Growers had growth ratio >1 , new-growers were only detected AC, and potential non-growers were only detected BC. No non-growers (growth ratio ≤ 1) were detected. K20 refers to incubation with Keller's enrichment media at 20 °C, and G20 and G10 refer to incubations with Guillard's *f/2* enrichment media at 20 °C and 10 °C respectively. Pooled indicates all AC samples combined.

To quantify the potential for this artefact in the analyses, a time-series qRT-PCR was conducted with heat-treated phytoplankton. Chlorophyll autofluorescence measurements in untreated and heat-treated samples over time confirmed that the heat-treated cells were dead or non-viable (Fig. S1). The qRT-PCR results showed an exponential decline in 18S V9 template concentration over time, with $>95\%$ DNA degraded after 11 days and $>99\%$ DNA degraded after 13 days (Fig. 5). With 1% dead cell DNA persistence in a 14 day MPN incubation and a 2,162-fold difference in BC and AC sample volumes, there would need to be $> 216,000$ dead cell DNA sequences in a BC sample to detect 1 sequence AC, much higher than the sequence counts for any taxa in the BC sample. As a result, dead cell DNA carry-over is not likely affecting the growth assessments.

3.6. Growability in MPN assays

Taxon-specific growability (GT) was calculated, which included rare taxa only detected AC after increasing through the incubation (new-growers). GT values were high for each condition (87–88%) and for all conditions combined (98%, Table 3). Community-specific growability (GC) was also calculated, which did not include new-growers because they did not contribute to the BC community weighting. The BC sequence counts were used to determine the relative BC abundance for each taxon, albeit with caution. Because the most abundant taxa were determined to be growers, and conversely because the taxa not determined to be growers were rare, GC values were very high for all conditions (99.94–99.97%) and for all conditions combined (99.99%). Though the calculations do not warrant four significant figures, they are necessary to show that values were not 100% (some rare taxa scored as potential non-growers). The phytoplankton community was dominated by two taxa, a diatom (*Thalassiosira* spp.) and a dinoflagellate (*Karlodinium* spp.) with 53% and 31% of the abundance respectively. The next six most abundant phytoplankton taxa each contributed 1–3% of the abundance of the community. Together these 8 most abundant phytoplankters comprised 95% of the phytoplankton community,

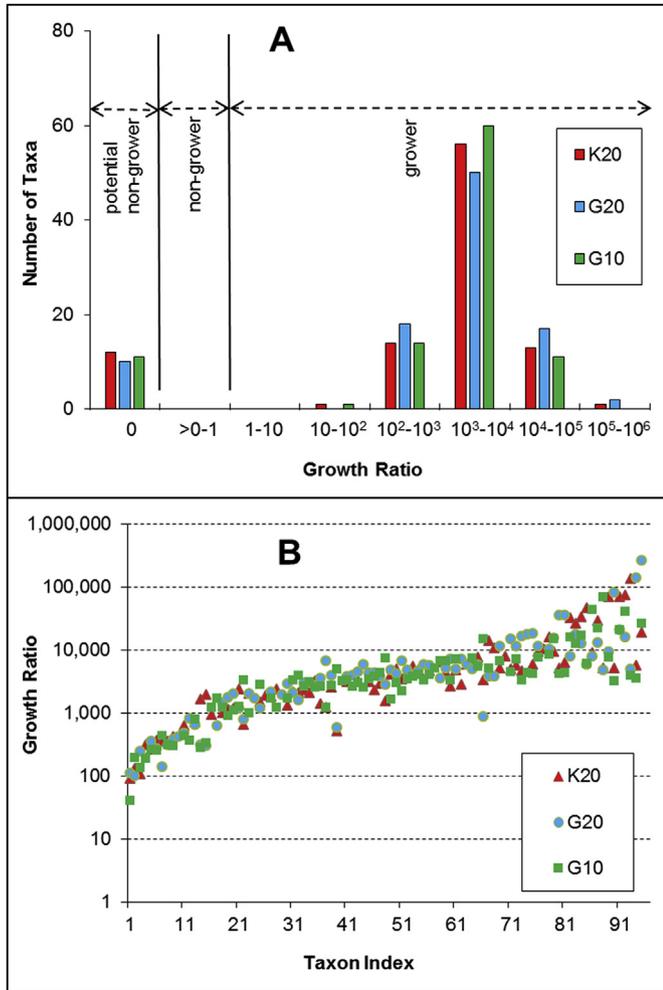


Fig. 3. A) Frequency distribution of growth ratios for phytoplankton taxa in different incubation conditions. Growth ratios were calculated as the AC/BC 18S template concentrations for a taxon. Note that potential non-growers were found (zero AC sequence reads) but no confirmed non-growers were found. Growth ratios for new-growers, taxa only detected AC, cannot be calculated. B) Growth ratios for the grower taxa in different incubation conditions. The x-axis categories are for each grower taxon sorted by magnitude of growth ratio. K20 refers to incubation with Keller's enrichment media at 20 °C, and G20 and G10 refer to incubations with Guillard's f/2 enrichment media at 20 °C and 10 °C respectively.

and all were growers. The remaining phytoplankters each had <1% of the phytoplankton abundance, and of these, only 2 were not growers.

4. Discussion

4.1. Measurements of growability

In the field of ballast water treatment, there is a need for information about the reliability of culture-based methods for enumerating viable organisms. Cullen (2018) developed a framework for validating organism status methods, identifying Detection Factors as key method performance factors because their complement, false negatives, correspond to invasion risk from living or viable organisms undetected in an assay of a treatment. For MPN assays, the Detection Factor is the property of growability, which can be determined by creating BC and AC taxa lists. Data generation using microscopic assessment has been constrained by logistic limitations due to the requirement for skilled taxonomists,

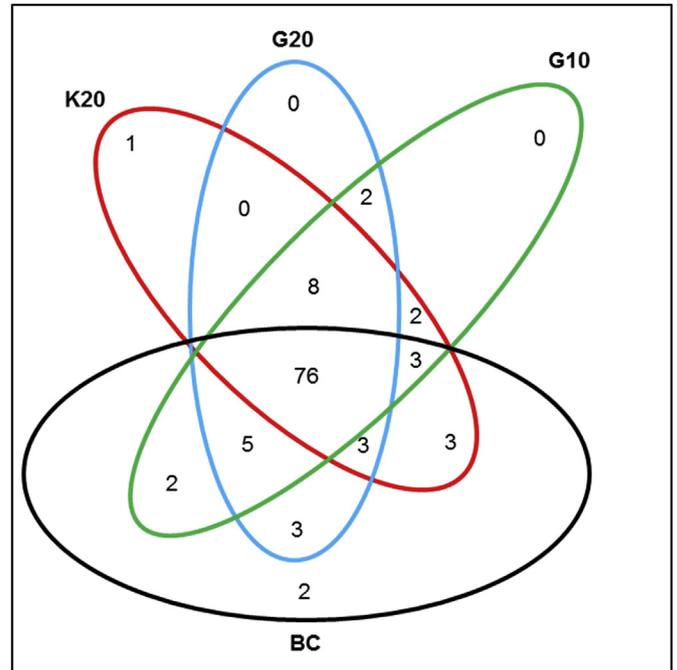


Fig. 4. Venn diagram showing the distribution of detected phytoplankton taxa based on metabarcoding of 18S V9 DNA sequencing, before-culture (BC) and after-culture with different incubation conditions. K20 refers to incubation with Keller's enrichment media at 20 °C, and G20 and G10 refer to incubations with Guillard's f/2 enrichment media at 20 °C and 10 °C respectively.

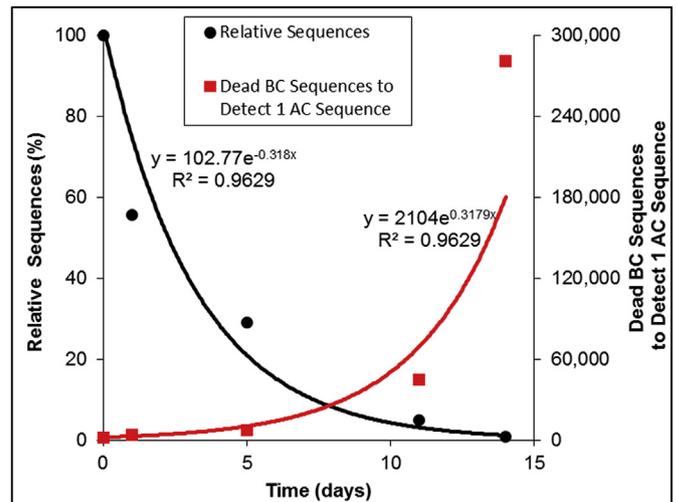


Fig. 5. Persistence of environmental DNA after heat-killing cultures shown as number of 18S V9 template copies measured by qRT-PCR, relative to untreated numbers at time zero (circles). Squares (right Y-axis) show the calculated number of dead BC sequences required to detect 1 AC sequence, based on the relative remaining dead cell DNA at any time and the original sample volumes used in BC and AC samples.

counting limits and sample preservation trade-offs (preservatives can destroy some taxa, but the examination of fresh samples is not always possible).

In this work a method for measuring growability of eukaryotic phytoplankton was developed that can alleviate these constraints using modern molecular genetic tools: metabarcoding for taxonomic assessments and qRT-PCR to quantitatively compare AC and BC samples for accurate growth determinations. The advent of NGS platforms that are reliable and affordable (Shendure and Ji, 2008)

Table 3
Calculated values of taxa-based growability (percent of taxa growing) and community-based growability (before-culture abundance weighted percent of taxa growing) for different incubation conditions and for all conditions combined.

| Incubation Conditions | | Growability (%) | |
|-----------------------|------------------|-----------------|-----------|
| Enrichment Media | Temperature (°C) | Taxa | Community |
| Keller's | 20 | 87 | 99.96 |
| Guillard's f/2 | 20 | 88 | 99.94 |
| Guillard's f/2 | 10 | 88 | 99.97 |
| All Combined | | 98 | 99.99 |

has led to the use of metabarcoding to assess species composition in many fields (National Research Council, 2007) and for diverse biological communities using eDNA (Neelakanta and Sultana, 2013), revealing remarkable diversity. Metabarcoding was used here for eukaryotic phytoplankton, to assess conventional MPN assays scored by chlorophyll autofluorescence, though this method could be applied to all protists with appropriate PCR primers. This method offers improvements over microscopic taxonomic assessment by alleviating the logistic limits noted above, based on the fact that DNA preservatives allow for reliable sample archiving for later assessment. In addition, the taxonomic identification is based on 18S sequences, hence results are comparable across sample times and sites by using common metabarcoding databases and assignment protocols.

There is considerable controversy over the culturability of phytoplankton (among other microorganisms), with some stating that most taxa cannot be cultured (First and Drake, 2013) but with MPN practitioners reporting that many taxa do grow (Madsen and Petersen, 2015). The paradox is resolved by understanding that growability refers to a short-term growth requirement that can be broadly successful, and culturability refers to a more rigorous requirement of isolation and perpetual maintenance (Cullen, 2018; Cullen and MacIntyre, 2016). Growability is the relevant property to determine growth bias in MPN assays, and measurements of growability in this work were high (growth bias was low). Taxa-based growability was 87–88% for individual conditions and 98% for multiple conditions combined. These high growability values are consistent with the concept that every sample has by design the “perfect” media because it is prepared with filter-sterilized water from the same source, and should thus have all required compounds for survival and short-term growth. This is in keeping with the explanations of high expected growability by Cullen and MacIntyre (2016), the observations of high growability but not culturability in similar applications by Thronsdon (1978), and the reports of high growability in these exact applications (Madsen and Petersen, 2015). These results are also consistent with work in other fields: Kaeberlein et al. (2002) showed that the growability of marine bacteria could be greatly improved by incubating with sample water rather than defined media, and metabarcoding of eDNA has been widely used to characterise the AC composition of complex microorganism communities that were previously considered essentially unculturable (Browne et al., 2016; Lagier et al., 2016).

It follows that different enrichment media should not impact growability, which is what was found in this work. Similarly, growability was not impacted by different incubation temperatures, also as expected. This simplifies the standardization of MPN assays, where the key features to standardize are the use of filter-sterilized sample water as the base for enrichment media and incubation in a near-ambient temperature range; in fact practices at BWMS test facilities are standardized based on these features (IMO, 2016). Differences in patterns of growing taxa among incubation conditions were likely due to heterogeneous sampling of rare taxa, which is enabled by the high resolution of the NGS method. In fact,

for the 9 cases of taxa detected in only one AC condition (Fig. 3), 5 cases had a single sequence read and all cases had 3 or fewer AC sequence reads. The growability of taxa under all incubation conditions combined provides strong physiological evidence of the ability for those taxa to grow in some incubation condition as opposed to being unculturable.

The high community-based growability values are a reflection that all of the abundant taxa grew. Community-based results are the relevant ones for evaluating methods used in ballast water treatment regulations, because the regulations are not based on taxa but on numbers of individuals within size classes. Community-based results are also the most reliable, because the high numbers of detected sequences provide the most reliable growth scoring calculations, and new-growers, which are rare taxa, do not contribute to the outcome. The growability of high abundance taxa is perhaps expected as these taxa are abundant because they were successfully reproducing in the environmental conditions, which were replicated in the laboratory. Finally, the similarity of growability values between incubation conditions was consistent with the enumerations produced by the MPNs which were themselves being evaluated, as the mean MPN responses between conditions were also not statistically different (Table S1, ANOVA, $p = 0.218$).

4.2. Method reliability and resolution

These analyses were designed to be robust - all singletons and doubletons were removed and strict sequence quality filtering was applied, thus all of the sequences retained in the analyses were meaningful. The rarefaction curves showed good sequence read depth for the characterization of the communities, consistent with the detections of rare taxa. DNA decay rates were measured in heat-killed cells which ruled out the possibility of artefacts in AC detections originating from dead cells in the BC sample. The use of qRT-PCR allowed BC and AC detections to be compared to make quantitative growth determinations. Though the quantification of sequences at two time points (BC and AC) does not allow the calculation of true growth rates during incubation, the growth ratios (R_i) were still high and allowed definitive growth scoring. Finally, the growth ratios could be biased by changes in the copy number of the 18S rDNA genes during the 14 day incubation period; however, while 18S copy number is known to vary among eukaryotes (Prokopowich et al., 2003), there is no expectation that the mean copy number would vary within a taxa over time, especially over an incubation period of 14 days.

For some taxa there was uncertainty in growth scoring: when a taxon was detected BC but not AC, it was subcategorized as a potential non-grower; and when a taxon was detected AC but not BC, it was classified as a new-grower with unknown growth ratio. Resolution can be improved by obtaining more sequences for the analysis, either by applying more sequencing effort, or by increasing the volume of original sample processed. Increasing the number of AC sequence reads will help resolve taxa categorised as potential non-growers (zero AC detections) into definitive grower or non-grower categories. Increasing the number of BC sequence

reads will help resolve new-growers (zero BC detections) into growers with a defined R_i . Assuming the same qRT-PCR values across samples, the resolution is determined by the ratio of volumes of original sample processed in BC and AC samples, V_{AC}/V_{BC} . The nature of MPN arrays leads to a low V_{AC}/V_{BC} ratio; in this work 80 mL of BC sample were processed, and though AC samples were 60 mL composites of 5 replicates at each of 3 dilutions, the amount of original sample was only 0.037 mL, resulting in $V_{AC}/V_{BC} = 1/2,162$. Thus taxa in potential non-grower and new-grower categories had $R_i < 2,162$ and $R_i > 2,162$, respectively.

It is unfortunate that the experimental design suffered from the probable contamination of K10 samples, but it was useful as a negative control for growth detection, showing that the method would generate appropriate negative growth outcomes. This condition was excluded from comparative analyses because it was atypical and would give false causality for growability differences from different incubation conditions. Even if the K10 condition were the cause of the lack of growth, it would not negate the results for growable taxa across all media combined. But it was unlikely to be the cause, as broad community growth was measured both for samples with the same enrichment media at a different temperature and for samples at the same temperature with a different enrichment media, and the likelihood that some interaction effect resulted in a lack of growth is low. These types of contamination events are rare in MPN assays.

4.3. Implications for method validation

For validating methods used to measure treatment performance (e.g. MPN and SM methods for measuring BWMS treatments), one of the most important metrics is the method Detection Factor, the complement of the false negative rate, as it is measurement bias that represents an invasion risk from living organisms not detected by an assay (Cullen, 2018). The method developed in this work was used to make Detection Factor estimates (growability) for validating phytoplankton MPN assays. Even though measured growability values were high, it is important to recognize that they are “worst-case” values. For phytoplankton in natural samples, their status is not known *a priori*. This method classifies taxa as non-growers, growers and new-growers with certainty, because of measured increases or decreases in PCR template concentrations. However, the distinction of potential non-growers is an important one, because, although they are counted as non-growers their true status is unknown. The absence of AC detections could be because the organisms were alive but not capable of growth, or they were dead in the original sample, or growth may have been undetected with the method resolution. For these three possibilities, classifying taxa as non-growers would give accurate, worst case, and worst case growability values, respectively.

This work was compared to the only other published validation of a BWMS assay for organisms in the 10–50 μm size category. Steinberg et al. (2011) reported low false negative rates of 1–2% for the SM method, however, for the complementary Detection Factors (98–99%) to be comparable to the Detection Factor values in this work (growability), the same classifying criteria must be applied. Steinberg et al. (2011) assumed that organisms not staining and not motile were dead and eliminated them from the counts, however their status was unknown and they could have been alive. For example, live but non-staining and non-motile diatoms such as *Amphiprora* spp. (MacIntyre and Cullen, 2016) would fall into this category and be mis-classified. These organisms are potential false negatives in the SM assay, analogous to organisms in the potential non-grower category in this work which were ultimately counted as non-growers. When adjusted to be comparable, the SM Detection Factor values for Steinberg et al. (2011) decline to 74–90%.

Higher Detection Factors (lower bias) for MPN than SM is consistent with the results of Molina et al. (2016) who reported higher concentrations of live cells measured by MPN than SM in all comparative samples, and with other comparative data (Blatchley et al., 2018; Cullen, 2018; Wright and Welschmeyer, 2015) where higher concentrations by MPN than SM were reported in a majority of samples.

5. Conclusion

A reliable method for measuring growability was developed and used to measure the performance of phytoplankton MPN assays used in BWMS testing. The high (>99%) community growability measurements verify that most phytoplankton grow to detection in MPN assays, even if many have not yet been brought into permanent culture. This in turn verifies that MPN assays produce accurate and unbiased results. The consistently high growability values for different enrichments and temperatures supports the conclusion that successful growth occurs because the media base is filter-sterilized sample water containing all necessary compounds, which provides a simple and elegant basis to standardize MPN assays. This work can provide regulatory confidence for broader acceptance of MPN assays without limitations, which will harmonize the sizing of UV-based BWMSs globally.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We wish to acknowledge the financial contributions from MITACs Elevate, MITACs Accelerate program and Trojan Technologies for PDF funding to Dr. Chaganti. We thank NSERC for providing operating grants in the form of NSERC Engage and NSERC Engage Plus grants as well as an NSERC Discovery Grant to DDH. We thank Dr. Gitte Petersen for providing split samples for this work, and Dr. Alastair Simpson for reviewing our chlorophyll assessments for protists. The comments of two anonymous reviewers and the editor helped to greatly improve the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.watres.2019.114941>.

References

- Adl, S.M., Simpson, A.G.B., Lane, C.E., Lukeš, J., Bass, D., Bowser, S.S., Brown, M.W., Burki, F., Dunthorn, M., Hampl, V., Heiss, A., Hoppenrath, M., Lara, E., Le Gall, L., Lynn, D.H., McManus, H., Mitchell, E.A., Mozley-Stanridge, S.E., Parfrey, L.W., Pawłowski, J., Rueckert, S., Shadwick, L., Schoch, C.L., Smirnov, A., Spiegel, F.W., 2012. The revised classification of eukaryotes. *J. Eukaryot. Microbiol.* 59, 429–493.
- Blatchley III, E.R., Cullen, J.J., Petri, B., Bircher, K., Welschmeyer, N., 2018. The biological basis for ballast water performance standards: “Viable/Non-Viable” or “Live/Dead”? *Environ. Sci. Technol.* 52 (15), 8075–8086.
- Browne, H.P., Forster, S.C., Anonye, B.O., Kumar, N., Neville, B.A., Stares, M.D., Goulding, D., Lawley, T.D., 2016. Culturing of ‘unculturable’ human microbiota reveals novel taxa and extensive sporulation. *Nature* 533, 543–546.
- Cangelosi, G.A., Meschke, J.S., 2014. Dead or alive: molecular assessment of microbial viability. *Appl. Environ. Microbiol.* 80, 5884–5891.
- Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-lyons, D., Huntley, J., Fierer, N., Owens, S.M., Betley, J., Fraser, L., Bauer, M., Gormley, N., Gilbert, J.A., Smith, G., Knight, R., 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* 6, 1621–1624.
- Carini, P., Marsden, P.J., Leff, J.W., Morgan, E.E., Strickland, M.S., Fierer, N., 2016. Relic

- DNA is abundant in soil and obscures estimates of soil microbial diversity. *Nat. Microbiol.* 2, 16242.
- Chaganti, S.R., Lalman, J.A., Heath, D.D., 2012. 16S rRNA gene based analysis of the microbial diversity and hydrogen production in three mixed anaerobic cultures. *Int. J. Hydrogen Energy* 37, 9002–9017.
- Coast Guard, Ballast water - living vs. viable, 2015. Coast Guard Maritime Commons. <http://mariners.coastguard.dodlive.mil/2015/12/07/1272015-ballast-water-living-vs-viable/> (viewed on Feb 03, 2019).
- Coast Guard, 2016. Final action on ballast water management system appeals, *Coast Guard Maritime Commons*. <http://mariners.coastguard.dodlive.mil/2016/07/12/7122016-final-action-on-ballast-water-management-system-appeals/> (viewed on Feb 03, 2019).
- Cullen, J.J., 2018. Quantitative framework for validating two methodologies that are used to enumerate viable organisms for type approval of ballast water management systems. *Sci. Total Environ.* 627, 1602–1626.
- Cullen, J.J., MacIntyre, H.L., 2016. On the use of the serial dilution culture method to enumerate viable phytoplankton in natural communities of plankton subjected to ballast water treatment. *J. Appl. Phycol.* 28, 279–298.
- Edgar, R.C., 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* 10, 996–998.
- Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., Knight, R., 2011. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27 (16), 2194–2200.
- Emerson, J.B., Adams, R.L., Roman, C.M.B., Brooks, B., Coil, D.A., Dahlhausen, K., Ganz, H.H., Hartmann, E.M., Hsu, T., Justice, N.B., Paulino-Lima, I.G., Luongo, J.C., Lympelopoulou, D.S., Gomez-Silvan, C., Rothschild-Mancinelli, B., Balk, M., Huttenhower, C., Nocker, A., Vaishampayan, P., Rothschild, L.J., 2017. Schrodinger's microbes: tools for distinguishing the living from the dead in microbial ecosystems. *Microbiome* 5, 86. <http://doi.org/10.1186/s40168-017-0285-3>.
- ETV, 2010. Generic Protocol for the Verification of Ballast Water Treatment Technology. NSF International for USEPA Environmental Technology Verification Program, Ann Arbor, MI.
- First, M.R., Drake, L., 2013. Approaches for determining the effects of UV radiation on microorganisms in ballast water. *Manag. Biol. Invasions* 4, 87–99.
- Guillard, R.R.L., 1975. Culture of phytoplankton for feeding marine invertebrates. pp 26–60. In: Smith, W.L., Chanley, M.H. (Eds.), *Culture of Marine Invertebrate Animals*. Plenum Press, New York, USA.
- Guillard, R.R.L., Ryther, J.H., 1962. Studies of marine planktonic diatoms. I. *Cyclotella nana* Hustedt and *Detonula confervacea* Cleve. *Can. J. Microbiol.* 8, 229–239.
- IMO, 2016. Review of the Guidelines for Approval of Ballast Water Management Systems (G8): Analysis Methods for Determining the Viability of Organisms in the 10 to 50 Mm Size Class. Originator: Denmark and Norway. IMO PPR 4/7, 12 October 2016, IMO PPR 4/7, 12 October 2016.
- IMO, 2017. International Convention for the Control and Management of Ships' Ballast Water and Sediments, 2004: Guidance on Methodologies that May Be Used for Enumerating Viable Organisms for Type Approval of Ballast Water Management Systems. BWM.2-Circ.61, 26 July 2017.
- Jarvis, B., Wilrich, C., Wilrich, P.T., 2010. Reconsideration of the derivation of Most Probable Numbers, their standard deviations, confidence bounds and rarity values. *J. Appl. Microbiol.* 109, 1660–1667.
- Kaerberlein, T., Lewis, K., Epstein, S.S., 2002. Isolating "uncultivable" microorganisms in pure culture in a simulated natural environment. *Science* 296, 1127–1129.
- Keller, M.D., Selvin, R.C., Claus, W., Guillard, R.R.L., 1987. Media for the culture of oceanic ultraphytoplankton. *J. Phycol.* 23, 633–638.
- Lagier, J.C., Khelafifa, S., Alou, M.T., Ndongo, S., Dione, N., Hugon, P., Caputo, A., Cadoret, F., Traore, S.I., Seck, E.H., Dubourg, G., Durand, G., Mourembou, G., Guilhot, E., Togo, A., Bellali, S., Bachar, D., Cassir, N., Bittar, F., Delerce, J., Mailhe, M., Ricaboni, D., Bilen, M., Dangui Niekro, N.P., Dia Badiane, N.M., Valles, C., Mouelhi, D., Diop, K., Million, M., Musso, D., Abrahão, J., Azhar, E.I., Bibi, F., Yasir, M., Diallo, A., Sokhna, C., Djossou, F., Vitton, V., Robert, C., Rolain, J.M., La Scola, B., Fournier, P.E., Levasseur, A., Raoult, D., 2016. Culture of previously uncultured members of the human gut microbiota by culturomics. *Nat. Microbiol.* 1, 16203.
- Lundgreen, K., Holbech, H., Pedersen, K.L., Petersen, G.I., Andreasen, R.R., George, C., Drillet, G., Andersen, M., 2018. UV fluences required for compliance with ballast water discharge standards using two approved methods for algal viability assessment. *Mar. Pollut. Bull.* 135, 1090–1100.
- Machida, R.J., Knowlton, N., 2012. PCR primers for metazoan nuclear 18S and 28S ribosomal DNA sequences. *PLoS One* 7. <http://doi.org/10.1371/journal.pone.0046180>.
- MacIntyre, H.L., Cullen, J.J., 2016. Classification of phytoplankton cells as live or dead using the vital stains fluorescein diacetate and 5-chloromethylfluorescein diacetate. *J. Phycol.* 52, 572–589.
- Madsen, T., Petersen, G.I., 2015. MPN Dilution-Culture Method as an Alternative Method for Measuring Treatment of Organisms in the 10–50 Mm Size Category. DHI Ref: 11097400 (Viewed on Feb 03, 2019). DHI. <http://mpnballastwaterfacts.com/wp-content/uploads/2016/01/DHI-Support-final.pdf>.
- Molina, V., Riley, S.C., Robbins-Walsley, S.H., First, M.R., Drake, L., 2016. Most probable number (MPN) assay to determine concentrations of ambient organisms $\geq 10 \mu\text{m}$ and $< 50 \mu\text{m}$ in oligotrophic waters. In: 19th International Conference on Aquatic Invasive Species, Winnipeg, MB Canada. http://www.icaiss.org/pdf/2016abstracts/ICAIS_Tuesday_PM_Session_C/250_Molina.pdf (Viewed on Feb 03, 2019).
- National Research Council (US) Committee on Metagenomics: Challenges and Functional, Applications, 2007. *THE NEW SCIENCE of METAGENOMICS: Revealing the Secrets of Our Microbial Planet*. National Academies Press, Washington, DC, 13: 978-0-309-10676-4.
- Neelakanta, G., Sultana, H., 2013. The use of metagenomic approaches to analyze changes in microbial communities. *Microb. Insights* 6, 37–48.
- Pawlowski, J., Lejzerowicz, F., Apotheloz-Perret-Gentil, L., Visco, J., Esling, P., 2016. Protist metabarcoding and environmental biomonitoring: time for change. *Eur. J. Protistol.* 55, 12–25.
- Prokopowich, C.D., Ryan, G.T., Crease, T.J., 2003. The correlation between rDNA copy number and genome size in eukaryotes. *Genome* 46, 48–50.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O., 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596.
- Ramakers, C., Ruijter, J.M., Lekanne, D.R.H., Moorman, A.F.M., 2003. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci. Lett.* 339, 62–66.
- Shendure, J., Ji, H., 2008. Next-generation DNA sequencing. *Nat. Biotechnol.* 26, 1135–1145.
- Steinberg, M.K., Lemieux, E.J., Drake, L.A., 2011. Determining the viability of marine protists using a combination of vital, fluorescent stains. *Mar. Biol.* 158, 1431–1437.
- Thronsdon, J., 1978. The dilution-culture method. In: Sournia, A. (Ed.), *Phytoplankton Manual*, vol. 6. UNESCO. Monographs on oceanographic methodology, pp. 218–224.
- Tuomi, J.M., Voorbraak, F., Jones, D.L., Ruijter, J.M., 2010. Bias in the Cq value observed with hydrolysis probe based quantitative PCR can be corrected with the estimated PCR efficiency value. *Methods* 50, 313–322.
- Wright, D.A., Welschmeyer, N.A., 2015. Establishing benchmarks in compliance assessment for the ballast water management convention by port state control. *J. Mar. Eng. Technol.* 14, 9–18.
- Yilmaz, P., Parfrey, L.W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., Schweer, T., Peplies, J., Ludwig, W., Glöckner, F.O., 2014. The SILVA and 'all-species Living Tree Project (LTP)' taxonomic frameworks. *Nucleic Acids Res.* 42, 643–648.